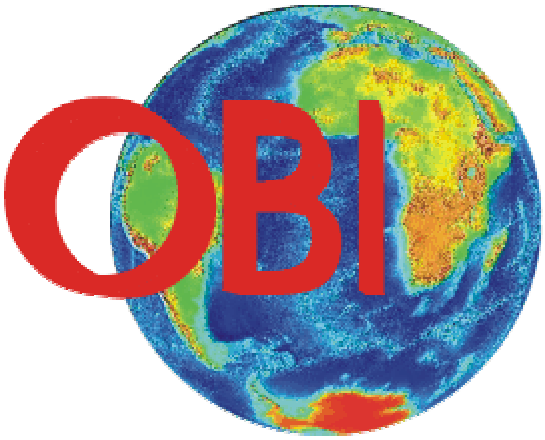
Ocean Biodiversity Informatics

Hamburg, Germany: 29 November to 1 December 2004

Book of Abstracts



Ocean Biodiversity Informatics

Hamburg, Germany: 29 November to 1 December 2004

BOOK OF ABSTRACTS

Organised by:

Intergovernmental Oceanographic Commission of UNESCO - International Oceanographic Data and Information Exchange (IOC/IODE)

International Council for the Exploration of the Sea (ICES)

Census of Marine Life - Ocean Biogeographic Information System (CoML/OBIS)

International Association for Biological Oceanography (IABO)

Taxonomic Database Working Group (TDWG)

Flanders Marine Institute (Vlaams Instituut voor de Zee - VLIZ)

Marine Biodiversity and Ecosystem Functioning EU network of Excellence - MarBEF

Federal Maritime and Hydrographic Agency (BSH, Germany)
– local organiser

With support from the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU, Germany), the European Union and the World Data Centre for Oceanography A, Silver Spring

Federal Ministry for the Environment, Nature Conservation and Nuclear Safety

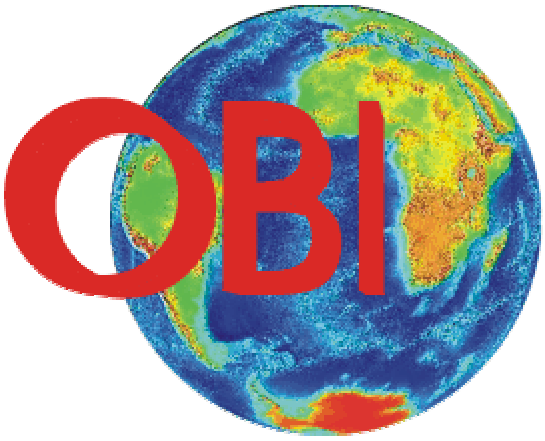
World Data Center for Oceanography Silver Spring

Table of Contents

Ocean Biodiversity Informatics

Hamburg, Germany: 29 November to 1 December 2004

Opening Session



From data to uncertainty - principles of data quality

A. Chapman

Australian Biodiversity Information Service, Australia

Data quality principles have become a core business practice in fields such as business, medicine and others over recent times, but they have barely touched the consciousness of the museum and taxonomic community. The rapid increase in the exchange and availability of taxonomic and primary species data has now made the consideration of such principles an important agenda item as users of the data begin to require more and more detail on the quality of this information. This paper examines principles of Data Quality as they apply to primary species data, and proposes a number of methods for cleaning and validating primary species data - museum, survey and observational.

Architecture and standards for Global Biodiversity Informatics - a GBIF and TDWG perspective

D. Hobern

Global Biodiversity Information Facility, Copenhagen, Denmark

In recent years many institutions have started using the Internet to provide access to their biodiversity data, particularly data on the occurrence of different taxa, based either on specimens held by museums and herbaria, or on field observations. Many of these projects are using the data standards and protocols developed through the Taxonomic Databases Working Group (TDWG) of the IUBS, including the Darwin Core and Access to Biological Collections Data (ABCD) exchange formats for taxon occurrence data, and the DiGIR and BioCASe protocols for handling remote queries against networked data resources. Many lessons have been learned from these experiences and TDWG is developing revised versions of these standards.

The Global Biodiversity Information Facility (GBIF) is an international organisation with the goal of providing integrated access to the wealth of biodiversity data distributed around the world. GBIF is using TDWG data standards to provide software access to these data and to support the development of web portals suited to the needs of different user communities. The flexible nature of the Darwin Core and ABCD schemas will mean that many different communities will be able to use a common architecture to share their data. This will ensure that effort applied to software development brings the greatest possible benefits.

GBIF is also developing a range of central services (including a registry of resource metadata, data index, schema repository, support for globally unique identifiers). These will assist users and organisations to discover and use all of these data resources. Although the first focus has been on taxon occurrence data, GBIF will soon be extending this model to support other classes of information, including information about taxonomic names, and descriptions of different taxa.

The standards and tools being developed are intended to be of use to all users of biodiversity data. Both TDWG and GBIF welcome involvement and contributions from all who hold such data or who need access to information on the world's biodiversity.

Fisheries impact on global marine biodiversity and ecosystems: inference from large heterogeneous data sets

D. Pauly [1], R. Watson [2]

[1] *Fisheries Centre, University of British Columbia, Canada*
[2] *University of British Columbia, Fisheries Centre, Sea Around Us Project, Canada*

Fisheries are often perceived as local affairs exploiting a limited number of species in a restricted area, with a matching time span usually restricted to one or two decades. This view strongly constrains the manner in which government laboratories analyse and present fisheries data. Since World War II however, marine fisheries have become increasingly global in scope, with distant water fleets roaming the high seas and also accessing, sometimes under dubious circumstances, the Exclusive Economic Zones of maritime countries.

Understanding the dynamics of this global enterprise and its impact on resources, markets and, increasingly, the food security of developing countries, requires that fisheries be analyzed on a global basis. Understanding their impact on the biomass of resource species and the collateral damage on the ecosystem in which these species are embedded requires, moreover, a long term perspective, rather than the few years generally considered when calculating next year's quota.

It is commonly argued that the data requirements for such global and long-term approaches are not available. We show here, however, that the creation of databases is possible, which allow mapping the spatial and temporal dynamics of the processes in question. Indeed, this is the only way the data and insights generated by more that one century of marine biology and fisheries science can be made accessible and useful for understanding major trends.

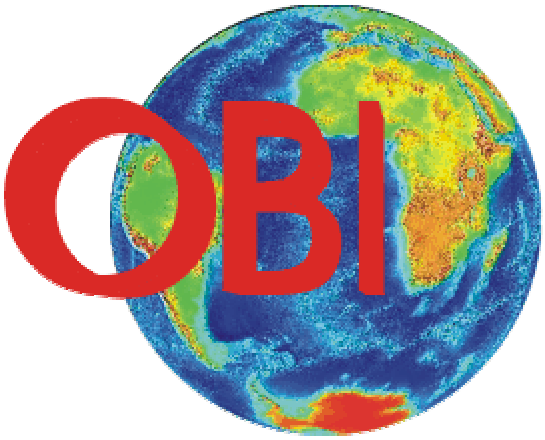
Specifically, we present here a GIS based approach for mapping world fisheries catches, which, combined with knowledge extracted from FishBase, the global, online encyclopedia of fishes and used to parameterise food webs (Ecopath models) could be used for reconstruction of long-term, ocean wide trends in marine fish biomass.

We also show how global catch maps can be used to refine previous analyses, demonstrating the occurrence of the 'fishing down the marine food web' phenomenon. Finally, combining the data sets mentioned above with others, notably socio-economic ones and other data will allow improved understanding of fisheries as a major force impacting on and eventually threatening marine biodiversity on a global basis.

Ocean Biodiversity Informatics

Hamburg, Germany: 29 November to 1 December 2004

Information system development



Emerging open source software, standards and protocols used for sharing and analysing marine biogeographic data

B.D. Best [1], P.N. Halpin [1]

[1] *Nicholas School of the Environment and Earth Sciences, Duke University, Durham, USA*

Development of marine biogeographic information systems requires common standards for data query, data discovery and data transfer. Information system developers must be able to attract both data providers and end users by offering highly flexible tools, processes and products. One of the most important approaches for attracting data providers is through the development of value-added products with dynamic, community-based websites. With software, protocols and user needs constantly evolving, a reliance on open-source and open-standard tool development provides maximum flexibility and responsiveness. Drawing from the OBIS-SEAMAP project (http://seamap.env.duke.edu), we will describe specific applications of open-source software packages (i.e. PostgreSQL, PostGIS, MapServer, DiGIR, Zope) used to store and interact with the marine datasets and species profiles, as well as the protocols (i.e. XML, GML, Z39.50) used to distribute biogeographic data and metadata using open standards (i.e. FGDC Biological Profile, ISO 19115, WMS, WFS) to clearinghouse portals (i.e. OBIS, GBIF, NSDI, Geography Network, NASA GCMD, and IOOS) and popular desktop software (i.e. Excel, ArcGIS and Matlab). We will also describe efforts to coordinate these emerging standards with the development of new data models under development for popular commercial GIS systems (i.e. ArcGIS Marine Data Model). We will further highlight the synthesis of real-time oceanographic and animal observation data through our internet mapping server, along with the visualisation and distribution of critical ancillary data, such as the associated lines of effort for dedicated surveys, inferred tracks of satellite telemetered animals, and polygons describing published distribution ranges.

The application of standardised data quality improvement methodologies to data describing marine biodiversity: three case studies illustrating the New Zealand experience

K. Duckworth

Ministry of Fisheries, Wellington, New Zealand

Poor data quality is the norm rather than the exception, but most organisations are in a state of denial about this issue (GartnerGroup, 1997).

Until the late 1990's the New Zealand Ministry of Fisheries work on improving the quality of the marine biodiversity data that it held was characterised by ad-hoc projects, often carried out by small groups of people working in isolation of each other, and intended to address specific immediate problems.

From 1998 the Ministry began to apply standardised methodologies, primarily developed by data managers working in commerce related fields, for improving data quality. These methodologies were found to be intuitive and easily transferable to the management of marine biodiversity data.

Key components of the methodologies include:

- Raising awareness within the organisation of the importance of data quality;
- Recognising the inherent "cost" (to the organisation and country) of poor data quality;
- Providing a useful definition of data quality;
- Providing a useful definition of the aspects (completeness, accuracy, format, timeliness, context and range) of data quality;
- Providing a framework for the development of plans to address all aspects of data quality, and prioritise the application of limited data quality improvement resources;
- Implementing these plans.

The most difficult aspect of implanting standardised data quality improvement methodologies was found to be measuring the quality of marine biodiversity data. This has made it difficult to quantify gains and measure the success of initiatives.

Three case studies of the application of standardised data quality methodologies to New Zealand's marine biodiversity data are described. These are with regard to:

- Catch and effort data reported by commercial fishers; and
- A web based GIS to display marine biodiversity data; and
- Catch and effort data reported by observers onboard commercial fishing vessels.

Key ingredients for developing a national oceans portal

K. Finney

National Oceans Office, Tasmania, Australia

Developing a widely accepted, innovative piece of software is a bit like baking a tasty cake. You either manage to gather all the right ingredients, mix them together in the right proportions according to a good recipe and cook them in a suitable environment to produce something that everyone wants to take a bite of, or you skimp on one of these processes and produce something that the vast majority would find inedible. In this latter case most people throw the cake out and start again, along the way losing consumer confidence and generally consumer willingness to re-invest in future baking ventures. Very mindful of this analogy, but sensing that the environment is right (both institutionally and from a technological standpoint), Australia's National Oceans Office (NOO) has taken on the challenge of delivering an Oceans Portal that will take its input from a distributed network of interoperable national/international data and service providers, linked through a web services catalogue. Whilst the project is still "cooking" and its success is not yet measurable, indications are that it has the capacity to lay an important foundation for the evolution of Australia's Spatial Data Infrastructure (ASDI). Whilst the Portal is an important instance of an end-user application that will provide a mechanism to discover and manipulate distributed resources, it is the back-end infrastructure encompassing a marine catalogue and the capabilities of the service providers that are at the heart of the project. This paper examines the key ingredients of Australia's Oceans Portal Project including: the building of a core community of interest; gaining agency level support, co-operation and commitment; funding; and selecting or developing the most appropriate catalogue, metadata, thesaurus, ontology, web services and technology standards. Blending these standards together within a suitable services-oriented architecture raises a number of interesting interoperability issues that really need to be solved through a globally cooperative approach. This paper discusses these issues and outlines design considerations that are required to underpin an infrastructure that will allow Australia's marine community to exchange and integrate socio-economic, biological, chemical, and physical oceanographic data related to the marine environment. It also explores how Australia's Oceans Policy, particularly its major implementation vehicle, regional marine planning, has created the type of environment conducive to tackling such an ambitious project.

Networking biodiversity data – online access to distributed datasources in GBIF-D

A. Hahn [1], A. Kirchhoff [1], W.G. Berendsohn [1]

[1] *Botanischer Garten und Botanisches Museum Berlin-Dahlem (BGBM), Freie Universität Berlin, Germany*

The Global Biodiversity Information Facility (GBIF) encourages, coordinates and supports the development of the worldwide capacity to access the vast amount of biodiversity data being held in biological collections such as natural history collections, botanical gardens and culture collections. GBIF is evolving into an interoperable network of biodiversity databases and information technology tools that opens up new possibilities for the utilisation of research data (*www.gbif.org*). GBIF participants (countries and international organisations) agreed to establish so-called network 'Nodes', with the joint mission of making the world's primary data on biodiversity freely and universally available via the Internet.

GBIF-D Botany (*www.gbif.de/botanik*), coordinated by the BGBM, is part of the German GBIF Node System (GBIF-D). The system, currently funded by the German Federal Ministry for Education and Research, consists of seven nodes, each focussed on a specific taxonomic segment but together covering all groups of organisms. The botanical node's tasks comprise setting up a communication structure for the community, supporting capture of text and image data in German botanical collections, and, as its central aim, providing integrated access to the existing data sources on botanical data by means of a central access system on the WWW.

Individual collections use a potpourri of database management and operating systems, and the data structures of collection and observation databases also vary to a great extent. Technology developed by the EU-financed project "BioCASE" (A Biological Collection Access System for Europe, *www.biocase.org*) allows to distribute a query to all registered data providers. A prototype WWW portal developed for GBIF-D Botany uses that technology to offer users a search form, and it is presenting the result to the user. The applied wrapper and XML technology allow use of a uniform query to access the data and to represent results in one and the same format. The different resources are thus 'virtually' integrated. The data access process in detail:

(1) For a given user query, the system consults a meta-database to identify relevant registered providers.
(2) The query is translated into a common query format that is BioCASe protocol compliant, using the UnitLoader software package of BioCASE. The query gets propagated to all relevant providers via http.
(3) Each provider database and web server is set up with a piece of software (the UnitWrapper) that translates the generic query into the native database language. This has to be set up only once, using configuration files.
(4) The UnitWrapper translates the database query results into an XML document (following the ABCD schema, www.bgbm.org/TDWG/CODATA/Schema/) and returns it for further processing and display.

In addition to a detailed and flexible representation of the primary object or observation data, ABCD documents allow the transfer of an extensive set of metadata, including information on the data source, the holding institution, and intellectual property rights. Both ABCD content specification and BioCASe protocol are also used by the international GBIF access system, so using BioCASe software tools ensures GBIF compliance.

Presently, GBIF-D Botany coordinates ten sub-projects that digitise data and provide databases of in-situ observations, living collections and herbaria representing the taxonomic groups of vascular plants, mosses and algae. Herbaria play an important role as biological archives, documenting the occurrence of organisms and communities in space and time. Two subprojects focus on collections of marine algae. Their data can assist research on water quality-changes within the last 200 years and may significantly contribute to the implementation of policies, e.g. the EU water framework directive. Through the GBIF-Network, their data will be integrated in a wide variety of further data networks, which opens new possibilities for data analysis.

The role of metadata in the management of data from the coastal zone

R. Lehfeldt

Federal Waterways Engineering and Research Institute (BAW), Germany

Information systems for the coastal zone provide options to systematically search and retrieve information to be used in decision-support systems for coastal management purposes, to meet reporting obligations to EU authorities and to support the communication process with the public. The web based North Sea and Baltic Sea Coastal Information System NOKIS (http://nokis.baw.de) relies on metadata, which describe contents, quality and accessibility of data in a formalised way, and gives indications of how to present the data.

This contribution presents a standardised metadata set for coastal zone purposes, which is a subset of the ISO19115 metadata standard. It was selected by the project group concerned with NOKIS and meets the requirements of both GIS and numerical modelling communities, which are important contributors to coastal information systems. It is currently being extended to incorporate monitoring aspects put forward by the European Water Framework Directive.

A user-friendly editor for this Coastal Zone Profile of the ISO standard is presented, which can be customised by local extensions of the XML schema describing the metadata model. This tool handles all documentary information of data and provides import filters for metadata exported by commercial GIS environments. In addition, a number of export filters have been implemented to serve the German environmental data catalog UDK and the German geo-portal gdi.de.

The information infrastructure of NOKIS relies on meta information prepared with this tool at distributed data archives, which is then replicated onto a central server and used in the NOKIS web portal.

Main focus is on web services in keeping with ISO 19119 provided by the central web portal. Commonly used visualisation and analysis tools, which can directly access the distributed data, are key features of the NOKIS methods base to support efficient workflows. A bathymetry viewer and a simulation viewer are prototype examples of metadata-driven web services, which are embedded tools in the web site of the German Coastal Engineering Research Council.

A semantic modelling approach to biological parameter interoperability

R. Lowry [1], L. Bird [1], P. Haaring [2]

[1] *British Oceanographic Data Centre, Bidston Observatory, Prenton, United Kingdom*
[2] *Directoraat-Generaal Rijkswaterstaat; Rijksinstituut voor Kust en Zee; Afdeling den Haag, Netherlands*

The BODC Parameter Dictionary currently contains over 16,000 terms of which nearly 11,000 pertain to biological parameters. The Rijkswaterstaat database in the Netherlands covers over 10,000 types of measurement, most of which are either chemical or biological. A requirement to populate a metadatabase described in terms of the BODC dictionary from the Rijkswaterstaat database meant that parameter interoperability between these information sources needed to be addressed. One technique for approaching this is manual mapping, working term by term through one of the information sources then searching for matching terms in the other. However, whilst this may be feasible for dictionaries containing tens of terms, it is totally unrealistic when the counts run into the thousands and so an alternative, automated approach was required.

Automation was initially attempted using a semantic matching tool developed at Rijkswaterstaat to offer a restricted list of BODC terms (preferably a single term) as the possible matches for each measurement. However, this met with limited success because the BODC dictionary consisted of plain language terms that not been written with machine processing in mind and had no constraints on either syntax or vocabulary. To appreciate the problem consider the programming required to recognise that 'Calanus abundance', 'Number of Calanus', 'Calanus count' and 'Abundance of Calanus' essentially mean the same thing. Further, no dictionary, especially a dictionary without vocabulary constraints, is perfect and there is a high risk that matches will be missed due to basic errors such as spelling mistakes.

The Rijkswaterstaat database is described in terms of a data model that qualifies measurements through associated attributes describing, amongst other things, what was measured and how it was measured. This is an example of a semantic model in which an entity is described in terms of discrete items of information, called semantic elements. Ideally, these elements are atomic, unambiguous and therefore ideally suited to machine interpretation. It was concluded that the only way a mapping could be achieved would be to develop a model along similar lines to describe the BODC dictionary and then map the two models.

A prototype semantic model based on three sub-models, each containing between 10 and 12 semantic elements, was developed to describe the biota, biota composition and chemical terms in the BODC dictionary and populated with approximately 13,000 terms. This was used as a basis for a two-stage mapping to the Rijkswaterstaat data model. The first stage was to set up a mapping between the semantic elements in the two models. For example, it was established that the 'Parameter' element in the Rijkswaterstaat model was equivalent to the concatenation of the 'Param' and 'Param_Comp' elements in the BODC semantic model. The second stage was to produce a mapping between the vocabularies used in each set of matched semantic elements. For example, the Rijkswaterstaat compartment term 'Surface water' mapped to the BODC compartment term 'water column'. Once these mappings had been established an automated term generation procedure was used to translate sets of Rijkswaterstaat semantic elements into BODC terms and identify matches.

The result was an automated mapping for approximately 90% of the Rijkswaterstaat measurement description terms. Of the remainder, all but a handful were matched by straightforward extensions to the vocabulary mapping. However, a small number of problems remained that could only resolved by querying Rijkswaterstaat, including ambiguity caused by homonyms that only came to light through standardisation of the BODC model to the ITIS taxonomic database. This exercise has shown that semantic modelling is a very promising technique for automating parameter interoperability between biological databases. However, without standardisation, particularly in the description of taxonomic entities, matches will be missed and there is a small but significant risk of false matches between parameters that are totally different.

Evolving concepts in the architecture and functionality of OBIS, the Ocean Biogeographic Information System

T. Rees

Commonwealth Scientific & Industrial Research Organisation,
Marine Research, Tasmania, Australia

The initial release of OBIS, the Ocean Biogeographic Information System went live in January 2002 via the OBIS Portal at Rutgers University, NJ (USA) (http://www.iobis.org/), as a portal to query distributed data sources for information on the distribution of marine species. This "version 1" of OBIS comprised a simple text-based name matching system which would retrieve data matching a user's query (e.g. species scientific name), or not, according to whether or not any matching data could be located in real time among the remote data providers (ref. 1).
Though the above mechanism is adequate to retrieve data where a match is known to exist in advance (and also, the user can correctly spell a potentially complicated scientific name), it contains no "intelligence" on behalf of the Portal as to the scope, coverage and completeness of actual OBIS data holdings which can be used to provide any useful generalised pre-query information to users. In addition, there is no facility to select data by taxonomic group ("all whales", "all barnacles"), to suggest potential matches to a user's partially entered or misspelled scientific name, or to supply useful "preview" information regarding the results of any search prior to undertaking it, such as number of records available, spatial or temporal coverage, etc.
Over the period 2003-2004 the initial user interaction pages of OBIS were completely redesigned, concurrent with the construction of a new "OBIS Index" which contains (a) a species name index (including a degree of synonym resolution), (b) a spatial index using the "c-squares" method (refs. 2, 3) suitable for spatial searching and display of dataset footprints direct from the Index in the form of "quick maps", and (c) associated metadata harvested in advance by parsing the remote data sources. A new search application was also constructed which incorporates a number of innovations including "fuzzy matching" on species names if relevant, and construction of a set of custom taxonomic categories considered to be useful for searching on by both specialist and non specialist users. The application also permits return of hundreds of matching records simultaneously as compared with the serial processing in the initial version of OBIS, along with substantial improvement in both retrieval speed (on account of the relatively small size of the Index compared with the source data) and information-rich user environment.
The opportunity was also taken to incorporate a local Cache of key fields on data points from the various remote data providers, for further improvement on search speed and robustness when requesting actual OBIS point data for other mapping, modelling, or download purposes. All of the above resulted in the release of “OBIS version 2” in March this year. These features will be demonstrated online (if possible), and the opportunity taken to discuss design decisions, and potential future development options, with conference participants as representatives of both system developers, and potential users of the OBIS system.

References:
[1] Zhang, Yunquing and Grassle, J. Frederick, 2003. A portal for the Ocean Biogeographic Information System. Oceanologica Acta 25:193-197. Electronic copy available at http://www.iobis.org/PDFs/2002Yunq.pdf .
[2] Rees, Tony, 2003. "C-squares", a new spatial indexing system and its applicability to the description of oceanographic datasets. Oceanography 16(1):11-19. Electronic copy available at http://www.marine.csiro.au/csquares/csq-article-Mar03-lowres.pdf .
[3] Rees, Tony, 2004. Use of c-squares spatial indexing and mapping in the 2004 release of OBIS, the Ocean Biogeographic Information System. Abstract and Presentation, EOGEO 2004, London, UK. Available via EOGEO website at http://www.eogeo.org/Workshops/EOGEO2004/ .

Pan-European network for ocean and marine data & information management – (SEA-SEARCH)

D. Schaap

Marine Information Service, HW Leidschendam, Netherlands

Background
National Oceanographic Data Centres and Marine Information Centres play a key role in management of marine and oceanographic data & information. Traditionally main priorities were well-organised archival and quality procedures. Nowadays key issues are developing and operating value-adding services and products, serving a large user community in a coherent mode, in particular via internet services. A prerequisite for this is an excellent cooperation and networking of the data centres and developing and operating jointly internet infrastructures for entering, quality controlling, storing, retrieving and exporting marine & oceanographic data from various sources.

Five years ago the initial Sea-Search network was established as a European cooperative network of 16 national oceanographic data centres and marine information centres from 14 European coastal states and EC within the framework of the EC-MAST programme. These activities were continued from November 2002 onwards as part of the 5th framework of the European Commission, now encompassing a Pan-European cooperative network of 33 National Oceanographic Data Centres and marine information services from 30 coastal states, riparian to all European seas.

Each Sea-Search partner is specialised in managing and giving access to extensive resources of oceanographic and marine data & information and is providing data management services & support to a wide range of institutes and research projects, both nationally and internationally.

Objectives
The primary goals of Sea-Search are to provide users with a central overview and access to ocean and marine data & information in a wider Europe as well as to harmonise quality standards and create improved conditions for data exchange and wider use in multidisciplinary applications.

Present Results
The 33 partners operate and further develop a network of partner websites and a joint European website (www.sea-search.net). At present the portal site hosts an array of catalogues, overviews and links and acts as central gateway to ocean and marine information & data resources in Europe, e.g:

- online directories of marine datasets (EDMED), marine research projects (EDMERP) and sea-going cruise summary reports (CSR) with input and regular updating of all research institutes in the countries, riparian to the European seas:
  - The European Directory of Marine Environmental Datasets (EDMED) is the de-facto standard directory in the oceanographic community for describing and locating datasets. At present the EDMED directory contains more than 3854 datasets from over 580 institutes in Europe
  - The European Directory of Marine Environmental Research Projects (EDMERP) is a relative new and fast growing directory, containing already more than 780 research projects from over 240 institutes in Europe. Next to EDMERP the Sea-Search website also hosts and maintains a Directory of EC supported marine research with more than 430 descriptions of EC Research Projects.
  - The Cruise Summary Reports directory is based upon the classical ROSCOP format for reporting on sea-going cruises, but within Sea-Search a more complete coverage of cruises by European institutes is emphasised.

- a Common Data Index (CDI) database to enable users to get highly detailed and up-to-date insight in the availability and geographical spreading of marine data across the different Sea-Search partners and possibly beyond. A pilot CDI database including a versatile geographical web interface will be launched in December 2004, covering an index to a number of major oceanographic data centres. Next to the CDI format also an XML / XSL format has been defined, compliant to ISO-19115 metadata standard, to facilitate updating and to enable easy participation by other data centres and data managing institutes.

Coming Developments and Results:
Further development activities of Sea-Search in the coming year will focus on:

- Streamlining the updating process and mechanism for the 3 directories (EDMED, EDMERP and CSR) by establishing ISO compliant XML formats and applying XML for direct exchange and implicit quality control between datacentres and central directories.
- Further roll-out of the Common Data Index database to cover all Sea-Search partners and beyond and to include pilots for direct access to data, managed at partner websites
- Seeking close cooperation and mutual tuning with specific user communities like Operational Oceanography, Remote Sensing, Coastal Management, Regional Marine Conventions, Water Framework Directive, …
- Preparing the next step (SeaDataNet) for further development into a pan-european and harmonised infrastructure for archival and provision of marine environmental data:
  - providing full and transparant data access by internet
  - streamlining overall process from data acquisition, QC and storage to online dissemination and distribution
  - adopting overall standard QC protocols
- providing standard automated data & information products

Potential exploitation by end users
Sea-Search provides improved overview and access to oceanographic in situ data for a very broad spectrum of uses, encompassing operational oceanography, physical modelling, ecosystem modelling, marine biodiversity, algal bloom studies, amongst others. Sea-Search supports not only research and education, but also coastal zone management, policy making, offshore engineering, fisheries, aquaculture, and recreational users. Sea-Search is very relevant for the execution of international protocols, conventions and agreements, that have been signed by coastal states for protection of the seas, like OSPAR, HELCOM, Barcelona and Black Sea Conventions. Sea-Search is also very relevant for implementation of Europe's environmental policy concerning Integrated Coastal Zone Management (ICZM) and the Water Framework Directive.

In addition, Sea-Search provides a support infrastructure for organisations and projects, dealing with ocean and marine data & information, for indexing, disseminating and promoting their data & information resources to a wide user community.

Reference: www.sea-search.net

Metadata exchange with use of ontologies

I. Shevchenko[1], G. Moiseenko[2], O. Vasik[1]

[1] *Pacific Fisheries Research Center (TINRO-Center), Vladivostok, Russia*
[2] *Russian Federal Research Institute of Fisheries and Oceanography (VNIRO), Moscow, Russia*

Collected marine biology data are usually treated as proprietary and no direct access for the information retrieval is available. Metadata describe different characteristics of data. Metadata bases provide data about data that allow one to understand if some data are of interest for a particular purpose There are several standards for metadata description. Some of the data characteristics are relatively static. However, there are metadata descriptors that have to be changed every time when a data set is updated and it is practically impossible to keep the metadata up-to-date without use of automated procedures.

We propose an architecture and implementation of a prototype of the system for the publication of metadata on the Web with use of a draft of the Marine Biology ontology. The approach allows us to integrate distributed metadata repositories by providing the semantic foundations for translations among different languages and representations.

OWL (Ontology Web Language) is used for the ontology implementations. In particular, we utilise ontology mappings to indicate that a class or a property in one ontology is equivalent to a class or a property in a second ontology. Protégé, an ontology editor, has an OWL Plug-in and provides an environment for creation of ontologies. Algernon, a rule-based inference system, has an interface with Protégé and is used for information retrieval.

References

1. Guarino N. and Giaretta P. Ontologies and knowledge bases: Towards a terminological clarification. In N.J.I. Mars, editor, Towards Very Large Knowledge Bases, — IOS Press, 1995, http://www.loa-cnr.it/Papers/KBKS95.pdf
2. Noy N.F., McGuinness D.L. Ontology Development 101: A Guide to Creating Your First Ontology.— Stanford University, 2001, http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html
3. Smith M.K., Welty C., McGuinness D. OWL Web Ontology Language Guide W3C Recommendation, — World Wide Web Consortium (W3C), 2004, http://www.w3.org/TR/owl-guide/
4. World Wide Web Consortium Issues RDF and OWL Recommendations.— World Wide Web Consortium (W3C), 2004, http://www.w3.org/2004/01/sws-pressrelease.html.en
5. Horridge M. A practical guide to building OWL ontologies with the protege-OWL plugin.— 2004, http://www.co-ode.org/resources/
6. Hewett M. Tutorial: Getting started with Algernon.— 2003, http://algernon-j.sourceforge.net/tutorial/
7. Ecological Metadata Language (EML) Specification — http://knb.ecoinformatics.org/software/eml/eml-2.0.0/
8. Macklin S.A., Megrey B.A. The North Pacific Ecosystem Metadatabase. PICES Press, Vol. 12. No. 2. Pp. 30-33.

How do you build an information system that works? Lessons from environmental case studies

K.I. Stocks [1], K.S. Baker [2]

[1] *San Diego Supercomputer Center, University of California, San Diego, USA*
[2] *Scripps Institution of Oceanography, University of California, San Diego, USA*

As the number and variety of marine data partnerships increases, and as integrated ocean observing systems are growing worldwide, information system development is emerging as an important frontier in ocean sciences. Both systems theory and past experience suggest that scientific expectations, data practices, and metrics of success are evolving as the field develops. Future growth may be facilitated by the development of an ocean informatics community that learns from and builds upon the experiences of past projects.

To this end, we present and characterise the development process of four projects building environmental information systems: 1) the Ocean Biogeographic Information System; 2) the Long Term Ecological Research community 3) the California Cooperative Oceanic Fisheries coastal monitoring study; and 4) the SeamountsOnline database of seamount biota. These systems range in age from relatively recent to more than half a century; they range in complexity from a single-researcher, centralised database to an international federation of data providers.

Rather than focus on technological aspects, we consider the organisation of these projects. For example, whether there is a clearly-articulated vision for the project and what metrics of success are used to track progress towards that vision, how decisions are made, to what degree and in what capacities the project has included participants from diverse disciplines (ocean/environmental scientists, field technicians, information managers, computer scientists, user communities, social scientists, and educators), and how the program partners at institutional and functional levels with other programs.

By comparing and contrasting across the four case studies, lessons emerged regarding characteristics that facilitate project outcomes and the specific mechanisms supporting those characteristics. These include:

Characteristic: flexibility in system design through time
- Mechanism: multiple prototype development for new components
- Mechanism: incorporating flexibility as an element of design

Characteristic: participant engagement/level of involvement
- Mechanism: consensus decision making in concert with hierarchical representation when needed for large bodies
- Mechanism: seed resources available for allocation to participants

Characteristic: usefulness of the system
- Mechanism (for long-term usefulness): clearly articulated and focused project vision/goals produced with broad input from users and participants
- Mechanism (for short-term usefulness): modular development with usable products at each step

Characteristic: Open communication
- Mechanism: establishing times timeframes and mechanisms for communi-cation (e.g. regular conference calls, meetings, email discussion lists, etc.)

- Mechanism: mixed representation on decision-making groups. E.g. having domain scientists participate in technical working groups, and technical members participate in scientific goal-setting.

- Mechanism: foster "bilingual, "multi-perspective" or "mediator" members able to understand and communicate with more than one discipline; recognise that the outputs of these members may not track traditional scientific measures of productivity (i.e. papers published or code written).

Characteristic: Sustainability

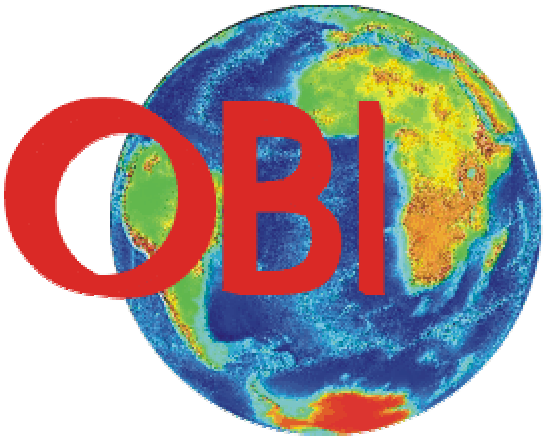
- Mechanism: recognising and providing rewards for participation at the institution, project, and individual level.
- Mechanism: develop cooperative relationships with related projects

Over the last decade the question 'how to build a data system' has broadened to 'how to build an information system that works' but emerges today as 'how to build a community' – a community organised to identify, address, and respond to the intertwined organisational and technical issues arising in the ongoing process of information system design.

Ocean Biodiversity Informatics

Hamburg, Germany: 29 November to 1 December 2004

Taxon-based systems



NeMys: an evolving biological information system, a state of art

T. Deprez [1], M. Vincx [1], E. Vanden Berghe [2], J. Mees [2]

[1] *Universiteit Gent, Vakgroep Biologie, Afdeling Mariene Biologie, Gent, Belgium*
[2] *Flanders Marine Institute (VLIZ), Oostende, Belgium*

The taxonomical online database system NeMys was originally developed to store literature based morphological and geographical information on Nematodes and Mysida. As the system was constructed as generic as possible it could also handle other taxa like e.g. phytoplankton, turbellaria, … and even non-marine taxa. New taxa also asked new requirements to the system. Therefore recently a new module on collection items was added. This enables users to add a collection item and add morphological, morphometrical and geographical data for specimens studied.

The Mysida dataset, the largest and most complete dataset running on Nemys uses all features of NeMys. Morphologically all genera have been defined and progressively also data on the species is added. Geographically information on more than 2000 stations where Mysida were found and about 5000 literature references, more then half of it in an electronical version, are available. Pictures giving a good morphological view of almost all species are progressively added. A start has been made with adding data on the most important worldwide mysida museum collections. All this together makes the Mysida dataset a good testing example for fine-tuning the online data-consultation and data-entry interfaces.

The online availability and use of common used technologies and database structures makes the NeMys system more then just a simple taxonomical dataset. NeMys can now (thanks to recent additions) be seen as fully functional biological information system which is implementable for any taxon.

The IOC-ICES-PICES harmful algal event data-base, HAE-DAT

H. Enevoldsen[1], M. Lion[2], C. Sexto[2], B. Sims[3]

[1] *University of Copenhagen, Botanical Institute, Department of Phycology, Copenhagen, Denmark*
[2] *Centro Oceanográfico de Vigo, IOC-IEO Science and Communication Centre on Harmful Algae, Vigo, Spain*
[3] *Intergovernmental Oceanographic Commission of UNESCO, Paris, France*

The Intergovernmental Oceanographic Commission of UNESCO (IOC) and the International Council for the Exploration of the Seas (ICES) have jointly developed the 'Harmful Algal Event Data base HAE-DAT' which holds information on harmful algal events such as origin of data and reference, general information, location and date, microalgal species, environmental data, harmful effects, and complementary information. At present the HAE-DAT covers the North Atlantic from 1987 onwards, but it is the objective to build a global harmful algal event database. IOC and ICES have invited PICES to become partner in HAE-DAT and to that effect PICES is testing the reporting format during 2004. The IOC networks on harmful algae in South America, the Caribbean and North Africa are preparing to participate. HAE-DAT has associated decadal maps of the occurrence of harmful algal event. These maps are offered in cooperation with the French Institute for the Exploration of the Sea (IFREMER). The IOC is presently developing a new software platform for HAE-DAT which will ease input of data and generation of maps.

Lessons learned in the design, data management and user needs of fishbase.

R. Froese

Leibniz-Institut für Meereswissenschaften, IfM-GEOMAR, Kiel, Germany

FishBase (www.fishbase.org) is a large information system with key data (names, size, photos, distribution, biology, population dynamics, trophic ecology, etc.) for all 28,500 fishes of the world. The backbone of FishBase is a relational database with more than 100 major tables and several thousand data fields. The database design balances the technical need for normalization with the need for user-friendliness, i.e., keeping related information in the same table so that advanced user who are familiar with relational databases have no problems writing queries. Other users are given access through an extensive online interface with over one thousand dynamic web pages which are generated on demand. Here the challenge was to combine simplicity andease of use with the speed and power of a relational database. In May 2004 FishBase received over 12 million hits from over 600,000 visitors per month. Most users come in with common names and look at summary information and photos of species. However, there is also substantial usage of specialist information such as taxonomy or population dynamics and of specialist tools such as the Life-history tool or the length-frequency wizard. Intelligent deep-linking into other relevant online databases such as Google, Genbank, GBIF, or the IUCN Red List make FishBase a one-stop portal for all you ever wanted to know about fish; the techniques behind this will be discussed.

Communicating the properties of marine organisms as the second dimension of marine biodiversity

W. Greve

German Centre of Marine Biodiversity Research (FIS), Hamburg, Germany

The abundance distribution of marine populations on all levels of abundance, including possible extinction, depends on the species-specific ecological niches, defining the requirements and the match/mismatch of requirements and supply. The definition of the degrees of freedom, their dimensions and numeric communication of these properties of marine organisms enables the understanding of abundance distributions.
The collection of taxonomic biodiversity distribution information offers the opportunity to expand the acquisition on the properties of the populations. This second dimension of biodiversity treated in ecology, ontogeny, physiology and behaviour is less organised than taxonomy. Within the strategic levels from (1) observation and definition, (2) discretisation, order and analysis and (3) synthesis and operative modelling the unifying concepts of the comparative property analysis are unsufficiently developed. A strategy for the comparative investigation of population properties is developed on the basis of a zooplankton example treating complex trophodynamics, ontogeny, physiology and physical forcing in a simulation model with standardised sensitivity analysis enabling the ranking of ecosystem parameters (population properties) according to the prognostic potential. The results are discussed in the light of recent phenological research changing the ecosystem paradigms. The development of strategies for the comparative investigation of population properties, computer accessible documentation, modelling and operative modelling within the framework of global biodiversity research is suggested.

Validating marine biodiversity: interconnecting web-based DNA taxonomy and geography

S. Lavery[1], H. Ross[1], M. Dalebout[1], M. Goode[1], G. Ewing[1], P. McLaren[1], A. Rodrigo[1], C.S. Baker [1]

[1] *School of Biological Sciences, University of Auckland, Auckland, New Zealand*

Efforts to describe global marine biodiversity have improved dramatically in recent years, largely due to the advances in integrating large web-accessible databases (e.g. OBIS, GBIF). A great benefit of these systems is the wealth of geographic information provided, although a weakness is the potential for database errors in taxonomy and consequential misinterpretation. In contrast, molecular (DNA) databases potentially have enormous power in ensuring correct taxonomy, but provide little or no geographic data (e.g. Genbank). As yet, these considerable, complementary data sources have not been brought together successfully, despite the great promise this may provide in both data validation (of both sources) and data exploration (e.g. in phylogeographic diversity). We present a web-based approach to DNA taxonomy that begins to address this interconnection with geography, and highlight some future mechanisms required to achieve this.

The identification of species using molecular markers is becoming a major bioinformatic application. www.DNA-surveillance provides a powerful tool for undertaking molecular taxonomy by utilising greatly improved DNA sequencing capabilities, and applying appropriate bioinformatic and phylogenetic analyses. This web-based application assists in the identification of species of unknown specimens by aligning user-submitted DNA sequences with a validated data set of reference sequences. This approach is of most immediate impact for groups of organisms that are widely distributed, currently exploited, and difficult to identify morphologically. The first taxonomic group implemented in DNA-Surveillance are the cetaceans (whales and dolphins). Our program, unlike the DNA-barcoding approach, takes a phylogenetic approach to species identification, using as a criterion strong support for a monophyletic grouping with reference sequences from one species. This application is currently being upgraded to incorporate recent advances in determining the certainty of identification using maximum likelihood. Such purpose-built molecular taxonomic applications offer great advantages over generic databases such as Genbank, which are relatively ill-suited to the task of species identification for a variety of reasons.

All specimens included in DNA-surveillance have accurate geographic records, and these can be used to ascertain the geographic coverage of validated specimens of a species. We have also implemented a large phylogeographic dataset for one species (the humpback whale). Together with future plans for geographic mapping tools, we will highlight the potential applications of this powerful interconnection between DNA taxonomy and geography.

FAO databases on marine species identification for fishery purposes

J. Lleonart [1], M. Lamboeuf [1], M. Taconet [1]

[1] *Food and Agricultural Organisation of United Nations, Rome, Italy*

The FAO Species Identification and Data Programme (SIDP) was initiated in the early 1970s to promote the upgrading of fisheries data by species through reliable species identifications in the field and to rationalise and expedite fishery work in all fields by furthering the use of correct scientific and standardised vernacular species names. To meet these objectives, the SIDP has so far published 10 Regional Guides (comprehensive inventory of the species of a wide region of the sea), 15 Field Guides (illustration based for quick identification guides for areas of one or few countries), 20 Catalogues (worldwide inventories of taxonomic groups), and 150 Synopses (comprehensive reviews of the present knowledge on single species). These Guides and Catalogues provide for species of importance to fisheries information on scientific nomenclature, international and local names, diagnostic features, geographical distribution, biology, fisheries and relevant literature. The geographical distribution information is presented as notes and maps.
The FAO Fisheries Global Information System (FIGIS) is a network of integrated fisheries information. FIGIS was conceived in a context of great stress on most major fishery resources and a worldwide concern about the state of the resources and some of their non-sustainable uses. For yet under-utilised stocks, the risk of over-utilization is amplified by globalisation of trade in fish and fishery products. When the 1995 Code of Conduct for Responsible Fisheries was approved as a basis for policies aimed at sustainable fisheries, a major need for reliable, high-quality and relevant information on the state of world fisheries was identified. FIGIS - the fisheries global information system - was established to address this need. With the adoption by the Committee on Fisheries of the Strategy for Improving Information on Status and Trends of Capture Fisheries (SSTF) on 28 February 2003, FIGIS becomes one of the privileged tools for its implementation.
SIDP and FIGIS are developing data-base tools to make this information available on-line. Four approaches are considered in this development, i.e. taxonomic, fisheries, ecological and management. The current objectives and uses are the following:
Taxonomic: Need to improve taxonomic knowledge regarding fish, crustaceans, molluscs, and other organisms important to fisheries.
Fisheries: Need to improve statistics (through species identification). Integration of Ecosystem Approach to Fisheries. Take into account new developing fisheries, (deep sea fishing), new methodologies, new commercial species and aquaculture production.
Ecological. Addressing biodiversity and biogeography issues, including cartography, GIS, species introductions, faunistic changes, tropicalisation, particular sensitive environments, such as: coralligenous, seagrass or maërl bottoms, coastal, lagoon, etc.
Management. Holistic approach to world fisheries. Threatened species and ecosystems. Highly vulnerable fisheries resources.
Considering the challenges for the future, the following issues should be taken into account:
Taxonomy: Enhance the on-line information of some taxonomic groups. Develop tools to identify “related species”: parasites, preys, etc. in order to assist the ecosystem approach. Develop tools to identify parts of individuals (otoliths, vertebrae, pieces of crustacean carapaces, etc.) particularly useful to analyse stomach contents.
Fisheries: Develop a common format for assessments. Fisheries modelling (including socio-economics and ecosystems): data and models.
Ecology and ecosystem. Species dynamics: describe the history (not only update data). Analysis of qualitative data. Changes in abundance. Changes in geographical distribution. Evaluation and assessment of biodiversity. Quantitative ecology and ecological modelling.
Management. Use of GIS. Take into account the history of fisheries and environments. Legislations (international, regional, national).
Transversal issues. Language, need for a common glossary. Vernacular names. Links between taxonomic, fisheries, ecology and management data bases

The information system of the marine animals collection (fish and invertebrates) in the Zoological Institute of the Russian Academy of Scienes

I.S. Smirnov [1], E.P. Voronina [1], A.L. Lobanov [1], A.V. Neyelov [1]

[1] *Russian Academy of Scienes, Zoological Institute, St. Petersburg, Russia*

The Zoological Institute (ZIN) in St. Petersburg occupies one of the prime places among the world's largest storehouse of zoological collections. The marine invertebrate collection contains over 100 thousand samples of 26 thousand species. The ichthyological collection contains over 160 thousand catalogued specimens of the 8700 species of marine and freshwater fishes and fishlike vertebrates of the world fauna.
The creation of the electronic databases (DB) started at the Zoological Institute in 1987, first on marine invertebrates, and then, in 1991, on fishes. The DB on different groups of animals are often interactive and successfully adding each other. Example of such combination is the DB on marine fishes and invertebrates. The material on both groups of animals can be collected at the same stations during the expeditions. Fishes and invertebrates are the main components of any marine biocoenosis, so their parallel research with using joint DB is very prospective to study biocoenotic relationships and fauna complexes.
Designed at the ZIN informational retrieval system (IRS) “OCEAN” consists of 4 main tables: the taxonomic table containing the name and nomenclature of taxa, the geographical table including the data of field books and catalogue of museum collections (locality of sampling: coordinates of stations, gear etc.), ecological part (biomass, depth, temperature, salinity, oxygen etc.) and bibliographic one.
In the course of organisation of faunistic and ecological data and IRS, two problems appeared: a) inputting and using of the scientific names, especially synonyms; b) the formalisation of geographical data. The first problem is being solved by using the classificators of scientific names of animals based on the user-friendly and periodically updated ZOOCOD standard, popular among the Russian institutions dealing with the biodiversity research. Coordinates and the developed geographical information system were used to solve the second problem.
Creating of the electronic DB and designing of the IRS carried out at the ZIN are the basis for the project “ECOANT” - “Creation of an IRS on ECology of benthos of the ANTarctica”. IRS “ECOANT” can promote the resolution of the following problems: refining the fauna structure of biota and its taxonomic features for the different areas; to get ecological information; to reveal the changes in the structure of fauna in the investigated regions under the influence of climatic and anthropogenic factors, which is one of the aims of a global ecological monitoring. The preliminary information about this project is available on the site http://www.zin.ru/projects/ecoant/index.html.
Now, the ichthyological part of the taxonomical table contains 5700 records; all high-level taxa including families, species of fish and fish-like animals of the Antarctic Region, Pleuronectiformes and some families of Scorpaeniformes of the world. As a list it is exhibited on the web site of the ZIN. The collection table includes data on the specimens of order Pleuronectiformes worldwide and of some families of order Scorpaeniformes. DB on marine invertebrates contains now more than 40000 records (more than 175 expeditions) on locality and the information on the chitons, bivalves and brittle stars collections.
The international projects to create DB and IRS sharing of biodiversity information on a global scale are developing. The system of data input in these projects usually is based on the Latin symbols, but not Cyrillic. It reduces the information of the original labels, hand-written in national languages, e.g. Russian. The IRS “OCEAN” provides the input data and the different queries using Cyrillic symbols alongside the Latin ones.

Support: Project N 11 “Exploration and research of the Antarctic Region”, grant RFBR № 02-07-90217 and the program “Information system on a biodiversity of Russia”.

TAXEX: TAXonomical EXpert System - history of development and technology of identification

S.G. Lelekov [1], Yu.N. Tokarev [1], V.V. Melnikov [1], N.P. Pakhorukov [1], A.M. Lyakh [1], E.Yu. Georgieva [1], S.A. Tzarin [1], V.F. Zhuk [1], Yu.B. Belogurova [1], O.N. Danilova [1]

[1] *Ukraine National Academy of Sciences, Institute of Biology of the Southern Seas, Sevastopol, Ukraine*

From the end of 80's in the department of Biophysical Ecology of Institute of Biology of the Southern Seas (Ukraine) Computer Taxonomical Expert Systems (TAXEX) for identification of marine organisms have being developed. At the first stage, when TAXEX was operated by MS DOS, the process of accumulation and storing information on marine organisms in databases was realized, and the technology of Taxonomical Expert Systems creation was developed and tested. In the beginning of 90's identifiers of the Black Sea Isopoda, Cirripedia and shrimps had been developed (only Russian versions). At the next stage, when MS Windows became widely distributed and Windows Programming with Rapid Application Development software became very simple, the main goal of the project was the information provision of expert systems. At this time identifiers of the Black Sea fish larvae, the Black Sea Bivalvia and Gastropoda were created. When Intranet and Internet became widely distributed, the first TAXEX versions written in Java appeared. In present, when huge arrays of knowledge about marine flora and fauna were collected, we integrated TAXEX with Taxonomical Information Systems and developed main components of such Taxonomic Information and Expert System (TIES).
The main advantages of TIES are:

- storing knowledge of expert-taxonomists in electronic format, and rendering of users by a comfortable interface for retrieving and managing of his knowledge;
- providing scientists with qualitative tools for identifying taxa using their external and internal features. The identification is possible without physical injuries of organism;
- training of new professional taxonomists by means of transferring knowledge, stored in knowledge base, to the next generation of specialists using special training software.

On the stage of realization fo the TIES project we still have some difficulties, mainly of which are:

- a lack of financing and as a consequence a lack of interest from proficient experts in developing such a system;
- loss of professional taxonomists and absence of a young generation of taxonomists.

At present time we are developing identifiers of Mediterranean Fishes and Chaetoceros of the Black Sea. Here we demonstrate the developed identifiers and describe the technology of biological objects identification using the expert system.

European Register for Marine Species version 2.0: data management, current status and plans for the future

E. Vanden Berghe [1], P. Bouchet [2], G. Boxshall [3], M.J. Costello [4], C. Emblow [5]

[1] *Flanders Marine Institute (VLIZ), Oostende, Belgium*
[2] *Organisms and Ecosystems, Paris, France*
[3] Natural History Museum, Department of Zoology, United Kingdom
[4] *University of Auckland, Leigh Marine Laboratory, Warkworth , New Zealand*
[5] *Ecological Consultancy Services Ltd,Dublin, Ireland*

ERMS was created by a project funded by the European Commission over 2 years, from 1998 to 1999; it resulted in a species list for European marine waters, which was published both on the www and as a book; a scientific society, the Society for the Management of European Biodiversity Data (SMEBD) was created to manage the intellectual property rights. Marine Biodiversity and Ecosystem Functioning (MARBEF) is a recently approved EU Network of Excellence; one of the objectives of MARBEF is to integrate existing, isolated datasets in one consolidated database. Because a standard species authority file is a critical infrastructure in this process, updating ERMS is part of the data management component of MARBEF. SMEBD has authorised VLIZ to manage ERMS and its updating within MARBEF.

The original website (ERMS 1.0) has now been replaced by the species register served from a relational database (ERMS 1.1, same taxonomy, different structure, added classification) hosted by the Flanders Marine Institute (VLIZ). Within the recently funded MARBEF project, ERMS has been transformed from a static online searchable publication, into a dynamic system. Approved editors of species lists will be able to access ERMS 2.0 online, such that the register will be changing according to need and editorial diligence. At regular intervals, a snapshot of the database will be archived; these archive copies will be kept available through the ERMS web site.

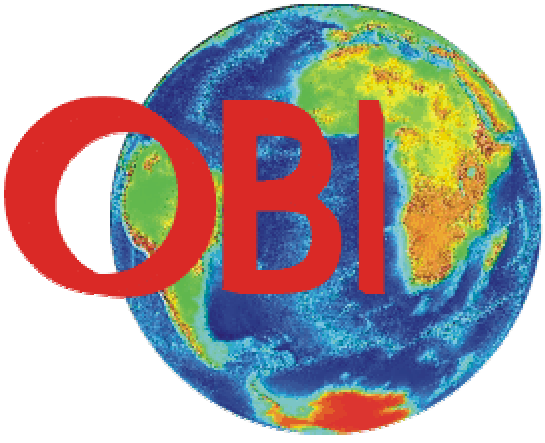
The global Species 2000 initiative has obtained EU funding for a project that will create an on-line “EuroHub” that will connect to the three main species databases for Europe, namely ERMS, Fauna Europaea and Euro+MedPlantBase. While the technical details of how this will be done remain to be finalised, it would seem that interoperability between on-line databases is now relatively easy to implement, using data exchange protocols such as DiGIR (Distributed Generic Data Retrieval, used by OBIS and GBIF) and BioCASe.

The structure of ERMS 2.0 is a variation on the structure of ITIS: an open hierarchy, where each taxon record points at a ‘parent’ record. Synonymy is implemented in a similar way. Information on when a record is created, updated or checked by a taxonomic editor is kept. Currently, there are approximately 47,000 taxa listed, of which 30,000 are valid species, 3,000 synonymous species names, and 14,000 names of taxa with rank of genus or higher.

Ocean Biodiversity Informatics

Hamburg, Germany: 29 November to 1 December 2004

Geography-based systems



MedOBIS: biogeographic information system for the Mediterranean and Black Sea

C. Arvanitidis [1], V.D. Valavanis [2]

[1] *Hellenic Center for Marine Research, Environmental Technology and Management Group, Heraklion (Crete), Greece*
[2] *Hellenic Center for Marine Research, Heraklion Crete), Greece*

An effort to assemble, format, analyse and disseminate surveyed marine biological data is currently under development at the Hellenic Center for Marine Research (HCMR), Greece for the Mediterranean and Black Sea region. The effort is supported by the MedOBIS project (Mediterranean Ocean Biogeographic Information System) in cooperation with the Department of Zoology, School of Biology, Aristotelian University of Thessaloniki (AUTH), Greece, the National Institute of Oceanography (NIO), Israel, and the Institute of Biology of the Southern Seas (IBSS), Ukraine.

The aim is the development of a taxon-based biogeography database and online data server with a link to survey and satellite environmental data.

The data assembly phase is based on the free contribution of biological data from various national and international surveys in the region. The data formatting phase is based on a Geographic Information System (ESRI ArcInfo GIS). Based on the geographic location of data stations, all data and their attributes are converted to shapefiles. The data analysis phase is based on GIS, e.g. species distribution maps, species-environment relations, etc. Finally, the dissemination phase is based on ALOV Map (http://www.alov.org/), a free, portable Java application for publication of vector and raster maps to the Internet and interactive viewing on web browsers. It supports navigation and search capabilities and allows working with multiple layers, thematic maps, hyperlinked features and attribute data.

The basic features of the MedOBIS online data server are its taxon-based search and mapping capabilities as well as its data navigation and downloading interface for the following regions: West and Central Mediterranean, Adriatic Sea, Levantine Sea, Aegean Sea, Bosporus Sea, and Black Sea.

The MedOBIS online marine biological data system (http://arch.imbc.gr/medobis/) will be a single source of biological and environmental data (raw and analysed) as well as an online GIS tool for access of historical and current data by marine researchers. It will function as the Mediterranean and Black Sea node of EurOBIS (the European node of OBIS Network, part of the Census of Marine Life).

The evolution and challenges of the ocean biogeographic information system

M. J. Costello [1], F. Grassle [2], Y. Zhang [2], K. Stocks [3], T. Rees [4]

[1] *Leigh Marine Laboratory, University of Auckland, Warkworth, New Zealand*
[2] *Institute of Marine and Coastal Sciences, Rutgers, the State University of New Jersey, New Brunswick, USA*
[3] *San Diego Supercomputer Center, University of California, San Diego, USA*
[4] *CSIRO Marine Research, Tasmania, Australia*

Science enters the millennium with a greater awareness of the global impact of fisheries and impending climate change, necessitating data and data analysis from global to local scales. Data are the foundation of science and their value increases with time as they become harder to replace. Such data include species descriptions, records of the presence or abundance of a species at a place and time, and associated environmental data. Computerised data management and analysis, called biodiversity informatics, provides the mechanisms for capturing, storing, organising, visualising, sharing, and analysing this information. Through collaboration between informatics experts and taxonomists, biodiversity informatics can increase the availability of authoritative information on species taxonomy. This increased communication will reduce taxonomic errors and enable taxonomists to spend more time describing new species instead of correcting errors of the past.

New technologies, including satellites and acoustic mapping, are providing data at global spatial scales on the physical conditions of the oceans. Physical oceanographers have cooperated at such scales for decades because of the high costs for their research and large spatial areas necessary to study physical processes and climate. In contrast, most research in marine biology and ecology has tended to be local in spatial scale, in part because so much can be discovered at such scales without the costs of sampling across oceans. However, informatics is enabling marine biologists to pool data, often collected for different purposes, to reveal patterns in biodiversity across ocean and global scales. Such patterns include global declines in fish stocks, large fish heading for extinction due to fisheries, and effects of climate change on open ocean plankton.

Marine biodiversity informatics has developed from publications on disk, to CD, and then to the web. The Ocean Biogeographic Information System (www.iobis.org) is mapping marine species over the World Wide Web. It has established a global network of Regional Nodes, and presently serves over 5 million records from over 30,000 species. It includes online data exploration tools, including a species name service, mapping species over selected ocean environmental data layers, and predicting where species may occur using two different modelling approaches. Such models have been used to assess the risk of invading species. OBIS has evolved from a distributed system to combined distributed and centralised system, and uses a crawler, cache and index to capture and organise data so as to improve the quality and speed of data searches.

The research challenges include development of improved data capture, exchange, cleaning, mapping, and quality control tools. It is anticipated that marine biology research will change so that (a) data sharing becomes a normal expectation of the scientific process, (b) online data publication becomes standard and accepted practice, (c) quality control systems (including peer-review) for web-publications will develop for scientific credibility, (d) there will be recognition of the value of on-line publication in research performance assessment, and (e) citation ranking systems for authoritative on-line publications will be established. However, the scientific and science publishing community need to be proactive to enable these changes.

Managing and distributing marine biodiversity data to meet the needs of marine conservation

J. Davies [1], S. Wilkinson [1], D. Connor [1]

[1] *Joint Nature Conservation Committee, United Kingdom*

Over the past decade, there have been a number of significant developments in the international arena that provide the mechanisms to protect and manage marine ecosystems. In particular the EC Habitats Directive requires Member States to designate and manage marine Special Areas of Conservation for listed habitats and species throughout their territorial waters, and then assess and report on their status. The Oslo and Paris Convention (OSPAR) has adopted an Annex that sets out to protect biodiversity and marine ecosystems, which has led to the identification of a list of threatened and declining habitats and species and the need to establish a network of marine protected areas for the OSPAR region. In attempting to respond to these international drivers and implement their measures in the marine environment, the UK government's statutory conservation agencies, including the Joint Nature Conservation Committee (JNCC), quickly identified a number of serious issues in relation to the availability of data and efficient mechanisms for its collation and delivery. Despite more than a hundred year's worth of effort in marine survey and monitoring in UK, it was not possible to answer simple queries such as 'What is the distribution of seagrass beds around UK' or 'Show the occurrence of species x'. Whilst any attempt to assess the status of marine benthic communities in UK seas was difficult, attempting such questions at a European scale was simply impossible. Clearly such questions are relevant at an ocean and global level and recent initiatives such as GBIF and OBIS are helping to provide answers.

The National Biodiversity Network (NBN) is a UK based initiative that sets out to improve the exchange of biodiversity information and specifically to make it available to those involved in environmental based decision making. It has now been running for approximately ten years and over that time has built up considerable experience in the field of data management and exchange. Recently JNCC has invested significant time and effort boosting the marine component of NBN. This paper will outline the technical solutions developed by the NBN to deliver marine benthic data to conservation managers and policy makers and include some recent practical examples of further developments. The issues to be addressed can be broadly broken into three areas:

- Insufficient desire or technical ability to exchange the data – this cannot be underestimated and in many cases there may be little that can be done to rectify the problem. The NBN has attempted to solve it by the development of standard tools for entering, editing and storing biodiversity data (for example the development of Marine Recorder) and also through developing reporting functionality (such as the NBN Gateway). They are also developing advice on best practice for holding and managing biodiversity data.
- The need for standard dictionaries and an ability to translate between them- different workers may know the same organism or habitat by different names but it is important that these differences are harmonised when the data are pooled. The NBN has standard species and habitat dictionaries and are currently in the process of increasing the engagement of the relevant experts to maintain these dictionaries and provide any necessary translations within them.
- Standard syntax and format for the data being exchanged – here the NBN have adopted two strategies; XML and a tightly defined tab separated text format. The former is used where there is a need for on-line transformations of data while the latter is used as a more standard route for exchanging data particularly where the volume is high and the expertise may be low.

During the past year, the UK Government has been preparing a 'State of the Seas Report' in which the JNCC have contributed to the sections on habitats and species using the marine data and the aggregation tools of the NBN. It was necessary to develop additional dynamic

tools to import the data into a geographic information system (ArcGIS) for analysis and reporting. Unfortunately this exercise highlighted the vast volume of marine data that remain inaccessible and thus present a considerable future challenge for its mobilisation.

Two new initiatives are tackling this data mobilisation issue at an international scale: Firstly, OSPAR Contracting Parties are sourcing data on 14 priority habitats which JNCC are collating and making available via the NBN on behalf of the OSPAR Commission. Secondly, in May 2004, a consortium of 12 partners from five countries in north-west Europe started an EU Interreg IIIB-funded project titled 'Development of a framework for the Mapping of European Seabed Habitats' (known as MESH). MESH will collate available marine spatial data and re-classify seabed habitat maps to a common format (EUNIS) to deliver consistent maps for north-west European seas via a web interactive GIS.

The paper will conclude with some further ideas on how marine biodiversity information systems should develop to satisfy the requirements of international conventions, agreements and legislation aiming to protect the marine environment.

Marine biodiversity conservation in Asia-Pacific: are scientific data effectively managed?

M.D. Fortes

Marine Science Institute CS, University of the Philippines, Diliman, Philippines

About half of the coastal resources in Asia-Pacific have either been lost or are severely degraded during the past half-century. The rate of degradation of the environment in the region is increasing as population increases. Within 100 km of coastline, the human population in Asia (excluding Middle East) and Oceania forms 93% of the total for the whole world (39%). If the present trend in population increase continues, coastal environmental degradation in Asia-Pacific will worsen, and the next decade will witness more alarmingly significant loss in the region's biodiversity.

Marine biodiversity loss will further result from the combined impacts of the following issues which require priority action: **climate change**, adapting to climate change and rising sea levels; **natural environmental disasters and climate variability**, requiring improving preparedness for and recovery from natural and environmental disasters; **coastal and marine resources**, protecting coastal ecosystems and biodiversity from siltation, pollution and over-exploitation; **energy**, or developing solar and renewable energy to lessen dependence on imported oil; and **tourism**, or managing tourism growth to protect the environment, cultural integrity, and biodiversity.

Capacity in the region to understand and address the issues is low, extremely limited and fragmented. The little knowledge that is available to traditional users of the sea and marine ecologists is not easily understandable, available, and accessible by those who need them most. In addition, data and information are largely descriptive (>60% of papers on seagrass worldwide), not synthetic, hence, with low predictive value useful to resource management. The results are great uncertainties when scaling up the knowledge produced locally, so that broad-scale assessment of biodiversity and its protection are faced with great difficulty. This low level of capacity prevents coastal managers from using a simple set of standards to guide all their policies and decisions.

In addressing marine biodiversity issues in Asia-Pacific, one recommendation that cross-cuts most others is the focus on data and information management issues. While there is a large amount of knowledge on the region's marine biodiversity, this has not been placed in context due to ineffective data management, making it difficult to use to assess the state of the regional coastal resource or to establish priorities for its management. Most countries in the region now realise that the key to strategy development and wise decision making on the sustainable use of biological resources and the equitable sharing of benefits, depends on having systematically organised data and information (e.g. inventories of biological resources, indicators of sustainable use, indigenous knowledge) (Reynolds and Busby 1996). Multi-stakeholder information systems ('information networks') supported by respected 'champions', address this challenge. They empower communities and individuals to collectively influence the decision-making process. Success in data management through these networks rests heavily on active **participation** of stakeholders, **cooperation** among networks of data providers and users, **support** for decision-making, and **custodianship**, selecting which institution is best suited to do it.

There is a critical and urgent need to move from unsystematic information gathering and planning to strategic solutions via data management. But unless there is a substantial change in the legislative agenda in the majority of Asia-Pacific countries, the lack of national commitments to support and encourage the development of data centers or networks and effectively integrating their results into policies for marine biodiversity conservation, will remain a major deterrent in the region's development effort.

OBIS-SEAMAP: developing a biogeographic research data commons for the conservation of marine mammals, sea birds and sea turtles.

P. Halpin [1], A. Read [1], L. Crowder [1]; B. Best [1], D. Hyrenbach [1], S. Freeman [1]

[1] *Duck University, Nicholas School of the Enviroment and Earth Scienes, Durham, USA*

In this presentation, we describe the Spatial Ecological Analysis of Marine Megavertebrate Populations (SEAMAP) program, a node of the Ocean Biogeographic Information System (OBIS) and a component of the Census of Marine Life. OBIS-SEAMAP is a digital database of geo-referenced marine mammal, seabird and sea turtle distribution and abundance data to augment the understanding of the ecology of these megavertebrates by: (1) facilitating the study of potential impacts on threatened species; (2) enhancing our ability to test hypotheses about biogeographic and biodiversity models, and (3) supporting modelling efforts to predict distributional changes in response to environmental change. To enhance the research and educational applications of this publicly available database, OBIS-SEAMAP provides a broad array of products (e.g., tabular data, maps and explicit meta-data) and services (e.g., web-based query, visualisation and analysis tools). OBIS-SEAMAP provides managers with the ability to place the habits and habitats of marine megavertebrates in an oceanographic context, which is essential to design effective conservation measures. The OBIS-SEAMAP information system integrates data from disparate perspectives (e.g., movement data, vessel-based surveys, remote sensing information) required to analyse design fisheries by catch mitigation measures, such as time-area closures and marine protected areas. Additionally, OBIS-SEAMAP provides educational products and analytical tools geared toward a broad audience of educators and students. The primary challenge to the success of this project is to stimulate the collaboration of diverse scientists, research institutions, and stakeholders. A multinational collaborative approach is essential for effective management, especially given the large spatial scales over which these highly migratory species operate. The integration of disparate data sets into a global data commons will enhance our ability to place the behaviour of these organisms in a large-scale oceanographic context and to design effective conservation measures.

A new environmental information system for tracking tagged marine organisms

D.A. Kiefer [1], F.J. O'Brien [2], M.L. Domeier [3]

[1] *University of Southern California, Department of Biological Sciences, USA*
[2] *System Science Applications, USA*
[3] *Pfleger Institute of Environmental Research, USA*

We will demonstrate new software for analyzing the data recorded by marine archival tags. This analysis system provides the tools to download data from tags and satellite imagery and integrate both types of information to obtain a more reliable estimate of the location of the tag than can be achieved by using the information from the tag alone. It also provides a dynamic, 4 dimensional (i.e. latitude, longitude, depth, and time) home for viewing this information. Our Marine Tracker software is an application of EASy (Environmental Analysis System) that has been developed specifically for marine applications by System Science Applications.

The processing of tags takes three steps, importing of tag data into a relational database and downloading imagery into a file server, computing the best path using a sophisticated search routine that we have developed, and displaying the results of the computation along with all associated tag and environmental information. Downloading and geo-referencing of tag data and imagery is automated. The computational scheme, which was written in visual basic and then plugged into the analysis system, matches tag information on the time series of temperature at the sea surface with time series information on the spatial distribution of temperature available from ocean thermal imagery. A global search of all possible paths is solved by finding the path that not only provides the best match between temperatures but also provides the most reasonable daily travel distances given the swimming speed of the tagged organism. All parameters that determine the best path are entered by the user into a simple graphical user interface. The graphical display of tracks superimposed upon imagery as well as the plots of tag data are displayed dynamically.

Trawl survey data stored at ICES

L.I. Larsen

International Council for the Exploration of the Sea, Copenhagen, Denmark

The ICES DAta TRAwl Survey database (DATRAS) comprises catch and effort data on fish from 4 research vessel surveys:

- International Bottom Trawl Survey (IBTS) - North Sea
- IBTS - Western and Southern areas.
- Baltic International Trawl Survey (BITS)
- Beam Trawl Survey (BTS) in the North Sea

These surveys were designed as abundance surveys for fish stock assessment work and this remains their main objective. However, the data are increasingly also being used for biodiversity studies and to report on occurrences of rare species.

The IBTS survey, dating back to 1965, is the longest and most extensive dataset of all survey data in ICES waters. Since 1974 the survey covers the whole North Sea, Kattegat and Skagerrak. The survey became fully standardised in the late 1980's.

The BITS survey dates back to 1991. The time series is longer, but the older data exist primarily in hard copy only. In 2000, the survey was standardised with one gear type and extended to fully cover the area of the cod stock. Abundance data for cod and flounder are provided by all seven participating countries. Many of the countries also cover all fish species found in the catches.

The Beam Trawl Survey and IBTS in the Western and Southern areas have only recently been incorporated in the ICES database and only parts of the existing data are stored at ICES; however, most counties performing these surveys have agreed to deliver data in the near future.

DATRAS data can be found at the ICES website *www.ices.dk*.

Access to data can be requested through the ICES Data Centre. Aggregate data are publicly available while access to detailed data is with the accept of all participating laboratories.

The main functionalities of the DATRAS database include:

- Submission and loading data
- Vetting data as a prerequisite for loading data
- Provision of standard abundance indices for use by ICES fish stock assessment working groups
- Extraction and aggregation of data
- Presentation of content status of the database

Building ocean profile-plankton database: progress since the first "International workshop on oceanographic biological and chemical data management"

S. Levitus [1], R. Gelfeld[1]

National Oceanic and Atmospheric Administration, World Data Center for Oceanography, Silver Spring, USA

We review the progress of the "World Ocean Database" (WOD) series since 1996. This global, integrated, scientifically quality-controlled ocean profile-plankton database now contains data from more than 55,000 cruises thanks to outstanding international cooperation. Data sources include the IOC GODAR Project, the IOC WOD project, and the IOC GTSPP project as well as usual international exchange. WOD and products based on it such as the "World Ocean Atlas" series have become standard products used by geophysics community. Some scientific results made possible by WOD are presented. We review the progress of the "World Ocean Database" (WOD) series since 1996. This global, integrated, scientifically quality-controlled ocean profile-plankton database now contains data from more than 55,000 cruises thanks to outstanding international cooperation. Data sources include the IOC GODAR Project, the IOC WOD project, and the IOC GTSPP project as well as usual international exchange. WOD and products based on it such as the "World Ocean Atlas" series have become standard products used by geophysics community. Some scientific results made possible by WOD are presented.

T.D. O'Brien

Building a global Plankton database: eight years after Hamburg 1996

T.D. O'Brien

Technology, Marine Ecosystems Division, Spring Silver, USA

At the "International Workshop on Oceanographic Biological and Chemical Data Management" (Hamburg - Germany, 1996), Linda Stathoplos and Todd O'Brien presented their initial efforts to include plankton in the World Ocean Database project. Two years later, this global collection of plankton data first became public with the release of World Ocean Database 1998. While this database followed the Hamburg 1996 metadata guidelines, it was difficult to use because of the complexity of the biological data itself, coupled with the complexity of the World Ocean Database data format. Over the next few years, this problem grew exponentially as the data and taxonomic content of this plankton database doubled. At that time, the collection contained 140,000 plankton tows and over 1.4 million plankton observations representing 4,000 unique plankton taxa. These observations were present in a broad range of not-readily-comparable units and sampling methods. While this global-coverage plankton database was now easily the largest plankton database available, it was complex and hard to use.

With the release of World Ocean Database 2001, two new supplemental fields were included with each plankton observation. A new Biological Grouping Code (BGC) allowed users to use a single indexing value to quickly access various factions of the plankton (e.g., "all copepods", "diatoms", "phytoplankton"), often composed of hundreds or even thousands of individual species. An additional Common Base-unit Value (CBV) was also included with each original measurement, providing standardized units of "per m3" for zooplankton and "per ml" for phytoplankton measurements. With addition of new plankton-specific access software, this global-coverage plankton database was now reasonably accessible to the user.

During the creation of the World Ocean Atlas 2001 plankton fields, it became evident that there were still access and usability issues. The creation of these fields represented the first focused effort at actually comparing and combining thousands of plankton measurements from different sampling methods and gear. This experience not only highlighted the necessity of complete metadata, but it also demonstrated the necessity in fully understanding and correctly translating the original plankton data and their meaning into any common database (centralized or distributed).

This presentation will review lessons learned since Hamburg 1996, and present updated guidelines for ensuring high-quality plankton databases (small or large, distributed or otherwise) and products.

The continuous Plankton recorder database: current uses and future directions

D. Stevens [1] , A. Richardson [1]

[1] *The Sir Alister Hardy Foundation for Ocean Science (SAHFOS), The Laboratory, Plymouth, UK*

The Sir Alister Hardy Foundation for Ocean Science (SAHFOS) is an international charity responsible for managing the Continuous Plankton Recorder (CPR) survey in the North Atlantic, North Sea and recently the North Pacific. The current computerised database currently contains data from 1946-2003 consisting of 192,193 CPR samples for ~ 400 species (2,275,857 taxonomic abundance entries).

This large data set has been instrumental in understanding impacts of climate change on basin-wide plankton biodiversity. For example, recent work at SAHFOS has described the northward movement by 1000km of warm water copepod communities. This work has required considerable expertise to bring together biological data from the CPR with high-resolution global physical data sets such as COADS. To facilitate this in the future, a standardised interface for extracting environmental and biological data across data provision sites would be beneficial.

Between January 2000 and June 2004, CPR data have been used by nearly 100 researchers world-wide and the demand is growing. This increased interest is linked to recent important discoveries (six papers in Nature and Science since 2002) and the increased ease of data access. It is imperative that the data are made easily accessible to further improve the use of CPR data.

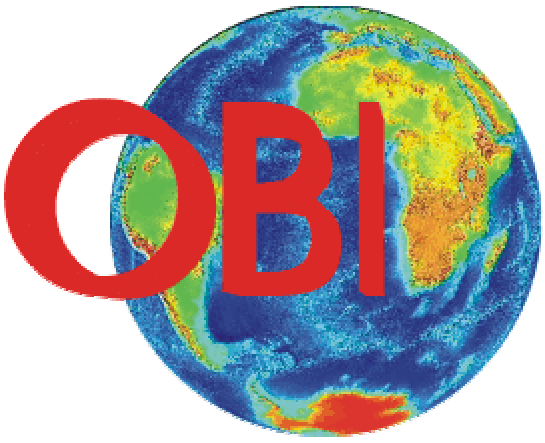
There are two main areas of current development in this regard. The first is providing improved visualisation tools for the CPR data set. Recent developments in this context include the new interactive CPR Plankton Atlas (CD), and the WinCPR North Sea browser (http://www.sahfos.org/winCPR.htm). Both provide easily interpretable graphics on the distribution of different taxa, providing researchers with valuable information quickly.

The second area of development is in allowing users easy and immediate access to available data, especially via the World Wide Web. This is already in place for two entities, Phytoplankton colour and *Calanus finmarchicus* monthly means, but needs to be developed further to allow researchers to input their own spatial and temporal requirements for all CPR taxonomic groups. By ensuring users have simple, efficient and on-demand access to CPR data we expect to see an increase in the use of the data, ensuring the continuation and expansion of the survey in the future.

Ocean Biodiversity Informatics

Hamburg, Germany: 29 November to 1 December 2004

Analysis



Using OBIS to provide reliable regional-scale estimates of population indices for marine species from research trawl surveys

R. Branton[1], D. Ricard[2]

[1] *Marine Fish Division, Bedford Institute of Oceanography, Dartmouth, N.S. Canada*
[2] *Dalhousie University, Halifax N.S., Canada*

Public provision of fisheries independent population indices for marine species from research trawl survey data although generally considered routine at laboratory and national scales is limited at regional and global scales by the timely availability of appropriate data. Furthermore, the Ocean Biogeographic Information System (OBIS), a world leading source of taxonomically authoritative marine biogeographic data products from specimen based sources like museum collections has yet to be shown to provide a similar capacity for observational sources like research trawl surveys. Recent developments in the form of a regional OBIS portal at the Bedford Institute of Oceanography at Dartmouth, Nova Scotia, Canada which provides the East Coast of North America Strategic Assessment Project (ECNASAP) trawl survey data set (276 species, 50,000 fishing sets) for the continental shelf of eastern North America from Cape Hatteras, North Carolina, USA to Cape Chidely, Labrador, Canada 1970-94 provides a practical although limited opportunity to examine how the OBIS technology might be used as a basis for the provision of regional scale estimates of population indices for marine species from research trawl surveys. A cursory examination of currently published population indices for selected groundfish species and stocks in the area covered by the ECNASAP datasets is given. A limited suite of the published indices as determined from stratified random trawl surveys (e.g. average numbers and weight per standard tow, average numbers at length per standard tow, proportions of immature fish at length) and their related reliability coefficients are examined in detail and used to map the distribution of mature vs. immature (juvenile) fish. Methods for combining indices across species are also examined. Established non-OBIS alternatives (Fishbase, ICES-BALTCOM, FAO-FIGIS and WorldFish-FIRST) for publishing trawl survey results are then examined against the range of published indices and the detailed subset. The current OBIS schema and portal are examined in detail as to how they might also be used to provide the range of indices and then compared to the non-OBIS alternatives. OBIS in its present form is clearly limited, hence a series of alternatives for extending the OBIS schema and portal(s) are devised and used as the basis for a working example on the regional OBIS portal at BIO. Alternatives that are considered include alternative uses of existing schema concepts, addition of new schema concepts, and the addition of new portal functionality using the ACON data visualization and the R statistical programming languages. These efforts if shown to be practical are expected to provide a basis for development for other regional OBIS portals and for the Future of Marine Animal Populations (FMAP) portal and as well for the Global OBIS portal at Rutgers University in New Brunswick NJ, USA, hence providing a global capacity for the provision of reliable regional scale estimates of population indices for marine species from research trawl surveys.

Spatial database model of ichthyofauna bioindicators of coastal environmental quality

J. Brenner [1], J.A. Jiménez [1]

[1] *Engineering Laboratory, Barcelona, Spain*

In the past decades the Catalonian coastal area (Mediterranean) has been subject to an intense human pressure due to the high concentration of activities and settlements. Although some of the impacts are evidents, others are not, or haven't been studied in detail in this region. In the present study we evaluate the environmental quality of the coastal/marine area through the coastal ichthyofauna diversity and distribution. We attempt to assess the potential influence of such activities on fish species identified as state indicators. This was done by identifying a number of land-originated indicators of pressure and relating them with their impact on fish. Special interest was paid to rare and special concern species that occur in Catalonian littoral waters.

We developed a spatial system that is based on a georelational database model and is implemented on a GIS environment. The GIS final application provides database maintenance, analysis and visualisation capabilities to the information system. The data model has three areas: biodiversity, physico-chemical environment and pressure/impact descriptors. The biodiversity area is based on species presence/absence occurrences and related to their ecological and conservation attributes. Physico-chemical area is based on marine environment dynamic variables. Pressure and impact descriptors are based on different data features and produced with impact extent algorithms proposed by the authors. In the three areas, features on spatial layers were attributed with behaviours as natural as they are supposed to be found in nature and/or anthropogenic environments. Fish spatial representation and ecological data were specially related using ecological domain attributes or relations on the GIS. This design provides to the database a more natural relationship among elements and environmental responses.

A major task of the study was to define species distributions along the coast with a confidence spatial resolution. The distributions were either included from reported literature or reconstructed from occurrence data using inference modelling and geostatistics. Several quality control indicators were used to define final presence/absence representation data to be included on indicator analysis. Finally, a few of the fish species that were present on the Catalonian coast were identified as indicators of environmental state. Canonical and other nonparametric analysis were performed to develop rules that led to the identification of indicators of state.

The system has two main modules: one intended for editing and analysis, and a second one for metatada visualisation. Project products include the identification of indicator species, different environmental state areas of the coastal area and areas of importance for the conservation of the icthyofauna.

An analysis of gaps in knowledge of marine biodiversity in Europe

M.J. Costello [1], C.S. Emblow [2], P. Bouchet [3], A. Legakis [4]

[1] *Leigh Marine Laboratory, University of Auckland, Warkworth, New Zealand*
[2] *Ecological Consultancy Services Ltd (EcoServe), Dublin, Ireland*
[3] *Organisms and Ecosystems, Paris, France*
[4] *National and Capodistrian University of Athens, Greece*

The organisation of information can facilitate objective analysis of gaps in knowledge and resources. Over two years, the European Register of Marine Species project compiled a list of marine species in Europe, a bibliography of marine species identification guides, surveyed species identification and taxonomic expertise, and surveyed the state of marine species collections in Europe.

The initial predictions of there being 20,000 to 25,000 marine species in Europe were a significant underestimate. In fact, the register now contains almost 30,000 species, and predictions of species not listed plus species yet to be described indicates there are probably 36,000 marine species in European seas. Rates of discovery are highest for the taxa that are small in body size, but have most species, such as the worms and copepod crustaceans.

Analysis of the bibliography of 842 identification guides found that there were fewer identification guides for southern European seas although there were more species there compared to northern Europe. There were only adequate identification guides available for fish for all of Europe's seas. New guides were especially needed for the species rich, but smaller sized, taxa, such as polychaete, oligochaete, and turbellarian worms, and harpacticoid copepods.

A database with over 600 experts in the identification of marine species (as stated by the individuals themselves), and a subset of these recognised by their peers as being taxonomic experts, was established. The average age of respondents was 47, and on average taxonomists tended to be older than ecologists. While there were generally more people identifying taxa with more species, there was no correlation between the number of taxonomists and species in their taxa. Some taxa with thousands of species have relatively few taxonomists.

A survey of holders of collections of marine species found that all collections, whether large or small, had a common problem of insufficient resources to maintain the collections as they wished. Most (64 %) of the collections were incompletely catalogued, and only 10 % had their catalogue in electronic form. New funding was therefore essential if knowledge on, and availability of, collections was to be made known on the Internet.

Gaps in marine biodiversity knowledge and resources, including species knowledge, coverage of identification guides, availability of taxonomists, and problems with marine species collections, have been identified. These gaps should be considered in both regional and national reviews of funding priorities for research and data management in Europe.

Integrating marine information and products using the Canadian Geospatial Data Infrastructure (CGDI)

P. Cousineau [1], J.R. Keeley [1]

[1] *Department of Fisheries and Oceans, Marine Environmental Data Service, Ottawa, Canada*

In the past, most large data centers have functioned independently of one another with limited reach and interoperability. This invariably led to a proliferation of formats and made data exchanges a complicated process. Furthermore, the increasing demand for integrated datasets with global coverage over long time scales poses a real challenge to data managers and software developers alike. Two-tiered architecture is still a popular design model when developing large scale application systems where performance, availability and robustness are paramount. However, as partnerships multiply so does the complexity of the resulting software needed to support them. Middleware now allows us to connect systems together more efficiently and provides a single portal of information originating from many different sources.

The Canadian Geospatial Data Infrastructure (CGDI) is the technology, standards, access systems and protocols necessary to harmonise all of Canada's geospatial databases, and make them available on the internet. Geospatial databases include: topographic maps, air photos, satellite images, nautical and aeronautical charts, census and electoral areas, forestry, soil, marine and biodiversity inventories. The CGDI builds on top of well-established standards such as ISO-TC 211 and the Open GIS Consortium (OGC) to deliver geospatial data and services to governments, scientists, industry and the general public.

The Department of Fisheries and Oceans Canada (DFO), in partnership with other marine organisations, has initiated the DFO Geoportal Project. Geoportal is a dynamic web mapping tool that allows clients to access disparate data sources from various government, private and research organisations. With minimal effort, DFO substantially improved the quantity and value of its geospatial products and services without incurring the additional costs of migrating and/or duplicating data.

This presentation highlights some of the core functionalities of Geoportal while taking a look at specific applications such as the National Registry of Aquatic Animal Health (NRAAH). The primary purpose of NRAAH is to provide Canada with the necessary infrastructure to establish and maintain all or parts of the country (i.e. zones) as being free of aquatic animal diseases of concern. The extent of the zones and their limits are managed using the Geoportal's Feature Editor (GFE), a web-based editor that allows users to add, delete and modify point, line and area features through a web feature server. The GFE also provides the ability to layer other geographical, biological, and climatological data on the same map in order to refine zone boundaries and achieve a more efficient and credible surveillance program.

MEDAS System for archiving, visualisation and validation of oceanographic data

V. Dadic [1], D. Ivankovic [1]

[1] *Institute of Oceanography and Fisheries, Oceanographic Information Service, Split, Croatia*

A marine environmental database of the Adriatic Sea (MEDAS) based on the ORACLE RDBMS has been developed. It serves for data management that includes processing of different types of oceanographic parameters (physical, chemical and biological) including archiving, validation, visualisation and presentation of data in row and graphic form. The system has capabilities of capturing historical as well as real time data. A special subsystem was developed for data validation of the classical oceanographic parameters in the framework of the MEDAR/MEDATLAS program and interface for transcoding of data from/to different formats received from eleven national and international sources.
As a result of data analysis by this subsystem many duplications, uncertainty and erroneous historical data of classical oceanographic parameters were recognised. For example, about 49.7% of BOT data was duplicated, 3.2% were outside the climatologically range, 0.7% of oceanographic stations was attributed to a wrong position, and about 17.4% of BOT data attributed as MBT data.
Based on the data analysis four different sub-regions of the Adriatic Sea with similar oceanographic properties and 41 standard oceanographic levels as suitable for climatologically analysis were recognised.
Interpolation of data on standard oceanographic levels was done by third order Newton method of finite differences modified by Reinger and Ross proposal. This method is specially recommended for oceanographic parameters (chemical parameters sampled by Niskin bottles and biological parameters sampled by special nets) with sampling on several discrete points along water column from the surface to the bottom.
Ordinary Kriging method and semi-variogram were used for spatial analysis of data. Satisfied results of climatological analysis were obtained if the distance among stations is less than 7.5 km in all regions except for Croatian channel waters and areas close to river mouths.
The MEDAS database with web interface represents a useful tool that performs quality control of the oceanographic data and various analyses, especially in the climatological domain. The web interface includes forms for data input and also has a java applet for visualisation of a GIS component. The intention is to develop a fully web-based database interface. Advances of web-based database interfaces are platform-independent and software independent database access, that only require an internet connection. The MEDAS database contains two main parts: a referral part (metadata) and a part with the measured data. Access to the referral part if free and is available on address: http://www.izor.hr/roscop/eng/. Access to the data part is under authorisation. Through data access, the MEDAS database has two main goals: providing public access to information about scientific research, hereby also improving scientific collaborations, and secondly, providing safe data storage including quality control and easy access and visualisation of data.

Predicting and understanding biogeographic ranges from occurrence records and correlated environmental data: a method-development study using clownfishes and their sea anemone hosts

D. G. Fautin[1], J. M. Guinotte [2], B.A. Maxwell [3], J. D. Bartley [2], A. Iqbal [2], R. W. Buddemeier [2]

[1] *Natural History Museum, Lawrence, USA*
[2] *Kansas Geological Survey, University of Kansas, Lawrence, USA*
[3] *Swarthmore College, Swarthmore, USA*

We illustrate the KGSMapper, a straight-forward user-controlled, web-served statistical method for predicting ranges and habitats based on georeferenced organism occurrence records and gridded environmental data. It can be used for identifying unsampled areas where the organisms may occur, for determining the environmental parameters that control the organism's occurrence, and for identifying regions where the organisms might become invasive nuisances. The KGSMapper can be used for comparison and evaluation of predictions derived by other methods, and can serve to help users choose among other modeling techniques. Using data on anemonefishes and their host sea anemones, we assess the relative robustness of predictions in the face of sparse data and data of uncertain quality, as well as how the accuracy of such predictions can be assessed or improved.

Global Ocean Data Analysis Project(GLODAP): results and data presentation through CDIAC data visualisation tools

A. Kozyr [1], R.M. Key [2]

[1] *Carbon Dioxide Information Analysis Center, Oak Ridge, USA*
[2] *Atmospheric and Oceanic Sciences, Princeton University, Princeton, USA*

The Global Ocean Data Analysis Project (GLODAP http://cdiac.ornl.gov/oceans/glodap/Glodap_home.htm) is a cooperative effort to coordinate global synthesis projects funded through NOAA, DOE and NSF as part of the Joint Global Ocean Flux Study - Synthesis and Modelling Project (JGOFS-SMP). From1990 to 1998 the global CO2 survey produced over 15 times more high-quality carbon measurements than previous survey efforts. These data were collected on more than 50 individual cruises by more than a dozen different analytical laboratories. For these data to be useful for evaluating global-scale issues (e.g. oceanic inventory of anthropogenic CO2) they must be unified into an internally consistent abd well documented data set. Wherever possible, we tried to include survey data from parallel international survey programs. The international data are extremely important for filling in ocean regions not covered by US cruises. The end result is a data set with nearly 100,000 unique sample locations. We have put a great deal of effort into evaluating the quality of the survey data and recommending adjustments where necessary. The evaluations were conducted at the basin scale starting with the Indian Ocean, then the Pacific, and finally the Atlantic Ocean.
The results of this extensive data compilation and evaluation, as well as an introduction to the unified data base via the GLODAP Electronic Atlas and web-based Live Access Server (LAS http://ferret.wrc.noaa.gov/Ferret/LAS/ferret_LAS.html), are given in this presentation. The LAS distributes both bottle and gritted GLODAP data. The GLODAP bottle data are also available as a comma-separated ASCII data files for each ocean as well as an Ocean Data View (ODV) Collection.

Integrating the OBIS schema into the information system that supports the Pew Global Shark Assessment

D. Ricard [1], R.A. Myers [1], L. Lucifora [1], F.Ferretti [1], A. Porter [1]

[1] *Dalhousie University, Department of Biology, Halifax, Canada*

Concerns about loss of biodiversity are global in scope and nature. At the same time, data collection campaigns, such as scientific surveys and fisheries log books, are often limited to certain oceanic regions and to a subset of marine species. Their use is therefore limited to specific systems and they cannot be readily used to investigate global trends for a multitude of species. The integration of a variety of data sources through a system of distributed databases is a significant step forward for researchers and practitioners studying biodiversity. A particularly alarming situation in the marine environment is the fate of elasmobranch populations (sharks, rays, skates) whose life history characteristics make them particularly vulnerable to industrial exploitation. The Pew Global Shark Assessment is part of the Census of Marine Life and its main objectives are: 1) to obtain baseline information about the historical population dynamics of elasmobranch species, 2) to determine key population parameters of elasmobranch populations, 3) to predict the outcome of various fisheries management and conservation practices and 4) to effectively communicate the findings to a wide audience. We present the preliminary information system that supports the Pew Global Shark Assessment and discuss the system's use of the OBIS schema. The integration of a diversity of data sources is essential to estimate global declining trends in shark abundance. Our system stores data from a wide variety of sources including scientific surveys, fisheries catch and effort data and voluntary surveys of shark sightings filled in by SCUBA divers. The OBIS schema offers a practical starting point to share this information with other researchers and also provides an opportunity for using data beyond the purpose for which they were originally collected. We stress the need to properly plan and design databases to allow their seamless integration with OBIS. Finally, we identify the benefits and limitations of using the current OBIS schema in our information system and provide suggestions about desirable features for subsequent versions of the schema.

Species naming curves versus species accumulation curves

G. Rosenberg

Academy of Natural Sciences, Philadelphia, USA

Species naming curves were plotted for Western Atlantic marine gastropods, and Indo-Pacific marine species, based on the data in Malacolog 3.3.2 <http://data.acnatsci.org/wasp/> and the Biotic Database of Indo-Pacific Marine Mollusks <http://data.acnatsci.org/obis/>. Year of naming was used as a proxy for year of discovery, with each species currently considered valid assigned the year that the oldest available name for it was introduced, thereby correcting for artifacts introduced by homonymy. For both faunas, the curves are almost linear since the 1840s, meaning that species have been discovered at a relative constant rate since that time. The 1840s marked a shift from works that attempted to document all known mollusks, to monographic works at the genus and family level, with a concomitant increase in rate of discovery. The lack of clear asymptotes means that it impossible to estimate how many species these faunas contain, but the true number is likely more than twice the number currently known. The curves suggest that discovery might have been accelerating since World War II, but whether this is a true increase in rate or merely lag time to synonymy is unknown.

Species naming curves differ from species accumulation curves in that estimates of sampling effort and rarity of species cannot be derived directly from the input data. Also, a species accumulation curve usually applies to a local area sampled with constant methods, whereas a species naming curve applies to a clade sampled with variable methods. Species naming curves should have a sigmoid shape, with an early exponential phase with appropriate sampling methods being developed, a linear phase with increasing global coverage and sampling effort, and a saturation phase where even massive effort has diminishing returns. This pattern is seen with global bird species. Fish and most groups of invertebrates are still in the linear phase of discovery, which makes it very difficult to estimate the total number of species.

With appropriate database design it may be possible to obtain estimates of the total number of species in various groups. A prerequisite for interoperability is a global database of names of species in groups of interest. Databases of specimen collections must be constructed to allow quantification of sampling effort and methods employed in various parts of the world. Average size of range of species and degree of overlap in areas of endemism in various groups will also be important for determining efficacy of sampling to date.

Availability of such data will allow minimum estimates of total diversity; minimum because it is not possible to predict the effect of new methods on perceptions of and discovery of diversity. For example, a shift to phylogenetic species concept from biological species concept would make a prediction based on the pattern of discovery to date an underestimate. Similarly, widespread application of molecular techniques might reveal high proportions of cryptic species among material already collected. This is yet another facet of the problem of estimating how many species there are: the lag time from collecting to naming, which can be as much as 30-50 years. Increasing globalisation of knowledge in the form of databases, interactive keys and new analytical tools will accelerate discovery of marine biodiversity.

Identify erroneous distribution data in OBIS using outlier detection techniques

W. Zhuang [1], Y. Zhang [1] , J.F. Grassle [1]

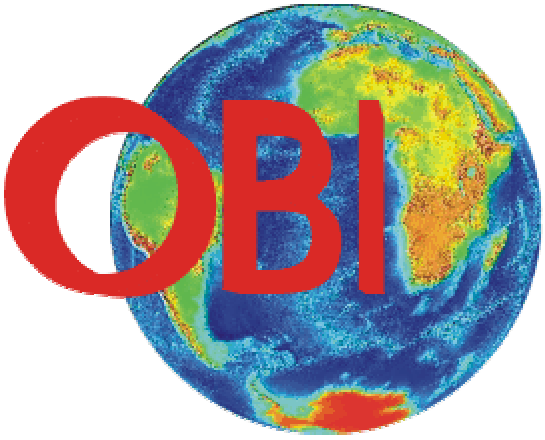
[1] *Rutgers University, Institute of Marine and Coastal Sciences, USA*

Global scale ocean biodiversity and its maintenance have not been studied extensively because worldwide species distribution databases are hard to assemble. Since 2001, Ocean Biogeographic Information System (OBIS) has been continuously growing and it now has over 5 million distribution data in its distributed network. Before further analysis is possible, the quality of the distribution data has to be evaluated and erroneous data taken out. Detecting outliers plays an important role in this process since outliers often indicate wrong data entries or inadequate georeferencing. Because we have yet to sample the world's ocean completely, a geographic outlier may be a sparse sample rather than an erroneous data point. Meanwhile, data points outlying the environmental niches may be legitimate observations since the niches are often times constructed from climatological environmental variables and thus may not reflect the conditions when the actual sampling is performed. Here we integrate these two outlier detection methods: we apply a density-based clustering algorithm (DBSCAN) to identify geographic outliers and statistical modelling to detect niche outliers. Our results have shown that with appropriate parameter settings this method could effectively detect the erroneous points in distribution data. It is not restricted to the way the data clusters and can deal with large spatial databases. The process can be easily visualised and understood. We will also analyse the relative importance of geographic and niche outliers in identifying erroneous distribution data and what environmental variables are most appropriate to construct niche models and identify niche outliers for different taxonomic groups. The general applicability of this method to data cleaning problems in biodiversity informatics will also be discussed.

Ocean Biodiversity Informatics

Hamburg, Germany: 29 November to 1 December 2004

Posters



Achieving data integration and solving interoperability problems when connecting databases to distributed systems

W. Addink [1], M. Brugman [1]

[1] *Universiteit van Amsterdam, Expert Center for Taxonomic Identification, Amsterdam, The Netherlands*

A *distributed database* is a set of databases stored on multiple computers that typically appears to applications as a single database. Consequently, an application can simultaneously access and modify the data in several databases in a network.

At present, distributed systems in biodiversity do not comply with this definition. This is because the current distributed systems consist of loosely coupled fully autonomous databases on which only distributed search and retrieval of data is possible. This is called distributed searching -- where a user sends a query to multiple sites at once, and the owners of the data return results from their own data stores.

Several architectures for distributed systems in biodiversity are already implemented or in development like the GBIF network, BioCASE, MaNIS, OBIS and BioMOBY. These systems show a lot of similarities but also differences. All current systems consider the autonomy of the datasources as vitally important, however the autonomy of data receivers (end users) is taken less into account. Data receivers are autonomous in the sense that they typically have different information needs and vary in the ways they perceive their particular domain of interest. For them a “one integrated schema fits all” is not satisfactory [1]. Also, end users often do have little need for ‘raw data’ like specimen or observations but instead need interpretations from calculated results based on this raw data.

Most current networks use a common concept schema like ABCD or Darwin Core. Less often implemented are standardisation of data itself, and things like duplicate elimination and missing value substitution. This leaves some data heterogeneity existing and therefore some interoperability problems still exist in the current systems. With the currently implemented distributed systems in biodiversity, development is moving into the direction of webservices.

Webservices seems to be one of the best solutions at current to achieve interoperability. Being application-centric, webservices are more scalable then a human-centric webportal approach, can be very language and system independent and are easier to establish than for instance a GRID network. Webservices have the promise to provide a new level of interoperability. Its interoperability goals are to provide seamless and automatic connections from one application to another. SOAP, WSDL, and UDDI protocols define a self-describing way to discover and call a method in a software application – regardless of location or platform. However developers have to interpret the meaning in parts of the specifications which still allows interoperability problems to creep in at the discovery, definition and request/response mechanisms [2].

References:
[1] Patrick Ziegler. USER-Specific Semantic Integration of Heterogeneous Data: What Remains to be Done? Technical Report ifi-2004.01.
[2] Frank Cohen. Understanding Web service interoperability. Issuses in integrating multiple vendor Web services implementations, February 2002.

BioFOX Zoo: A software for storing and analysing zooplankton sampling data

A. Lyakh[1], I. Agarkova-Lyakh[2]

[1] *Ukraine National Academy of Sciences, Institute of Biology of the Southern Seas, Sevastopol, Ukraine,*
[2] *Tavrical National University, Department of Geography, Saki, Ukraine*

There are many Taxonomical Information Systems but few systems exist for analysing biological data. The very popular and universal analytical software, widely used by biologists in developed countries, is MS Excel. But universal doesn't mean convenient for use. The main goal of BioFOX Zoo project is to create software especially designed to store, retrieve and analyse zooplankton sampling data.

The zooplankton or phytoplankton sampling data are the data obtained by microscopic observation of individuals in a sample droplet. These data include organism taxonomy, their dimensions and quantity. Zooplankton sampling data also include organisms age (adult, larvae, egg), sex (only for adults and final stages of larvae development), and stage of larvae development. In addition, to calculate organisms abundance and biomass, the catching information is needed: a trawled water volume, a sample volume , etc.

BioFOX Zoo, using primary and catching data, calculates abundance and biomass of each taxon found in the sample. For each sample the dominated species and dominated high taxa are determined, and Shannon and Simpson biodiversity indexes are found. Also BioFOX Zoo analyses dynamics of taxa in a group of samples, which is combined by some criteria. Software calculates the minimal, maximal and average abundance and biomass of each taxon in the group, and calculates indexes of individual popularity and density. In addition BioFOX Zoo has a function of Taxonomic Information Systems. Users can manage a list of taxa, which is used to quick enter the organism taxonomical rank, and coefficients for translating an organism dimension to weight.

BioFOX Zoo is approved in the Institute of Biology of the Southern Seas (Sevastopol, Ukraine) and "Kerzhinskiy" preserve (Nizhniy Novgorod, Russia).

Micro-Mar: a marine prokaryotes database to correlate habitat with taxonomy

J.C. Alba-Casado [1], R. Pushker [1], F. Rodríguez-Valera [1]

[1] *Universidad Miguel Hernández, Division de Microbiologia, San Juan de Alicante, Spain*

The global ocean ecosystem is deeply dependent on the activity of its large population of prokaryotes. However, the small size relatively diluted environment and reluctance to be grown in pure cultures makes marine prokaryotes also the less known group of microbes. PCR based approaches have provided an invaluable wealth of information about marine prokaryotes. Most of the DNA databases such as GenBank provide very little information to correlate ecological or oceanographic parameters with the biodiversity estimated by DNA sequences.

The Micro-Mar database is a marine prokaryotes database to correlate habitat with taxonomy and has been created to collect DNA diversity data of marine prokaryotes to make it amenable for ecological and biogeographical analysis. The database currently includes over 2500 entries corresponding to 16S, ITS or 23S rRNA sequences of marine prokaryotes, together with the available geographical information (geographical origin, latitude and longitude) and ecological information (depth, temperature, salinity, currents, chemical and ecological properties). rRNA sequences were chosen because they are most commonly used for taxonomic purposes but other sequences such as gyrases etc. will be incorporated in the near future. Ideally each entry comes from a pure culture but assuming the low cultivability of marine prokaryotes, many entries are either PCR or cloning products.

The web page has a search option to query the database for various parameters like GenBank ID, geographical origin, latitude, longitude, depth, temperature, salinity, taxonomic unit, similarity percentage, phylogroup, etc. In the taxonomy field, the best BLAST hit to any marine prokaryotic sequence and the best BLAST hit to a taxonomic unit (more often a pure culture) with similarity percentage is indicated. It also has a Geographic Information System (GIS) where a user can get information regarding various ecological and biogeographical parameters by choosing a particular location on the world map. An online BLAST search against Micro-Mar database can also be done.

The database is updated every week to include the most recent marine prokaryotic sequences. Although some conclusion can already be drawn based on the current database entries, the available information on ecological parameters is still limited. As more entries are incorporated we will be able to correlate accurately the bacterial biodiversity with geographical and ecological factors.

A taxonomic and biogeographic information system of marine species in the southern North Sea developed by Flanders Marine Institute

W. Appeltans [1], E. Vanden Berghe [1] and J. Mees [1]

[1] *Flanders Marine Institute (VLIZ), Oostende, Belgium*

Aphia is a register of marine species that the Marine Data and Information Centre of the Flanders Marine Institute (VLIZ) has built and maintains since its establishment nearly five years ago. This species register (Aphia) serves as a taxonomic backbone in the management of biological data obtained by our own efforts and those coming from various national and international projects and databases. Currently, Aphia holds 20,000 taxa, including 10,000 species names of which 2,200 are relegated to synonymy. Remaining names are those of higher hierarchical levels from the systematic classification. Besides pure taxonomy, Aphia contains about 3,000 vernacular names (mainly Dutch, English, French and German) and 4,000 notes regarding species biology, ecology, distribution, habitat, taxonomy and systematics among others. All the information is referred to its source, which can be a publication, an expert or a database.
More recently, VLIZ started storing distribution records for the Southern North Sea. We now have nearly 40,000 distribution records for 400 localities. By screening literature, from recent works to those dating back to the mid nineteenth century, we intend to get a more complete overview of our current and historical biodiversity. In addition, documenting nomenclature changes, recording synonymy and storing every species record will help in correctly interpreting old and grey literature.
The structure of the database corresponds to the international formats of both the European Register of Marine Species (ERMS) and the International Taxonomic Information System (ITIS). It will be possible to give taxonomic input to ITIS and ERMS. Our biogeographic system is now already linked with and provides data to the European node of the global Ocean Biogeographic Information System (EurOBIS).
Maintaining our species register and storing biogeographic information is a work in progress, which falls within VLIZ's core activities of documenting and integrating marine biodiversity data.

A marine observatory network for the Basque Country (n.e. of Spain)

A. Aramburu [1], A. Borja [2], I. Galparsoro [2], Y. Sagarminaga [2]

[1] *Basque Government, Spain*
[2] *AZTI Foundation, Spain*

A Marine Biodiversity Observatory network for the Basque continental shelf and estuaries is under development with funds from the Basque Country Government (Northern Spain). This project seeks to meet some of the requirements put forward by the Basque Environmental Framework Programme (REF), which identifies the need for a system that provides rapid access to marine data and their analyses for territory management, scientific studies and conservation purposes.

Up to now, many small projects have been carried out by different institutions, companies and universities all along the basque littoral. All this work has produced a great amount of multidisciplinary data, which is actually dispersed and under different formats.

Thus, the first task of the Marine Observatory has been to integrate several different datasets, such as: (i) littoral land cover cartography, of different years and resolutions, to estimate the anthropic pressure evolution over the littoral; (ii) general bathymetry and bathymetry of specific sites, in order to see morphological changes; (iii) geology and geomorphology, including seabed cartography and habitat mapping; (iv) biology and fisheries; (v) hydrography obtained from oceanometeorolgical measure stations, buoys and CTDs; (vi) meteorology; (vii) pollutants (organic and inorganic); (vii) uses and management, including submarine installation characteristics, dredging sites and quantities, dumping sites, etc.

All these data are being standardised and integrated into a georeferenced database together with standardised metadata. Once settled up the outcoming integrated dataset will be the basis for a public environmental information service hosted by the Basque government. This service aims to be a valuable tool for managers, marine scientists, coastal users, and general public.

Through this paper we will present the general characteristics of the project and the tasks and results achieved so far.

Thermal fronts in the Baltic Sea and their influence on spatial distribution of ocean colour phenomena

K. Bradtke [1], L. Szymanek [1], A. Krężel [1]

[1] *University of Gdańsk,Department of Physical Oceanography, Gdynia, Poland*

Spatial and time qualitative distribution of phytoplankton in the Baltic Sea results, among others, from different temperature requirements of individual taxa. Spatial temperature distribution depends primarily on light conditions. However, dynamic processes are also of great importance since they cause a local advection of different water types. At the border of different water temperatures, frontal zones are formed.
The main goal of the investigation was to analyse the frequency of frontal zones, occurrence in temperature field and on sea colour maps in the Baltic Sea, and their time persistence (stability) on the basis of satellite images (AVHRR- and SeaWiFS-derived data). To determine the fronts, one-dimensional Webster method (Davis 1986) was modified to analyse two-dimensional maps obtained from satellite maps. The method appeared to be less sensitive to local noises than the standard techniques of the edge detection type.
The fronts were observed mainly as a result of coastal up welling and river inflows. Since the waters from land run-off that are raised to the surface by up welling, frequently differ in temperature, salinity and other physico-chemical properties from the surrounding waters, the frontal zones could constitute boundaries separating the habitats occupied by different communities of organisms. To explore the interactions between physical and biological oceanic processes, the occurrence conformity of frontal zones on the sea colour maps (SeaWiFS-derived data) with thermal fronts was investigated. The results obtained indicated that thermal fronts are usually a sufficient barrier to prevent waters of different properties from mixing. However, a more complicated image of front lines on the sea colour maps also indicates an influence of other factors on spatial distribution of ocean colour phenomena.

A next step in the emergence of self-funded OBIS regional nodes: industry sponsored data product development on the CMB-BIO internet portal

R.M. Branton [1], L. Van Guelpen [2]

[1] *Marine Fish Division, Bedford Institute of Oceanography, Dartmouth, Canada*
[2] *Atlantic Reference Centre, Huntsman Marine Science Centre, New Brunswick, Canada*

Environmental assessments for offshore industrial activity use marine biological data from published literature, industry projects and/or from government archives, which tend to be complex and subject to local differences. The Centre for Marine Biodiversity (CMB) and Bedford Institute of Oceanography (BIO) together are developing and promoting the CMB-BIO Internet Portal as a collaborative facility for authoritative marine biological data from the North Atlantic, Arctic and North Pacific Oceans and be integral to the International Census of Marine Life (CoML) - Ocean Biogeographic Information System (OBIS) and Canadian Geo-spatial Data Infrastructure (CGDI) initiatives. BIO is funding basic its operations following OBIS and CGDI metadata standards, whereas CMB has designated it as the focal point for its members' data. Multi-disciplinary initiatives such as the CMB Discovery Corridor and the CoML Gulf of Maine Ocean Data Partnership would also use it thus providing a one-stop self-serve shop to discover, assess and acquire marine biogeographic data. BIO has installed the required computing infrastructure and CMB has facilitated funding of basic bio-geographic datasets including: species lists, specimen collections, research trawl surveys, and lists of rare and endangered species. An application to support standardisation of species names from the multiple data sources served by the portal is also in preparation. There are many more datasets that could be added, in terms of marine fin fishes these include: ichthyoplankton surveys, tagging studies, commercial landing statistics, fishers' logbooks, port samples, at-sea observer samples and industry surveys. Start-up funding however is exhausted; hence proposals for industry funding calling for an inventory of BIO's marine finfish data holdings that would be placed on the portal and used by a panel of environmental consultants to generate a prioritised work plan of future activity. The proposal also includes training seminars for environmental consultants and public outreach materials such as fishers' atlases. Other data, particularly those collected by private environmental consultants and conservation groups would be solicited at all stages. Although additional data sources would require planning and funding, the cost for data already in or close to the CGDI and OBIS standards would be minimal. Open standards will increase end-user confidence in the data and enable cross-jurisdictional mapping and analyses not possible with current ad-hoc methods. In the case of government data, users could expect that standards based access will allow them to make better and more efficient use of their own resources and of government staff's time. Continuing reliance on ad-hoc arrangements, particularly as many of the present generation of government scientists retire are clearly bound to deteriorate. The CMB-BIO Internet portal is being developed because offers as a clear solution for these problems as well as a means of levelling the playing field for environmental consultants.

Biodiversity informatics and Indian Ocean: challenges and potentials

V. Chavan [1], M.V.M. Wafar [2], S. Krishnan [1]

[1] *Information Division, National Chemical Laboratory, Pune, India*
[2] *Biological Oceanography, National Institute of Oceanography, Goa, India*

The Indian Ocean is the third largest ocean in the world. It encompasses 21% of the world's sea area and 14% of the earth's surface. The Indian Ocean region includes 36 littoral and 11 hinterland states, together accounting for 30% of the earth's population, much of which lives in developing nations.

The Indian Ocean region is a significant contributor to the production of living marine resources with an estimated annual yield of 8 million tons of captures fisheries and 23 million tons of culture fisheries, equivalent respectively to 10 and 90% of the world production. The tropical countries of Indian Ocean are also known for the high coastal and marine biodiversity which they sustain.

Despite the high potential biodiversity, the Indian Ocean region is the least studied, and hence very little is known about the state of its biodiversity. Lack of taxonomic expertise and hesitation to engage in cataloguing biodiversity are the stumbling blocks that constrain widening of our knowledge basis. This lacuna is compounded by a lack of communication and information exchange between the regional countries and, when such possibilities exist, a reluctance to share knowledge.

In this paper we discuss the challenges of documenting biodiversity information about Indian Ocean. However, we also elaborate on the plans in offing's. While doing so we would discuss the potentials of Information and Communication technology in documenting Indian Ocean biodiversity information. With the recent initiatives such as Ocean Biogeographic Information System (OBIS) component for Indian Ocean, it is feasible to develop a mechanism that has the ability to synergise networking of physical components and human minds to acquire, store, analyse, predict, forecast and disseminate state of biodiversity within Indian Ocean region. It is our belief that emerging discipline such as biodiversity informatics would contribute significantly to the understanding of the past and the present, in order to learn about the future of life in the Indian Ocean. Indian Ocean component of OBIS will become a prime provider of biodiversity information on the Indian Ocean, and make this information available in a multidimensional geographic context; promote communication and awareness to user groups at all levels, using the appropriate information tools; and enable informed decision making process leading to a sustainable use of natural resources.

An integrated approach to managing data, graphics and reports

G. Coleman

Australian Institute of Marine Sciene, Australia

The Australian Institute of Marine Science Long Term Monitoring Team has collected ecological data from reefs spanning the length and breadth of the Great Barrier Reef for over twenty years. For its entire history the project has had the luxury of a full time data manager. The result of this has been a well-structured database with easy to use data entry screens. In recent times, this project has been taking digital photographs to complement the scientific data collected. The data entry processes have been altered to allow the photographs to be attached to the relevant data. Time and spatial information is entered for each dive. Scientific results and images are associated with each dive. Images dragged and dropped onto the data entry screen are associated with the current dive. The images can be dragged from windows explorer or any of the various thumbnail viewing packages available (ACDSee, ThumbsPlus, NikonView etc). For most photos, Captions are entered into the database. Keywords and taxonomic names are extracted from these captions for searching purposes. Other projects using a modified version of this image database will be discussed briefly. For example a project more focused on taxonomy does not have space for captions but has pull down species lists.

Results of the project are reported on the Internet and an annual CD based report is generated from the information on the Internet. Data is summarised, analysed and plotted using programs written in S-PLUS. The resulting plots are stored in the database. A custom written program allows users to view the plots, interrogate the data and write a few paragraphs about each reef. The data entry has a spell checker and the ability to find and italicise species and genus names. These paragraphs are stored in the database. The photographs, plots, data and text are all extracted from the database and collated into an attractive and informative web page for each reef. These “reef pages” are generated by a Java server-side process. This simplifies maintenance of the web site. Changes to the database are automatically reflected on the Internet. A change to a template file will change the look of every reef page. A snapshot is taken from the web site for the annual CD. This process utilises a web crawler called “Teleport Pro”.

One advantage of creating the report directly from the database is that there is little delay formatting the document after it is written. Another advantage is the ability to capture information about the photographs. It is often difficult to find the time to catalog and tag images. Since the photos and captions are used in the report, there is more incentive to catalog the photos and write the captions. The key to the process is in the
data entry. Data entry programs must be designed to allow maximum information captured for minimal effort. This is achieved by the integrated approach which avoids double handling

Fisheries and ecosystems information system: a tool for the implementation of the ecosystem approach to mediterranean fisheries.

S.R. Coppola [1], V. Giacalone [2] , T. Bahri [2]

[1] *Food and Agricultural Organisation of the United Nations, Rome, Italy,*
[2] *Food and Agricultural Organisation of the United Nations - MedSudMed, Mazara del Vallo, Italy*

The Fisheries and Ecosystems Information System (FEIS) is the information component of the FAO regional Project MedSudMed which operates in the Strait of Sicily and undertakes research activities in cooperation with the four participating countries (Malta, Tunisia, Italy, and Libya). Fisheries resources are shared by some of these countries; however, no clear information is available to date on the spatial distribution of the main target stocks, nor on the interaction between fishery resources and the environment, and therefore no concerted management measures can be planned and eventually adopted at regional level. The research topics deal with spatial distribution of demersal and pelagic resources in relation to environmental parameters. Regional stock assessment and stock identification were also identified as priorities. Data are provided by joint cooperative trawl and acoustic surveys following a standardised sampling protocol and involving all the participating institutes.

The FEIS is designed to support scientific communities and countries in the development of a system for monitoring fisheries resources and ecosystems through the organisation and standardisation of the data collected in the MedSudMed Project area complemented by accessory data and information obtainable from the web and duly structured to enable joint processing. The aim of this system is to act as an analytical support tool to study marine ecosystems, natural phenomena and fisheries, by providing a framework within which the project participants can share information and data. It will also enable the creation of applications to support the management, analysis and representation of data relative to fishery resources and their environment. The key information concerns biological aspects of the resources, environmental parameters, fishery statistics and accessory data. The primary motivation is to standardise, aggregate and analyse the data, and enable them to be exported on a Geographical Information System (GIS). The primary users and beneficiaries would be the partner countries of MedSudMed, as well as the scientific institutions of the General Fisheries Commission for the Mediterranean (GFCM) and FAO.

The system structure is constituted by i) a **corporate database** which contains data and basic information, maps, documents and other from both the participating Institutes and public sources, ii) **database applications** including, among others, several modules (Data Entry Module to assist the user in the input of data into the corporate Database; Data Management Module for the management of the database; Data Analysis Module for the data queries, analysis and data export), iii) a **web Interface** which allows visibility and accessibility on the web of agreed information contained in the system.

The FEIS is destined to host the common pool of data of countries participating in the MedSudMed Project and considered useful for the fulfilment of the objectives of the Project's activities. The database contains information that institutes wish to share, with different levels of accessibility according to the confidentiality of the data. An interface is foreseen to extract information from already existing databases or software developed by the participating institutes.

The effect of the hemispheric climatic oscillations on the Adriatic ichthyofauna

J. Dulčić [1], B. Grbec [1], L. Lipej [2], G. Beg-Paklar [1], N. Supić [3] & A. Smirčić [4]

[1] *Institute of Oceanography and Fisheries, Laboratory of Ichthyology and Coastal Fishery, Split, Croatia*
[2] Marine Biological Station Piran, National Institute of Biology, Piran, Slovenia
[3] Ruder Boskovic Institute, Center of Marine Research, Rovinj Division, Croatia
[4] Hydrographic Institute of the Republic of Croatia, Croatia

The hypothesis of an expansive northward movement of thermophilic species and changes in marine biodiversity is nowadays supported with numerous records of fish species previously characteristic to the more southern areas. Obviously, this is happening in the Adriatic sea as well, where numerous new species in the area or in the northern sectors were recorded during the last thirty years. Good correlation between mean annual air and sea surface temperature and yearly total number of specimens, as well as between annual sea surface temperature and species richness is obtained for the period 1973-2003. The variations in Adriatic temperature conditions correlate well with the North Atlantic Index (NAO) variations showing that local temperature changes at least partly result from the hemispheric one. Variations in sea surface temperature conditions mostly result from heat flux exchanges on the air-sea interface, and since net heat flux is under NAO influence there is no doubt that recent changes of Adriatic ichthyofauna are partly controlled by hemispheric climate changes. Distribution of warm-water fish records is influenced by overall cyclonic circulation in the Adriatic Sea.

Stock identification of European anchovy (*Engraulis encrasicolus*) using morphometric and meristic characters

Z. Erdoğan[1], C. Turan[2], H.T. Koç[1]

[1] *University of Balikesir, Department of Biology, Balikesir, Turkey*
[2] *University of Mustafa Kemal,Hatay, Turkey*

Morphologic differentiation among stocks of European anchovy, Engraulis encrasicolus, throughout the Black, Marmara and Aegean Seas, was investigated using morphometric and meristic characters. Degree of morphologic differentiation among populations of anchovy was evaluated with the Truss network system using Principal Components Analysis and Discriminant Function Analysis. The Aegean Sea (Izmir Bay) and The Marmara Sea (Bandırma Bay) samples were the most isolated samples from the others for morphometric and meristic characters, respectively, which may indicate existence of a distinguishable anchovy stock in the areas. Multivariate analyses indicated that differences among samples based on morphometric characters seemed to be associated with body height. In meristic analyses, the highest contribution to multivariate analyses were associated with the number of gill rakers.
Keywords: European anchovy; meristics; morphometrics, stock identification; Engraulis encrasicolus

BIOCEAN – a new database for deep-sea benthic ecological data

M.C. Fabri [1], J. Galeron [1], G. Maudire [1]

[1] *Institut Francais de Recherche pour l'Exploitation de la Mer, Centre Brest, Plouzane, France*

The Biocean database was designed to gather the extremely large volume of data collected from different deep-sea ecosystem studies conducted by Ifremer's departement « Environnement Profond ». This database comes in a five application package; two of them are used onboard research vessels to collect operational data, while the others are used to link with a core database back on land. The latter are used to: 1) manage the taxonomic nomenclature, 2) monitor the identification of faunal collections and 3) add or extract data from the database. Biocean was designed to facilitate ecosystem studies in the deep sea. It represents an important new resource for deep-sea ecologists, will have wide applications in biogeography, and biodiversity studies.

Over the past 30 years, huge numbers of samples have been collected on deep-sea cruises dedicated to benthic community studies. These faunal samples are usually sieved through a series of mesh sizes and separated into taxonomic groups at the phylum, class or order level before being dispatched to taxonomists for identification at the most precise taxonomic level. The occurrence of species that are new to science is very frequent. Species-level identification and taxonomic description may takes five to ten years after a cruise. The Biocean database was created to help manage the several steps required between sampling and species identification.

In addition to high faunal diversity, the deep-sea environment can also be characterised by very strong habitat gradients, and close links between species distribution and physical or chemical parameters. For examples, chemosynthetic communities (such as those found on hydrothermal vents or cold seeps) live in fragmented ecosystems distributed either along ocean ridges or on active or passive margins. The environmental factors characterising these ecosystems represent a wide range of data, which can be stored in the Biocean database alongside faunal data.

The Biocean database contains data covering 30 years of French deep sea oceanographic research, and include 78 cruises, 223 dives and 371 moorings. All samples and measurements from these cruises are geographically referenced.
Overall 3055 biological samples, 383 water samples, 210 sediment samples, 445 in situ temperature measurements, 64 continuous measurements datafiles and 4262 photos are indexed. Post-treatment data (e.g. chemical factors or taxonomical identifications) represent 3513 values, 4476 deep-sea benthic species and 69 datafiles.

A mediterranean Geo-database of cetacean strandings (MEDACES): an implement for research and conservation*

D. Garnería [1], M. Fernández [1], J.A. Raga [1]

[1] *Cavanilles Institute of Biodiversity and Evolutive Biology, University of Valencia, Marine Zoology Unit, Valencia, Spain*

A marine mammal stranding is usually a natural event as a consequence of starvation, illness or, simply, aging. Cetacean stranding is a well-known phenomenon throughout the ages. Stranded cetaceans are a very important, and sometimes the only, source of information about the anatomy, ecology, genetics, pathology, parasitology or toxicology of the species in question. Recently, studies on cetaceans have been seen extended thanks to the employment of modern technologies, e.g. molecular tools that allow small parts of these animals to be used for genetic studies. In addition, monitoring cetacean strandings over an area can, on occasions, provide information on the health status of populations and identify important problems, such as epizootics resulting in mass mortality. In endangered species, rehabilitation is very important for the survival of the species. Rehabilitation efforts are directly linked to the human concern for conservation of cetaceans and of the marine biodiversity, in general. In a number of countries, where the concern for the protection of cetaceans is more apparent, stranding networks have been developed over the last decades of the 20th century to gain more knowledge on the biology and conservation of these animals. The monitoring of cetacean strandings in the Mediterranean countries is very heterogeneous. In some countries there are no active groups devoted to cetacean research, whereas in others there have been well-established networks, some of them governmental, for at least 10 years. A Mediterranean Database of the Cetacean Strandings (MEDACES) has been set-up at the University of Valencia (Spain) to co-ordinate all national and regional efforts for riparian countries. This database was created under the Barcelona Convention. Recently, MEDACES has been expanded to cover regions adjacent to the Mediterranean, as the Black Sea and the contiguous Atlantic waters as defined in the Agreement on the Conservation of Cetaceans of the Black Sea, Mediterranean Sea and contiguous Atlantic area (ACCOBAMS). Distribution maps will be produced using a Geographical Information System (GIS). Annual or biannual reports will be accessible online through a web site (http://medaces.uv.es). Given the differences in the level of coverage of cetacean strandings between the Mediterranean countries, two levels of data collection are proposed. At the first level, basic information will collected and this will be common to all the stranding networks. The second level concerns more complex data and may vary as a function of the logistic and technical possibilities of each country. MEDACES will gather this information on cetaceans strandings provided by a National Focal Point in every country, through the Regional Activity Center for Specially Protected Areas (RAC/SPA). An agreement between the ACCOBAMS and the RAC-SPA has been developed to co-ordinate the information of cetacean strandings of the ACCOBAMS countries and the MEDACES. In conclusion, cetacean strandings provide not only relevant scientific information but also the key to their conservation. The establishment of a national cetacean stranding network does not necessarily demand high expenses, but it must be adapted to the economic circumstances of every participating country. In the Mediterranean context, the Barcelona Convention represents a nucleus of coordination and diffusion of information about the strandings through the MEDACES. This database pretends to become a frame where researchers and people interested in cetacean conservation may exchange data and information.

*Supported by the Ministry of Environment of Spain.

Adriatic-possible influence of climate oscillations over the northern hemisphere

B. Grbec [1], J. Dulčić [2], M. Morović [1]

[1] *Institute of Oceanography and Fisheries, Split, Croatia*
[2] *Institute of Oceanography and Fisheries, Laboratory of Ichthyology and Coastal Fishery, Split, Croatia*

In order to establish the connection between the hydroclimate variables and pelagic species, year-to-year fluctuations of small pelagic fish landings in the eastern Adriatic coast were compared to climatic fluctuations over the Northern Hemisphere and to salinity fluctuations in the Adriatic. Using this approach, basic climatic oscillations were determined for both hydroclimate and biological data. The main climate oscillation period was approximately 80 years and an interrelation between climatic fluctuations over the Northern Hemisphere and small pelagic fish landing data was found. The results suggest a linkage between the advection of Levantine Intermediate Water, which is controlled by the pressure distribution over the wider area, and fish abundance in the Adriatic Sea.

Data management and communications (DMAC) for the U.S. integrated ocean observing system (IOOS)

S. Hankin [1], L. Dantzler [2], R. Cohen [2], F. Grassle [3]

[1] *National Oceanic and Atmospheric Administration, Pacific Marine Environmental Laboratory, USA*
[2] *Ocean.US Office / National Oceanic and Atmospheric Administration, National Oceanographic Data Center, USA*
[3] *Rutgers University, Institute of Marine and Coastal Sciences, USA*

Data management and communications within the marine environment present great challenges due in equal parts to the variety and complexity of the observations that are involved; and the complex relationships among community participants. At present there is no coherent cyberinfrastructure that effectively integrates these data streams across organisations, disciplines, and spatial and temporal scales. The resulting lack of integration of data denies society important benefits, such as improved climate forecasts and more effective protection of coastal marine ecosystems.

The US Congress has requested that the US marine science communities come together to plan, design, and implement a sustained Integrated Ocean Observing System (IOOS). Central to the vision of the US IOOS is a Data Management and Communications (DMAC) Subsystem that joins Federal, regional, state, municipal, academic and commercial partners in a seamless data sharing framework. The DMAC framework in the US must integrate seamlessly with its international counterparts, and with data management frameworks employed by related disciplines. Designing the DMAC Subsystem goes beyond solving problems of software engineering; the most demanding aspects of the solutions lie in community behaviour.

Ocean.US, the IOOS national office, established the Data Management and Communications Steering Committee (DMAC-SC) in the spring of 2002. In May 2004 Ocean.US published the plan developed by the DMAC-SC, "**Data Management and Communications Plan for Research and Operational Integrated Ocean Observing Systems, I. Interoperable Data Discovery, Access, and Archive"** (http://dmac.ocean.us/). This poster presents an overview of the DMAC Plan.

Web mapping services for biological data

F. Hernandez [1], B. Vanhoorne [1], R. T'Jampens [1], E. Vanden Berghe [1]

[1] *Flanders Marine Institute (VLIZ), OOSTENDE, Belgium*

When marine data is being shared online through the web, often the need rises for interactive mapping tools. Whether it is for displaying data, or for providing geographical selection options for this data, whether for biological parameters (distribution records) or for abiotic, oceanographic parameters: a nice, interactive map gives a clear added value to the application. At VLIZ we have several of these online databases running, and through the last few years we have developed a certain amount of program code. This paper describes our experiences, the tools developed so far, and how they will be made available to the larger scientific community.

After some experiments with clickable bitmaps we chose for an SVG based technology, because of some clear benefits: SVG is an W3C specification, it is Open source, it is free and it has very good scripting possibilities. However, when using SVG for something more than a simple, one file application, several problems arise. The generation and documentation of base maps, the need for different geographical projection systems and the performance issues when using higher resolution maps are some of the problems that need to be solved.

Our solution to these problems is the 'Svg Sniplet Server'. What we basically did is to upload our charts into an SQL database, and install a webservice application that queries the database, projects the retrieved coordinates and generates small pieces of SVG code (called sniplets) that can be combined to form one SVG file. The paper further explains how you can take benefit of this service for you own applications.

In a next phase we want to investigate the possible interaction with other standards like GML and other ISO 191xx standards, and the forthcoming MarineXML standard.

The BODC taxonomic browser – a powerful search tool to locate taxonomic information

M. Hughes [1], R. Lowry [1]

British Oceanographic Data Centre, Prenton, United Kingdom

The British Oceanographic Data Centre (BODC) stores a wide range of data sets, covering physical, chemical and biological information. To ensure reliable archiving and retrieval of this data, every record held at BODC is described by an entry in a parameter dictionary. The dictionary is constantly expanding and currently has over 16,000 entries. Of these entries, about 11,000 refer to biological organisms.

The Taxonomic Browser is a web-based search tool which eases navigation through these biological entries. On entering a scientific name, the Taxonomic Browser will retrieve every dictionary entry for the search term and for all organisms taxonomically related to the search term. This works even if the actual search term is not present in the BODC dictionary. For example, a search for “Cetacea” (an Order of marine mammals) will return every Genus and Species of dolphin that is held in the dictionary even though the word “Cetacea” does not occur in any part of the dictionary.

This type of search works because the Taxonomic Browser uses the Integrated Taxonomic Information System (ITIS, http://www.itis.usda.gov) as a hierarchical framework for its own searching. For a given search term, the Taxonomic Browser scans a local copy of the ITIS database for all organisms related to the search term. A list of organisms is produced which is then checked against the dictionary. All this can be done in a very short time via mappings of the BODC parameters to a unique ITIS serial number for each of their taxonomic entities.

ITIS is used for a number of reasons:

- It provides an authoritative and up-to-date source of taxonomic information with which we can standardise our taxonomic terms (e.g. adjusting spellings, following the ITIS taxonomic hierarchy).
- It is free and easy to download a local copy of the entire ITIS database and update this copy as necessary.
- ITIS provides tables of common names and synonyms, and allows linkage between invalid taxonomic terms and their valid equivalents.

The Taxonomic Browser includes a number of extra features that make use of the ITIS database:

- Search by common name – uses a table of ITIS vernaculars to interchange between common and scientific names.
- Search by ITIS taxonomic serial number (TSN) – searches for BODC parameters corresponding to a particular TSN.
- Search by BODC parameter code – searches for dictionary entries corresponding to a unique BODC parameter code.
- Modifies a search to find all synonymous taxonomic terms – uses a table of ITIS synonyms to find all valid equivalents for a taxonomically invalid search term and vice versa.

Marine biodiversity data in Peru

A. Kameya [1], M. Quiñe [1], E. Delgado [1], S. Sánchez [1], A. Chipollini [1]

1 Instituto del Mar del Peru, Lima, Peru

The marine ecosystem biodiversity has received little attention up to now, the advances in the knowledge and conservation of the marine ecosystem are scarce, nevertheless its enormous potential, not only as a source of proteins but also as a source of a great variety of active principles for medicine and industrial applications.

The Humboldt Current Ecosystem (HCE) at the Peruvian coast is one of the main marine systems of the world showing greatest complexity, variability and productivity. On the other hand, the northern Peruvian coast shows a high species richness with a large number of Panamian origin species and absence subantarctic origin species; this is due to the transition and water masses mixing of equatorial and tropical origins. Up to 50% of the recorded species are restricted to the transition zone between the HCE and the Pacific Central America Coastal Ecosystem.

There is an increasing interest in the conservation and management of the Mangrove ecosystem, in spite of it covers only near 0.01 percent of the total peruvian coast the inlets are inhabited permanently or temporally by a rich diversity (mainly, fish, molluscs and crustacean) IMARPE, is carrying out the inventory of species in mangrove ecosystem.

The deep sea is an important ecosystem inhabited by different associations of organisms. During the recent years, the species diversity on the deep-sea ecosystem is being studied by IMARPE joint to Japan Deep Sea Trawler Association; most of species are new records. The information about this research will be published in a catalogue.

In Peru, several scientists have published systematic keys, cataloges and checklists about fishes, crustacean and molluscs: 1070 fishes, 1024 mollucs and 480 crustaceans (mainly decapoda and Stomatompoda), 754 phytoplanktonic and macroalgas were registered.

Besides IMARPE is analizing the time series data collected during the pelagic (80s, 90s, 2000s) and demersal (1998-2003) research surveys. IMARPE joint to the World Data Center For Oceanography (WDC) are carring out a international proyect to recover phytoplankton and oceanographic data with their variability along the Peruvian coast, including the ENSO from 1961 to 2000. Another activity in recording the marine biodiversity is Marine Biodiversity Collection, one of the most important in South America. Additionally, IMARPE is developing a database of fisheries including its taxonomy data of each species to promote conservation and sustainable use in an equitable way.

In Peru, IMARPE has been designated to carry out a Marine Thematic Node through National Enviromental Comission (CONAM) according to the Convention on Biological Diversity (CBD) to organize this specific topic to promote and facilitate technical and scientific cooperation, within and between countries with a global mechanism for exchanging and integrating information on biodiversity. First of all, it is necessary to develop a human and technological network as well as to build the technological framework.

Mapping world-wide distributions of marine mammal species using a relative environmental suitability (RES) model

K. Kaschner [1], R. Watson [2], A.W. Trites [3], D. Pauly [1]

[1] *University of British Columbia, Fisheries Centre, Vancover, Canada*
[2] *University of British Columbia, Fisheries Centre, Sea Around Us Projekt, Vancover, Canada*
[3] *University of British Columbia, Fisheries Centre, Marine Mammal Research Unit, Vancover, Canada*

We developed a large-scale habitat suitability modelling approach to map global distributions of 115 species of marine mammals. Predictions were generated by first assigning each species to broad-scale categories of habitat preferences with respect to depth, sea surface temperature and ice edge association based on synopses of published qualitative and quantitative habitat preference information. Using a global grid with 0.5 degree lat/long cell dimensions, we generated an index of the relative environmental suitability (RES) of each cell for a given species by relating quantified habitat preferences to locally averaged environmental conditions in a GIS modelling framework. RES predictions closely matched published maximum range extents for most species, suggesting that our model-based approach for identifying habitat represents a useful, more objective alternative to existing sketched distributional outlines. In addition, raster-based predictions provided more detailed information about heterogeneous patterns of potentially suitable habitat for species throughout their range. We validated RES model outputs for four species (northern fur seal, harbour porpoise, sperm whale and Antarctic minke whale) from a broad taxonomic and geographic range using at-sea sightings from dedicated surveys. Observed relative encounter rates and species-specific predicted environmental suitability were significantly and positively correlated for all species. In comparison, observed encounter rates were positively correlated with < 3 % of 1000 simulated random data sets.
Mapping of suitable habitat for marine mammals using this environmental envelope model is helpful for evaluating current assumptions and knowledge about species' occurrences (especially for data-poor species) and allows the investigation of large-scale biodiversity patterns of these high level predators. Moreover, RES modelling may help to focus research efforts on smaller geographic scales and usefully supplement other, statistical, habitat suitability models.

Dealing with the challenges in presenting taxonomic information online: an introduction to PLANKTON*NET@AWI

A. Kraberg [1], F. Buchholz [1], K. Wiltshire [1]

[1] *Alfred-Wegener Institute for Polar and Marine Research, Biologische Anstalt Helgoland, Helgoland, Germany*

Phytoplankton taxonomy requires comparison of large volumes of information including images of taxa from different geographical areas. The internet should be ideally suited for this task. However, despite its advantages compared with traditional dissemination methods and the huge array of different online taxonomic resources, it lacks the evaluation and validation mechanisms of traditional resources and 'ground rules' for the treatment of taxonomic data have not yet been established. This poster will introduce the online database PLANTKON*NET@AWI, one of several online taxonomic projects at the Alfred Wegener Institute. PLANTKON*NET@AWI contains more than a thousand plankton images from the North Sea and different collections from all over the world. The database can be searched alphabetically or via collections. Each record can be viewed as a standardised data sheet with images and taxonomic descriptions. Comment functions are also provided but their administration has yet to be discussed. Images from different collections can be compared, facilitating the detection of taxonomic inconsistencies and geographic variations in morphology. PLANKTON*NET is a collaborative project with partners at Roscoff and in Woods Hole, but the individual sites are not yet networked. We are currently exploring mechanisms for future database formats and ways of networking existing resources to maximise the benefits for taxonomic research. Our favoured approach will be to follow the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) as an application-independent interoperability framework based on XML technology, so that the integrity of local taxonomic initiatives can be maintained, while sharing content but this will also require discussion within the wider scientific community.

Linear referencing as a tool for analyses of organic material deposition along a sandy beach of Gdańsk – Sopot – Gdynia (Polish coast of Baltic Sea)

M. Kędra[1], J. Urbański[2]

[1] *Institute of Oceanology, Polish Academy of Sciences, Sopot, Poland*
[2] *Institute of Oceanography, University of Gdańsk, Gdynia, Poland*

The main aim of the project was the use of linear referencing to answer the question of what kind of natural processes may influence the variability of organic material deposition along the shoreline of a sandy beach of Gdańsk – Sopot – Gdynia. Linear referencing allows to create a database especially for the purpose of collecting and analysing data connected with the seashore. Still, such applications allows to enlarge the datasets and to create databases for any kind of global process that occurs along the various linear features.
Sandy beaches, with its socio-economical value and environmental role as active biological filters, are of great importance. Biota associated with the interstitial system (sandy beach) play a great role in processing organic material and in returning nutrients back to the sea.
Due to increasing euthophication in the Baltic region, dense algal mats are covering the shoreline and algal blooms became a serious problem and creates much concern. It is essential to assess the amount and volume of organic material and to analyse biological processes connected with algae debris along the shore line. It is necessary to organise data in a database and use appropriate tool to analyse datasets.
GIS provides a linear referencing technique that helps in dealing with data associated with any kind of linear feature. Linear referencing provides a tool that enables associating multiple sets of attributes to portions of linear features. It allows to analyse dynamically changing objects with relative position on linear features. The objects are sampled in two dimensions, using x, y coordinates. They may have different attributes, cover each other and consider different scales of processes. Linear referencing reduce the effort of maintaining, organising, analysing and controlling any data describing processes that occurs along existing linear feature. It allows to associate any multiple set of objects with linear feature. Then, it enables querying, editing or analysing attributive data sets without affecting linear feature.
Considering marine environmental types of data sets, the shoreline might be treated as a basic linear feature while natural processes connected with shoreline should become objects connected with the seashore line. In this case organic material occurrence, changes of species diversity connected with algae debris and dependent on organic material supply, might be considered as natural processes.
Data were collected through the whole summer season along an intensively used sandy beach of Gdańsk – Sopot – Gdynia shoreline. Geographic position and volume of organic material washed ashore were measured using GPS and portable GIS system. Process of washing algae debris ashore and offshore was observed. The data concerning waves, short time sea level changes and wind speed and direction were recorded. The data also concern diversity of sandy beach species connected with organic material and algae debris.

What information on biodiversity can be derived from ichthyoplankton surveys in the western Baltic Sea ?

B. Klenz

Federal Research Centre for Fisheries, Institute for Baltic Sea Fishery Rostock, Rostock, Germany

Ichthyoplankton provides information on the species composition in an area of investigation and leads to a better understanding of the entire fish community.
Since 2000 every year in April/May an ichthyoplankton survey in ICES-Subdivision 22 and 24 has been carried out to sample the plankton community in the western Baltic Sea. The stratified sampling has been performed in order to develop a hypothesis for abundance and distribution of the fish larvae in the context of highly variable environmental factors and recruitment. Another objective was to acquire indicators of possible changes in the natural structures of the fish community.
The time series derived from the Bongo-Net samples is too short to show a trend in larval densities. Up to now samples in the western Baltic Sea yielded only low mean densities.
According to the reproductive biology the fish species were divided in three major groups:

- Fish species with a long developmental phase in the plankton community
- Small short lived species with benthic eggs and a reduced planktonic phase
- Guests without local spawning populations.

Species were identified, which were absent in the plankton of the western Baltic Sea in the 1960s and 1970s.

Development of an updating information system on Decapoda Crustacea Museum collections, useful in education and scientific research

A.Koukouras [1], M.S Kitsos [1], I. Kirmitzoglou [1], C. Arvanitidis [2]

[1] *Aristotle University of Thessaloniki, School of Biology, Laboratory of Zoology, Thessaloniki, Greece*
[2] *Hellenic Centre of Marine Research, Environmental Technology and Management Group, Heraklion (Crete), Greece*

In the Zoological Museum of the Aristoteleio University of Thessaloniki large collections of Decapoda Crustacea from all habitats of the Aegean Sea and other Mediterranean areas are deposited. Recently, all these decapod specimens along with the information on the habitat from which they were collected were computerised and a dynamic database was created, also suitable for use through the internet. In this database, a complete bibliographic list of the decapod species known from the European coasts and the Mediterranean Sea is also included, as well as the relative information for their habitat.
In this presentation, we demonstrate the procedures through which a user of this database can: (a) derive original scientific information for a species or a group of species; (b) assess, through the comparison, his own relative data; (c) borrow specimens for comparative study and (d) educate students and young researchers.
In the near future, in this database will be registered all information concerning the systematics, zoogeography, ecology, biology, conservation and cultural status of every existing species and of every new species that will be collected. This database can be connected with other proportionate, or of wider content, databases in order to have a more efficient diffusion of the contained information and thus, to contribute to the international scientific cooperation.
In this presentation two examples of the database use are also given. The first concerns the acquisition of information on the vertical distribution of a species or a group of species, as it results from the registered data of the collections, and finding whether these data give new information in relation to the existing in the literature.
The other example refers to the capability of finding whether the collection of a new specimen of a certain species (e.g., a Lessepsian migrant) from a certain area, constitutes an expansion of the known geographical distribution of this certain species.

Plankton dataset from Lebanese seawater

S. Lakkis

Lebanon University, Faculty of Sciences, Martine Research Center, Plankton Laboratory, Beirut, Lebanon

Long-time plankton series were collected at different periods from 1965 up to date, at several nearshore and offshore stations in Lebanese waters (East Mediterranean). Sampling process including vertical profiles for phytoplankton and zooplankton with standard plankton nets, was accompanied with hydrobiological parameters such as temperature, salinity, dissolved oxygen, nitrate, phosphate, water transparency, chlorophyll-a and PH. These data were taken during monthly or seasonally oceanographic cruises, providing about 1500 phyto and zooplankton samples. Most of these samples were submitted for taxonomic and biodiversity study and quantitative analysis. All oceanographic and biochemical data were integrated in the database of the MEDAR/ MEDATLAS II Project. Metadata were stored in a standard format and submitted to quality control before being integrated in the Dataset of the system.
The Lebanese waters, including the Levantine Basin is situated in a warm temperate region and are characterised with a high oligotrophy and impoverished plankton. The shortage of nutrient input induces a low primary productivity and thus leads to poor zooplankton. The annual hydrothermic cycle of the seawater is characterised with two annual phases: a cold thermohaline phase in winter (December-March) and a warm phase in summer (June-November). A short interseason (April-May) splits the two annual phases. During winter the water is characterised with isothermic conditions of T=16-17ºC in the whole water column and low phytoplankton standing crop and small species diversity. The warm phase is marked by high surface temperatures up to 30-31ºC, a high salinity (39.75 %º) and a sharp forming thermocline in the waterlayer 35-75 m. The plankton is poor in diversity and in biomass. During the interseason, phytoplankton growth may reach small blooms in May, showing a high cell density. Up to date 400 phytoplankton taxa were identified and there were more than 750 zooplankton species.
The big amount of data should be integrated in a database system to facilitate exploitation and exchange of data.

DOME – database on oceanography and marine ecosystems

J. Larsen

International Council Exploration of the Sea, Copenhagen, Denmark

What
DOME is a relational database which aims to store oceanographic, environmental and fisheries data at the ICES Data Centre in one database, with one common interface to the database, thereby supporting a marine ecosystem approach at the data level. The database is currently under development for biota, fish disease, sediment and water bottle data. In the second phase the aim is to include high-resolution oceanographic data such as CTD data and biological community data. Some fisheries data will be considered for a 3rd phase of DOME's development.

Why
User requirements
The primary motivation for developing an integrated multidisciplinary database has been an increasing demand for integrated products for use in implementing an ecosystem approach. A typical request of this kind could be biological data from a specific area in correlation with data on the physical environment, e.g., fish disease data from oxygen depleted areas.
Furthermore, it is desirable to unify access to data from many disciplines through common interfaces.

Data Centre requirements
Another major driving force has been to lower the high maintenance cost of running a large number of very different database and coding systems within the ICES secretariat. At the same time, the ICES Data Centre gains flexibility by harmonising the databases and data handling routines, enabling the staff to more freely share the workload.

How
The overall aim is to standardise data structure, data codes and data presentation.

The data structure includes two common parts; the top level (e.g. time, position, sampling platform, etc.) and the code part which is organised in a single structure. DOME organises the actual measurements and many of their qualifiers in separate structures, one for each of the major data categories. Separate data structures were necessary because the nature of the data differs among marine data. Therefore, relevant information to be stored and their respective structures are different for each data category.

Standardising codes and data handling routines for data from different scientific communities is essential for integration. Where codes can be translated between disciplinary data without problems, this is implemented. However, the design of DOME has not resolved how to convert incompatible codes. The codes in DOME, while stored in the same structure, include sections of codes that are only relevant to a particular category of data. This mainly creates problems where two disciplines measure the same type of data and there is a need to combine these data.

Data presentation will be through common user interfaces. The web interface that is planned for DOME will enable users to download data through a unified web service directly from the ICES web site. The system shall handle data download through password protection, allowing for different levels of data access restriction.

Technology of distribution phytoplankton geometric information

A. Lyakh [1], I. Agarkova-Lyakh [2]

[1] *Ukraine National Academy of Scieneces, Institute of Biology of the Southern Seas, Sevastopol, Ukraine*
[2] *Tavrical National University, Department of Geography , Saki, Ukraine,*

Phytoplankton geometric parameters, its biovolume and surface area, are important to the study of phytoplankton ecology. Microalgal biovolume is commonly calculated to assess the relative abundance of co-occurring algae varying in shape and size, to study cell cycle processes, or to convert phytoplankton biovolume to carbon. Phytoplankton surface area essential for estimating surface/biovolume ratio. This ratio is an important estimation of microalgae metabolism, rates of uptake of light and nutrients and release of waste products and heat.

Calculation of phytoplankton biovolume and surface area is based on the method of the approximation of the cell shape by the set of geometric models. Every phytoplankton species is associated with some set of models, and volume and area of the models are used as estimation of microalgal biovolume and area. Only a few attempts have been made to standardize the procedure of geometric calculation, but standardized sets of primitives have not been created yet.

In the framework of the project “Development of the system for collection, storage, transferring and exchanging of biological information” of the Institute of Biology of the Southern Seas we developed the technology of distribution of the phytoplankton geometric information through the Internet. This information includes the description of the combination of approximating primitives and the list of taxa with codes of combinations corresponded to each taxon.

The description of approximating primitives consists of:
- Images of the combination in different projections;
- Schemes of the measurement of the linear dimensions;
- Formulas of volume and area calculation;
- Additional information.

Sharing of the information of phytoplankton shape approximation has the following advantages:
- We can produce standards for the calculation of geometric parameters of phytoplankton cells;
- We can decrease number of necessary measurements improving measurement scheme;
- We can check and use current sets of primitives for calculating biovolumes of given species;
- We can choose the best approximation set(s) for the given species;
- We can obtain help of the world internet-community in case of difficulties with formulas construction.

Also, in the frame of the mentioned project, we developed a new alternative method of approximation phytoplankton cells shape: 3D geometric modeling of phytoplankton organisms. The main idea is that we create a set of *base 3D-models*, which imitates typical forms of phytoplankton cells, and then deform these base models to obtain shapes of real organisms. Deformation process consists of displacement of some control points that leads to deformation of a base model. So, for representing shape of any phytoplankton cell we need to know initial base model and vectors of the control point displacement.

Using database to study the taxonomic distribution of organisms inhabiting macrophytes of the Black Sea

S.A.Mazlumyan [1], E.A. Kolesnikova [1]

[1] *Ukraine National Academy of Sciences, Institute of Biology of Southern Seas, Sevastopol, Ukraine*

Creation of databases based on taxonomic distribution of benthic organisms is essential because the dimension of relevant faunistic information is great. A benthic sample may contain ten taxa and more, each taxonomic group may number about 1000 individuals. According to different methods, samples are collected in several replications (up to 5 samples) that also adds to the dimension of data. Increasing geographical and temporal investigation scale would lead to increasing dimension of the data matrix. Huge batches of data are available about coastal algal communities and their inhabitants – the contour biotopes of the World Ocean. In the favourable environment fauna occurring in the growth of macrophytes is diverse and the density of macro- and meoibethic populations is high. First qualitative studies of Black Sea meiobenthos date back to the 19th century, quantitative studies to 1955.

The database has been created based on taxonomic data obtained for a geographical area restricted to a bay (southwestern Crimea) for 35 years. The study focused on the communities found in the growth of Cystoseira crinita in Kruglaya bay. Three sampling sites with different environmental conditions were located in the open coastal seawater (I), at the mouth of the bay (II) and amidst the bay (III). At each sampling site three samples were collected from 0.7 m depth each month from April 1990 to March 1992. In each sample the fouling organisms were counted, identified and their numbers determined per 1 kg algal wet weight. Altogether, 190 samples were collected and handled, and ten taxa, Turbellaria, Nematoda, Polychaeta, Acarina, Harpacticoida, Ostracoda, Amhipoda, Bivalvia and Gastropoda examined.

The database on the taxonomic distribution of meiobenthos proved to be an efficient tool in addressing same aspects of complex study of biodiversity, for example, short-term (interannual), seasonal and long-term variations. K-dominance analysis was made; its results point out that the examined sites differ in the environmental conditions. All the variations were studied as specifically related to each of the sites.

The data base was created using seasonal index determined for the main taxonomic groups of the benthos inhabiting the algal growth; comparative analysis were made to understand variation of the index during 23 months of the investigation and the variation of the index found between identical taxa occurring in different localities of the bay; the limits in which the index seasonally varied were also analysed.

Seasonal peaks were in accord with seasonal and ecological changes which developed in biotope. This is represented by population density estimates obtained for the three sampling sites: the greatest estimates are found amidst the bay, where the tidal effect is the least. Acarina, Harpacticoida and Amhipoda dominated at either of the three sites. In the algal growth safe from anthropogenic impact the estimates of population density of macro- and meiobenthos were higher than in the open sea. Taxa distribute unevenly: Nematoda prefer the safe central part of the bay, Acarina – the mouth of the bay, Harpacticoida are abundant everywhere. In the samples taken from the same locality Amhipoda either may be abundant or completely absent. Turbellaria, Polychaeta, Izopoda were found all the year round, Gastropoda and young Bivalvia – only during spring and summer.

Compared between the years 1990 – 1992 the abundance estimates for Harpacticoida and Acarina decreased in the open coastal seawater area, for Nematoda in the coastal seawater and at the entry of the bay, and for Amphipoda at all three sites. The decreased was owing to unfavourable environmental background in Kruglaya bay during 1990 - 1991. Analysis of interannual (1991 - 1992) average abundance estimates for the benthos has shown that the numbers of Gastropoda was low and Bivalvia vanished from the thallus of macrophytes.

Comparison between the recent data ad those obtained in 1955 and in 1970s confirms the trends of the structural changes of phytophilous benthos. During the 1990s all the main taxa of phytophilous benthos decreased in the numbers; among the taxonomic groups Harpacticoida and Acarina became dominant, the percentage of Nematoda drastically dropped.

Operational integration of biodiversity and physico-chemical data: experience at the BMDC

A. Meerhaeghe [1], K.D. Cauwer [1], M. Devolder [1], S. Jans 1, S. Scory [1]

[1] *Schelde - estuarium, Brussel, Belgium*

Over the past few years MUMM, the Belgian federal administration responsible for the marine environment, has built up a centre of human and technical expertise to ensure a smooth and scientifically sound data flow between data producers and end-users. The Belgian Marine Data Centre (BMDC) serves as national repository and processing centre for marine and environmental data collected in the frame of research or monitoring programmes. Many of these data are linked to scientific projects of the “Sustainable management of the North Sea Programme” - funded by Belgian Science Policy - but datasets originated from other national and international programmes have been included too. The covered data include most domains of oceanology as there are: physico-chemical parameters, biodiversity, hydrodynamics, geography and human interest. Most of the datasets relate to the Belgian Continental Shelf, the Scheldt estuary and its surrounding areas.

Since 1984 the real-time data acquisition system 'ODAS' (Oceanographic Data Acquisition System) on board on the RV Belgica (managed by the MUMM) collects up to 30 parameters. These quality-controlled physical and chemical parameters are kept in a relational database management system together with CTD profiles and current and wave measurements taken during long-term anchorage in the seventies and eighties. On the other hand, the database on the quality of the marine environment, IDOD (Integrated Dynamical Oceanographic Data management), mainly contains the concentration values of numerous substances in air, water, sediment and biota as well as biodiversity information. These values are the result of measurements and observations in situ and analyses performed in laboratories. These values and their meta-data, documented and verified, constitute a coherent and unique source of information for scientists, policy makers and other marine stakeholders.

To render these data to the public, different tools have been made available and are constantly being developed. Through our website (http://www.mumm.ac.be/datacentre) most of the data stored in the IDOD-database can be queried and downloaded for further use and analysis. A spatial tool is provided to map different data or to use in different GIS-applications. Interpolation and statistical analysis of the selected data are also available online. The strongest feature of this system is the common underlying structure for different kinds and sources of data. This opens the possibility to compare biodiversity data, physico-chemical data, geographical data and historical data.

Our presentation will concentrate on the structure of the database and verification methods of the marine environmental data. Some specific case studies will demonstrate the benefits gained by this unique structure and will be described in more detail.

JCOMM ETDMP Pilot Project – the prototype of the "End to End" marine data management technology - basic solutions and development status

N. Mikhalov [1], E. Vyazilov [1], S. Belov [1], S. Sukhonosov [1]

[1] *Hydrometeorological Information, Obninsk, Russia*

This year the Pilot Project "The Prototype of the "End to End" Marine Data Management Technology" (hereinafter referred to as E2EDM technology) has been started. This Project is one in the set of pilot projects to be implemented by the Expert Team for Data Management Practices of the Joint Commission for Oceanography and Marine Meteorology (JCOMM).

The "end to end" data management is considered as the coordinated and inter-connected combination of the following components:
(i) the ocean and marine meteorological data management systems (local DM systems) which are available and being developed under various ocean study programmes and services;.
(ii) the integration technology as an "umbrella" that comprises local DM systems and provides communication and "transparent" interaction between metadata, data and products resulted from these DM systems and also an end-user access to any data and information generated by DM systems.

The goal of ETDMP Pilot Project is to develop solutions (distributed data model, functional, technological and other components) and test them through the integration of selected local data systems. The E2EDM technology prototype should undertake real-time data fusions from distributed oceanographic and marine meteorological data sources into sample products of interest to JCOMM users and demonstrate "end to end" data management opportunities.

The E2EDM technology prototype will handle the following real-time and historical observation data and climate and analysis/forecast products:
(i) In-situ data, including marine meteorological data – air temperature, sea surface temperature, pressure, wave (height and direction), wind (speed, direction) and ocean data – temperature, salinity, oxygen, and some nutrients;
(ii) satellite data – ocean colour imagery data.

On the basis of the analysis of the currently available system and technologies for the development of the prototype it has been decided to use the following: DiGIR Portal and Provider software and Protocol, OPeNDAP data exchange protocil and Java-utilities, SeaSearch CDI, WMO Core metadata, ESIIMO software, data and metadata models and sytructures, and NERS DataGrid metadata and data models.

The conceptual framework of the E2EDM technology is the developed scheme of distributed information resources comprising coordinated metadata and data subschemes (models).

The main components of the E2EDM prototype technology are the E2EDM Integration Server и E2EDM Data Providers, which provide technology services:
access to distributed and heterogeneous data, using some of the local data system so that local data become data sources;
management of metadata, system dictionaries and tables of codes and other;
decomposition of a user request for individual Data Sources, connection with Data Sources, aggregation of data from different Data Sources and provision of them to other services;
end user interface and other.

By now the basic agreements and solutions have been determined and the draft of the technical specifications of the technology component information and software based on the distributed information resources scheme has been developed. The development of the technology programme applications-services and testing of specifications in practice are under way.

Integration of biological information with other information classes - the experiences of the IOC/IODE MarineXML initiative

K. Millard [1], F. Hernandez [2], E. Vanden Berghe [2]

[1] *HR Wallingford Ltd, United Kingdom*
[2] *Flanders Marine Institute, Oostende, Belgium*

It is uncontroversial that improved management of the marine environment requires integration of biological data with data from other sources. XML technology can be regarded as providing a consistent 'data grammar' that marine biologists, chemists and physical oceanographers - in addition to instrument and software manufacturers - can use to bring different genres of marine data together for meaningful analysis.

The IOC of UNESCO has looked to harmonise the application of XML to the marine community through its Marine XML initiative. One project under this umbrella is part-funded by the European Commission. This paper reports on the work undertaken by the EU marinexml project in developing xml-schema based on the geographic mark-up language (GML) developed by the OpenGIS Consortium (OGC) . This GML application schema can be used for data interchange within and across the marine community.

Recent years have seen much activity in the convergence of protocols for data exchange, in particular the harmonisation and adoption of OGC standards within the ISO TC211 (ISO19000) series of standards. The adoption of GML as ISO 19136 means that the marine community no longer has to make a choice between the ISO or OGC path. Given this convergence, the new challenge is to create the tools to support the interoperability of data models across overlapping domains.

MarineXML has concluded on concepts and implementation strategies for the implementation of GML application schemas in overlapping domains. In particular, the use of a Feature Type Catalog as a Web Services enabled registry, able to support emerging libraries of component schema building blocks as well as higher-order semantics of relationships between Feature Types. Three ‘test-bed’ applications have been defined; one of these deals with biogeography – combining species distribution records with information on the physical state of the environment.

Any activity of this kind must strike a balance between maintaining the community view on the data and proving a view on the data that enables it to be understood by other communities; including the technologies supported by these communities. This paper presents the practical results obtained from adopting this approach and the 'hands on' activities require by marine biologists in using and adopting this approach. In addition information is given on how the standards can evolve in the future

Mesozooplankton from Ría de Vigo (NW Spain) and its adjacent shelf between 1995 and 2003

A. Miranda [1], C. Eirín [1], G. Fernández [1]

[1] *Ministero de Ciencia y Techologia, Instituto espanol de Oceanografia, Centro Oceanografico de Vigo, Galiza, Spain*

The mesozooplankton (200 µm to 2 mm) community from monthly samples collected at two stations in Ría de Vigo (39 m) and its adjacent shelf (97 m) between 1995-2003 is described. The mesozooplankton was dominated by copepods. Inside the Ría, meroplankton (larvae of gasteropods, lamelibranchiae, echinodermate and cirripids nauplii) and other holoplanktonic organisms, such as apendicularia, cladocera, siphonophora and medusae, were also well represented. The copepod group inside the Ría was dominated by *Acartia clausi*, juveniles of *Calanus helgolandicus* , *Paracalanus parvus, Pseudocalanus elongatus, Temora longicornis, Oithona plumifera, Oncaea media, Euterpina acutifrons* and copepodites of *Centropages* spp. In the offshore station, the dominant copepods species were the same species as these from the inshore together with other oceanic species, such as *Paraeuchaeta hebes, Calanoides carinatus* and *Aetideus armatus*. Biomass values within Ría de Vigo ranged between 2 and 151 mg dry weight /m3 and between 2 and 314 mg dry weight /m3 in the offshore station. Abundance ranged between 149 ind/m3 in January 1995 and 41,814 ind/m3 in August of 2003 in the inner station, and between 22 ind/m3 in January 1996 and 93,410 ind/m3 in May 1999 in the outer station. The most outstanding events were the high levels of the Mediterranean copepod *Temora stylifera*, which usually rarely occurs in Galician waters, recorded in the area in 1997 and 1998, and the increased levels of cirripid larvae observed inside the Ría.

The IOC taxonomic reference list of toxic Plankton algae

Ø. Moestrup [1], H. Enevoldsen [2]

[1] *Botanical Institute, University of Copenhagen, Department of Phycology, Copenhagen, Denmark,*

Problems pertaining to toxic microalgae are being worked on and discussed by many categories of researchers: ecologists, taxonomists, toxicologists, by persons involved in monitoring the environments for harmful effect of toxic algae, by politicians, journalists, etc. It is a source of confusion that the scientific ('Latin') names of the algae sometimes change when the organisms become available for more detailed studies due to improved methodology or after establishment of the organisms in pure cultures, and such changes may give rise to misunderstandings. To ameliorate the situation the IOC Intergovernmental Panel on Harmful Algal Blooms decided to establish a Task Team on Algal Taxonomy with the aim of providing an agreed reference list of harmful algal species, including correct citation of the author(s), date of valid publication, and a list of synonyms. The team is comprised of Ø. Moestrup (Chair), G.A. Codd, M. Elbrächter, M.A. Faust, S. Fraga, Y. Fukuyo, G. Cronberg, Y. Halim, F.J.R. Taylor, and A. Zingone. The Team has established, updates and further develops the 'IOC Taxonomic Reference List of Toxic Plankton Algae' at http://www.bi.ku.dk/ioc/default.asp. The IOC Reference List contains the correct name according to the International Botanical Code of Nomenclature, followed by author(s) and year it was given its present name, the basionym (the first name applied to the species, including year), reference to the article in which the species was first described, reference to the article in which the species was given its present name, type locality, synonyms including year, main harmful effects, up to (usually) 3 references with information on toxicity, toxins or toxic effects, notes (if applicable). The list of eukaryotic algae was completed in early 2002. The cyanobacteria (blue-green algae) will be added when the list of freshwater algae has been completed (mainly blue-green algae). The on-line version of the IOC Taxonomic Reference list of Toxic Plankton Algae is provided in cooperation with and hosted by the Institute of Biology, University of Copenhagen. Initiative has been taken to collaborate with the International Centres of Marine Microbes, OBIS and 'Planktonnet' and thereby expand the information linked to list and for the list to be a contribution to the mentioned initiatives.

Marine biological data management at the British Oceanographic Data Centre

G. Moncoiffe

British Oceanographic Data Centre, Bidston Observatory, Prenton, United Kingdom

As the UK's National Oceanographic Data Centre, the BODC began banking biological data in the late 1980s. This was in response to the challenges posed by increasingly sophisticated multi-disciplinary programmes such as the UK-NERC North Sea Project and the UK-JGOFS Biogeochemical Ocean Flux Study. The activity provided a unified approach to data management, delivering an integrated quality controlled data set, which also included the physics and chemistry of the water column.

Since then, BODC's biological data collections have grown steadily, integrating a large number of highly diverse data from UK- and European-funded multi-disciplinary fieldwork programmes in the Atlantic Ocean, Arabian Sea, western and north western European margins and the continental shelf seas around the UK. BODC's banking of oceanographic data is supported by two main database systems and a common parameter coding system designed to cover an extensive range of parameters. The structure of the databases and the fully multi-disciplinary coverage of the parameter coding system has enabled the development of a multi-disciplinary oceanographic data bank capable of integrating the ever growing diversity and complexity of oceanographic measurements. Current biological data holdings include spot measurements from discrete water or sediment sampling and net tows, and also continuous profiles and time-series. Parameters include taxon-based and non taxon-based measurements of abundance, biomass, production rate, loss rate, biochemical composition, molecular information and biogeochemical cycling for planktonic and benthic organisms ranging from viruses and bacteria to micro- and meso-zooplankton.

Spatial and temporal coverage is expected to increase rapidly within the next few years as stricter enforcement of data management policies by the Natural Environment Research Council (NERC), the UK main funding body for environmental research, will ensure that copies of most data collected at sea on NERC's research vessels are deposited with the BODC and made available via its databases. Such an influx of data will represent a challenging task for BODC and new strategies are currently being developed to optimise its data handling and data delivering facilities. New channels for a greater collaboration between oceanographic researchers and BODC data scientists have been put in place while on-going development work on the BODC's parameter dictionary, including standardisation of its taxonomic parameter codes with ITIS, and the development of searchable web interfaces to its data holdings will all benefit biological oceanographic data management and accessibility in the UK.

MED POL Phase III database

V. Myroshnychenko [1], S.C. Polat Beken [2]

[1] *Institute of Marine Sciences, Middle East Technical University, Erdemli, Turkey*
[2] *UNE Mediterranean Action Plan, Athens, Greece*

Development of the MED POL Phase III Database has been initiated by MED POL in order to increase its data storage and management capabilities according to the needs of MED POL Phase III and to achieve Internet access to the Database. As a first step, the needs and requirements were identified and the conceptual model was developed (2001). The model was thoroughly discussed and approved during the expert consultation meeting (Athens, 2002). Development of the first version of the MEDPOL Database has been carried out in IMS METU, Turkey in 2002-2003. All available data were loaded and the Internet version of the Database has been published at the UNEP MAP Web Site. The work on the database upgrade continued in 2003-2004 and mainly concentrated on quality control issues and fixing errors that were found during the exploitation. The routine loading of monitoring data has also been fulfilled.

The database was developed using Microsoft Access DBMS, which provides an end-to-end solution for all development tasks, starting from the creation of tables, reports, forms and queries and finishing with publishing the database information on the Internet. The database consists of 4 main tables – Stations, Samples, Analyses and Data – and about 30 auxiliary tables such as dictionaries with MED POL codes, formats, limits, etc. Database Data Management and Administration module implements the following functionality:

- Loading of data reported in standard formats
- Data browsing
- Data editing
- Selection of data on different criteria
- Generating different reports
- Visualisation of data (plots and maps)

MS Access does not possess a tool for drawing maps, so a special Map Module was developed using Borland Delphi for the visualisation of station positions on a map of the Mediterranean. A map Module is also available for generating map images in the web version of the database.
MED POL Phase III Database contains:

- data on trend (and state) monitoring of contaminants and loads on matrices such as biota, sediments, effluents, waters, atmosphere;
- data on monitoring of biological effects and compliance control of pollutants,
- certified material analysis and quality assurance data.

Every single parameter value in the database is supplied with a quality flag, which can be assigned during data loading or after expert control.

Eight Mediterranean countries have been submitted data to the MED POL Phase III Database. These are Albania, Croatia, Cyprus, Greece, Israel, Slovenia, Tunisia, and Turkey. At the moment, the database contains: 457 monitoring stations, 308 of which have data; 4209 samples; 26892 monitoring parameters values; 247 records on CRM analysis.

Use database on the Black Sea benthic diatoms for study of its biodiversity

E.L. Nevrova [1], A.N. Petrov [1]

[1] *Ukraine National Academy of Sciences, Institute of Biology of the Southern Seas NASU, Sevastopol, Ukraine*

Microphytobenthos is the major primary link in the trophic relationships of sublittoral ecosystems and study of this taxonomic group is insufficiently developed in biodiversity investigations of the Black Sea. Microphytobenthos is almost entirely consisting of diatom algae *(Bacillariophyta).* In contrast with phytoplankton, the researches of the Black Sea benthic diatoms were performed mostly for the Western and Northwestern sectors, whereas the water areas nearby shores of Crimea and Caucasus are relatively poorly investigated. The information about diatom's flora is almost lacking for the Southern and Southeastern parts of the Black Sea.

The main number of works is devoted to the description of seasonal dynamics and species composition of diatom flora, smaller amount of articles is dedicated to the study of structural patterns of taxocenes and only a few ones to measures of biodiversity based on traditionally used indexes, basically Shannon (H'), Pieloy (J), etc. Meanwhile, it has been shown (Warwick, Clarke, 1998; Warwick, Clarke, 2001), that the application of these indexes is inexpedient for analysis of historical data in a large-scale spatial and temporary scale, when frequency of number replication of samplings and size of samples are unknown or the quantitative data are absent (when the species list is only available).

The estimation of changes in diversity of benthic diatoms is complicated also by ambiguous response of diatom taxocen on the various environmental factors. In some cases species richness or eveness can remain at a previous level or arise under increasing of environment disturbance extent.

In this connection, any cooperation and combining of present information regarding diatom's flora of the Black Sea will be much needed. The combined database can be used for assessing biodiversity on the basis of relatedness of species (taxonomic distinctness index by R.M. Warwick & K.R.Clarke (1995, 1998, 1999, and 2001).

At present time the taxonomic database (created in "Access" package) of benthic diatoms has been worked out. The total updated list of species has been prepared based on available materials. By most recent evaluations the total species richness of the Black Sea benthic diatoms consists of about 750 species and intraspecies taxa.

The highest species richness of diatoms is registered near Crimea that makes about 61 % of total number of the Black Sea benthic diatom species. Regarding other investigated coastal areas this relative index was much lower: 36,6 % (Bulgarian coast); 47,3 % (Romania); 35,6 % (Caucasus); 46,2 % (Northwestern shelf, but without consideration of species from brackish-water estuaries and lagoons). By reviewing of all species dwelling in hypersaline and brackish-water lagoons, total updated list of diatoms from NW coast includes 604 species and intraspecies taxa (i.e. about 80% of total number of species registered for the Black Sea).

Under comparing diatom species composition of Crimea with other Black Sea regions, the highest extent of species similarity was revealed with Northwestern region, where Bray-Curtis (Chekanowsky) similarity index was 71.3% (for presence/absence transformed species data matrix). This index had a little lower value for pair "Crimea - Caucasus" - 67,4 %. The similarity index of diatom flora between Crimean and Bulgarian coasts was 52,7 %. Among all investigated regions, the lowest degree of species composition similarity was marked between NW sector of the Black Sea and Bulgarian coast (46,1 %).

The most updated total list of benthic diatoms of the Crimean coast, including 456 species and untraspecies taxa, 81 genera, 45 families, 24 orders, 6 subclasses and 3 classes of division Bacillariophyta, has been prepared. By results of studies implemented for the last 10-15 years, 48 new and 21 rare species have been found along the Crimean coast, five of them were recognized as newly-found for the whole Black Sea and 4 species were new for science.

NMFS-COPEPOD: An online, investigator-friendly, global plankton database

T.D. O'Brien

Technology, Marine Ecosystems Division, Spring Silver, USA

The Coastal & Oceanic Plankton Ecology, Production & Observation Database, **COPEPOD**, is a collection of globally-distributed coastal and open-ocean zooplankton and phytoplankton abundance, biomass, and composition data. It features easy online access and searching of over one hundred thousand plankton tows from hundreds of cruises and projects, and is available in a common format with supporting documentation and access software.

COPEPOD data content is displayed and packaged in various sizes and compilations. The user can graphically review and select data by individual project or cruise, or download an entire region or ocean. The user can also select from pre-made global compilations of commonly studied plankton data groups such as “copepods”, “diatoms”, or “total biomass” measurements. **COPEPOD** also offers global mean fields of annual and seasonal mean biomass and abundance.

Credit and acknowledgement of the investigator(s) and/or institutions responsible for the original data collection is a special focus of **COPEPOD**. In addition to being stored within the data itself, this information is summarized in each data set summary and linked within the searching system. Users can search database content by responsible institution, project, or investigator.

The **COPEPOD** online component is static-html. This means that no special interface or server software is required to access it. It will run on any standard web browser, online or from a CD-ROM.

COPEPOD is available online at: www.st.nmfs.noaa.gov/plankton .

Macrobenthic communities in relation to long-term oil pollution of the coastal bottom environment (Black Sea)

T.S.Osadchaya[1], S.V. Alyomov [1], J.G.Wilson [2]

[1] *Institute of Biology of the Southern Seas, National Academy of Sciences of Ukraine, Sevastopol, Ukraine*
[2] *The Uiniversity of Dublin, Trinity College, Zoology Department, Dublin, Ireland*

Overall pollution of the Black Sea and its coastal aquatoria is the major danger for biological diversity of the flora and fauna and negatively influences recreational and other resources at the Black Sea coast. One of the most urgent areas of Crimea is the Sevastopol Bay because of a big port with widely developed industrial infrastructures. As this port structure has been conversed, infrastructure has been divided between the Russian and Ukrainian Navies, cargo shipping and passenger transportation has increased and Sevastopol has developed as tourist and recreation area, it is extremely necessary to develop methods of ecosystem evaluation including designing of models for making possible change prognoses and elaboration of methods of rehabilitation.
Bottom sediments, in which most of the polluted substances that enter the coastal aquatoria, accumulate and transform, are of primary importance to the ecological status of the Sevastopol Bay (which is an inland seawater area subjected to continual human impact). River inflow overuse, poor water exchange due to the recently constructed pier, eutrophication provoked by the untreated or insufficiently treated sewage from numerous and diverse sources of pollution, have badly changed the ecosystem of the Bay. The adverse changes in the environment have affected the marine organisms at every taxonomic level as well as hydrological and hydrochemical characteristics of the Bay. Having accumulated in excess, the pollutants of bottom sediments triggers an imbalance between production and destruction leading to higher trophic level of all water body. Pollutants are injurious to biota through qualitative and quantitative disturbances, which they create in community structures. Simultaneously, biogeochemical processes going on in the environment result in decay and transformation of the received pollutants.
One of the aims of our work, conducted in the framework of INTAS Project "Evaluation of coastal pollution status and bioindicators for the Black Sea (BIOBS)" N 01 – 788, was the investigation of integral pollution level of the Sevastopol Bay on the indices of biological quality (BQI) and pollution load (PLI).
The samples of bottom sediments were collected during four surveys (7 stations altogether) in October 2002, January, April and July of 2003 respectively. Biological (species number, abundance, biomass) and chemical (concentrations of oil hydrocarbons) parameters were analysed on each site. Comparison between recent (BQI) data with available data from the 1980 – 1990s has shown some improvements in "biological quality" of the Sevastopol Bay in the last 15 years (namely: BQI = 0.65 in 1982; 1.15 in 1988 and 1.88 in 1997).

N.P. Pakhorukov

Fish biodiversity of raising in the World Ocean

N.P. Pakhorukov

Institute of the Biology of Southern Seas (IBSS), National Ukrainian Academy of Sciences, Sevastopol, Ukraine

The specific diversity of fishes, which live on underwater raisings, was not yet exposed to a strict quantitative estimation. R.Wilson and R.Kaufmann's (1987) reported about 450 species belonging to 90 families. During our research in eight detailed regions, catches and visual observations revealed 411 species, which belong to 222 genera from 92 families. The highest number of species belonged to the genera *Caelorinchus* (14), *Physiculus* (12), *Hymenocephalus* (9), *Etmopterus*, *Bathypterois* and *Ventrifossa* (on 8), *Epigonus* (7), *Chlorophthalmus, Nezumia* and *Centrodraco* (on 6), *Polymixia* and *Decapterus* (on 5). If to add this list with data on two large ridges - Northwest (Imperial) in the northwest part of Pacific Ocean (Fighter, 1986) and Madagascar ridge in the southwest part of the Indian Ocean (Collette, Parin, 1991; Parin et al., 1993), it is possible to add 141 species, 45 genera and 7 families. Thus, the list of "mountain" fishes of these raisings, which in any way should not be considered complete, totals about 550 species from 267 genera including 99 families. It is possible to believe, that the total number of bottom-near-bottom species on underwater raisings in the tropical zone of the World Ocean makes in any way less than 600-650, i.e. approximately 25 % of the species composition of mesobenthal and mesobenthopelagic ichthyofauna, which, according to N.V. Parin (1988), accounts for about 2,600 species.

Representation of separate genera on separate raising strongly varies. Only 3 genera - *Caelorinchus* (Macrouridae), *Beryx* (Berycidae) and *Epigonus* (Epigonidae) are recorded on all seven raisings. Species of genera *Etmopterus* (Squalidae), *Aldrovandia* (Halosauridae), *Synaphobranchus* (Synaphobranchidae), *Hymenocephalus* (Macrouridae) and *Prometichthys* (Gempylidae) are observed on 6 raisings. *Bathypterois* and *Chlorophthalmus* (Chlorophthalmidae), *Malacocephalus* and *Nezumia* both (Macrouridae) and *Benthodesmus* (Trichiuridae) - on 5 raisings. *Gadomus* (Macrouridae), *Dibranchus* (Ogcocephalidae), *Hoplostethus* (Trachichthyidae), *Scombrolabrax* (Scombrolabracidae) and *Ruvettus* (Gempylidae) on 4 raisings. The representatives from 21 genera are found on three raisings, and 61 genera on two raisings. At the same time prevailing the majority of genera (122, i.e. 55 % of list structures) were caught only in one of 7 investigated areas. 5 families were found on all raisings (Squalidae, Chlorophthalmidae, Macrouridae, Berycidae and Epigonidae), 4 (Holosauridae, Synaphobranchidae, Gempylidae and Trichiuridae) on 6 raisings, 2 (Moridae and Scorpaenidae) on 5 raisings, 10 (Scyliorhinidae, Nattastomatidae, Congridae, Alepocephalidae, Platytroctidae, Ophidiidae, Ogcocephalidae, Trachichthyidae, Oreosomatidae and Scombrolabracidae) on 4 raisings, 18 on 3 raisings, 20 on 2 raisings and the others 33 families on one raising.

A variety of fish fauna (number of species and genera) on separate raisings obviously depends on the variety of biotopes. This variety is determined first of all by bathymetric peculiarities of the raising, its extent and area, variety of soils, characteristics of water weights and biological efficiency of waters. Ranking all the raisings that were investigated on the number of genera and species, would result in: mountain Equator (25/35), raising Sierra-Leone (27/29), ridge Vavilov (27/36), mountain Error (28/30), height Rio-Grande (44/66), ridge Kyushu-Palau (109/151), ridges Nazka and Sala-y-Gomez (126/171) (our and literary data). Thus the mountain Equator represents a separate "point" peak, ridges Vavilov and Error, as well as the raisings Sierra-Leone and Rio-Grande concern to number mesoscales, and ridge Kyushu-Palau and complex ridges Nazka and Sala-y-Gomez are characterised by the greatest extent and as against other investigated raisings incorporate rather shallow mountains.

New biotechnologies for sampling the ecological diversity of the oceans: the informatics challenge

C. Palacios [1], B. Olsson [3], A. Boetius [1], P. Lebaron [2]

[1] *Max Planck Institute for Marine Microbiology, Bremen, Germany*
[2] *Observatoire Océanologique Banyuls-Sur-Mer, Laboratoire de Microbiologie, France*
[3] *Marine Biological Laboratory, St. Woods Hole, USA*

Exploring the ecological diversity is taking the study of biodiversity one step further by unravelling the relationships among the organisms and with their habitat. But ocean biodiversity seems so large and our knowledge of it too small. New technologies might help to decrease the gap helping to better understand the ecology of the vast oceans. Molecular methods continue revolutionising biodiversity studies, particularly challenged in the field of microbiology. In the last decade the application of relatively rapid profiling methods like terminal restriction fragment length polymorphisms (T-RFLP) or denaturing gradient gel electrophoresis (DGGE) to characterise the ribosomal DNA (rDNA) products of the polymerase chain reaction (PCR) experiments have been the general choice for microbial ecology studies. But these methods are not exempt of difficulties; they are generally used in conjunction with sequencing to provide information about the phylotypes in a given community. Moreover, although attempts exist to extract relative abundances from band areas or pick heights that result from these fingerprinting studies, fluorescence in-situ hybridisation (FISH) remains the most reliable method to determine active counts of probed organisms in a given sample. FISH, however, is not precisely a high throughput method. A new set of technologies are being developed that might overcome the technical difficulties that microbiologists still face to study the ecology of natural environments. Their key is the use of the intrinsic phylogenetic information contained in genetically hypervariable regions of the 16S ribosomal RNA molecule, which is present in all living forms, to extract phylotype information of the community directly from sequencing, avoiding the difficulties and assumptions underneath band profiles. In other words, these new methods allow summarising all kinds of organisms present in natural samples providing as well with its relative numbers. This information is becoming essential for ecology research. However, ecological diversity imposes a new challenge to the study of biodiversity; it implies the treatment of large amount of data with the consequent need of its computerisation. In this context, we will present our latest advances on the informatics treatment of data from one of these new methodologies, SARST-V6 (Serial Analysis of Ribosomal Sequence Tags of the V6 rDNA hypervariable region). This method renders concatemers of the V6 region increasing by at least 6-fold the yield of each sequence. With the aim of facilitating the analysis of sequences to SARST-V6 users, a website is being constructed for rapid and accurate split and classification of the concatemers. We have successfully applied this method in various natural environments. Our next challenge will be to study the microbial community composition and functional diversity of a largely unexplored habitat, sunken woods in deep marine water environments. These are very interesting habitats to explore biodiversity. Degradation occurs by specific microbes and animal communities adapted to the use of woods or other high carbon content materials as substrates. It has been suggested that, at least for metazoans, these environments could act as stepping stones for chemosynthetic communities that inhabit hydrothemal vents and cold seeps. However, nobody has demonstrated the presence of anaerobic communities in sunken woods.

Towards an information system on aquatic invasive species with early warning functions for European coastal waters

V.E. Panov [1], V.S. Shestakov [2]

[1] *Zoological Institute of the Russian Academy of Sciences, St.Petersburg, Russia*
[2] *Pacific Fisheries Research Center*

Development of the open databases and information systems on aquatic invasive species (AIS) is essential for effective international cooperation in data and expertise sharing, and provides support for risk assessments, management and control efforts. Internet-based information systems on AIS may serve also as essential elements of early warning systems, providing timely access to geo-referenced data on invasive species distribution from monitoring efforts. In the Baltic Sea region, open informational resources on AIS are located in several national and regional databases, including the Baltic Sea Alien Species Database (Klaipeda, Lithuania, *http://www.ku.lt/nemo/mainnemo.htm*) and on-line GIS application "Invasive Species of the Baltic Sea" (St.Petersburg, Russia, *http://www.zin.ru/rbic/*). Despite formal status of last two databases as parts of the HELCOM project on development of open informational resources on invasive species for the Baltic Sea area (at present supported via GEF-funded Baltic Sea Regional Project), there are several organizational and technical constrains for free flow of information, which can be resolved only in case of active involvement in the national and regional databasing of all holders of primary biodiversity and/or invasive species monitoring and survey data.
Currently we are developing a system for collecting and analyzing scientific information on AIS in inland and coastal waters of Europe. This system will combine four main blocks: Web portal, relational database server, analytical tools and electronic scientific journal. Combining advantages of these blocks will allow to enhance efficiency of delivering relevant information to scientific community, authorities and other stakeholders, providing actual early warning service. In order to make the developing system effective, establishment of Pan-European electronic scientific journal on biodiversity and invasive species records for publication of primary data is a priority objective. The journal will communicate with the authors by means of distributed applications and databases. Articles are intended to be created with the application in database table format. After obtaining such an article by the server, author's rights are registered following by opening data with corresponding reference on the Web portal of the journal. Experts of the European Research Network on Aquatic Invasive Species (ERNAIS, *http://www.zin.ru/rbic/projects/ernais/*) would participate in reviewing such publications. Ideally, suggested approach may shorten time between obtaining primary data by the scientist and their electronic publication to few weeks.

Main features of the system include:

- collection of scientific information approved by designated experts;
- author's rights protection;
- storing data on single server of the Web portal (MS SQL Server);
- access to data from the Internet (client-server applications using ASP or .NET technology);
- data analyzing and result performance using analytical tools and GIS applications located on the Web portal (applications of ArcGIS family: ArcView with ArcGIS Spatial Analyst extension, ArcIMS server);
- publishing and using applications simulating and forecasting ecosystem processes;
- technical support of databases and original software installed on the computers of the data providers;
- developing and distribution of applications available for use on computers with different characteristics of the system and software as well as different regime of access to the Internet;
- dissemination of information on records of invasive species among decision makers.

At the first stage, we suggest to use this approach for the Baltic Sea area, the region highly sensitive to AIS introductions, intensively studied by high-level experts in frameworks of national and international initiatives, including ongoing EU Sixth Framework Program and Baltic Sea Regional Project.

The Marine Life Information Network (MarLIN)

J. Parr [1], D. Lear [1]

1 Marine Biological Association, The Laboratory, Plymouth, UK

The Marine Life Information Network for Britain and Ireland (MarLIN) started its work in 1998 with the intention of providing information for environmental management, protection and education. Those objectives have been achieved through the MarLIN Web site (www.marlin.ac.uk). New approaches are being developed to make that information even more useful to decision-makers by ensuring that account is taken of relevant directives, conventions and statutes in presenting that information and that results are displayed in a form that has become 'expected' – on the Web, using GIS and with links to relevant other information.

As the marine node of the United Kingdoms National Biodiversity Network (NBN), the MarLIN programme is actively pulling together the widest range of possible data sources from a wide variety of data holders. This is achieved through visits to data holders and through the offer of assistance with data mobilisation and conversion. With the National Marine Biological Library, MarLIN provides a Marine Life Data Centre for the UK where survey data sets can be deposited and maintained in perpetuity for groups that are unable or unwilling to manage their own data holdings. MarLIN is collaborating with other groups within the UK to develop a centralised internet mapping initiative for marine data in the UK and through it's links with the NBN. MarLIN data is available to the GBIF and OBIS initiatives.

Concise, targeted, quality assured information is of paramount importance for environmental management decisions. The MarLIN Web site includes a wealth of information on key species and habitats, UK BAP species and habitats. On-line decision support tools are included to support decision-making under present imperatives such as the Habitats Directive, SAC management planning, and the Water Framework Directive.

The MarLIN programme has also made progress in integrating information on the likely sensitivity of species and biotopes with marine survey data in marine spatial planning tools, using GIS, to prepare 'sensitivity maps' for use in local and regional environmental management and emergency response.

MarLIN is an Internet initiative of the Marine Biological Association (www.mba.ac.uk)

Marine biodiversity data management and dissemination: a case study from Gulf of Mannar marine biosphere reserve, Southeast Coast of India

E.J.K. Patterson

Suganthi Devadason Marine Research Institute (SDMRI), Tuticorin, Tamil Nadu, India

The Gulf of Mannar (GOM) is one of the four major coral reef areas of India, covering an area of about 10,500 sq. km from Tuticorin to Rameswaram, Southeast coast of India. The area includes a chain of 21 coral islands situated between Lat. 8' 47 - 9' 15' N and Long. 78' 12 ' - 79' 14' E. GOM was declared as Biosphere Reserve in 1989 and it is the first Marine Biosphere Reserve in South and Southeast Asia. The area is predominantly a high-biodiversity coral reef ecosystem housing more than 3600 species of associated flora and fauna, including 117 species of corals, 13 species of mangroves, 147 species of seaweed's, 13 species of seagrass, 17 species of sea cucumbers, 510 species of fin fishes, 306 species of mollusks etc. Due to the biological and economical importance of this precious ecosystem, research and monitoring works on various aspects have been carried out for over five decades. In order to conserve and manage the biological diversity of GOM, UNDP-GEF has recently sanctioned a project and for the first time in India, an autonomous trust called "Gulf of Mannar Biosphere Trust" has been formed for proper conservation and management of this ecosystem. However, until now, an exclusive Marine Biodiversity Data Base and dissemination of information for this ecosystem is in the initial stage with little input. Recently, a "Reef Ecosystem Information System (REIS)" has been initiated within the ongoing Coral Reef Degradation in Indian Ocean (CORDIO) program and through REIS the database management and dissemination of reef ecosystem research would be made for all four-reef ecosystem in India with initial priority to GOM. The REIS would also function in association with the SDMRI's already existing "UGC – Network Resource Centre" especially in the dissemination of information. Within REIS, necessary information on reef ecosystem research from throughout the world will also be collected, stored and would be disseminated to researchers, students, managers and to all concerned departments on request. The database management through REIS would mainly carried out in four topics viz; 1. Corals, 2. Associated Resources, 3. Environment and 4. Conservation and Management. The concept of complete data base, dissemination modalities and networks within REIS will be discussed.

Evaluation of diversity and assessment of technogenic pollution impact upon taxocene structure of benthic diatoms in coastal zone of the Black Sea

A.N. Petrov [1], E.L. Nevrova [1]

[1] *Ukraine National Academy of Sciences, Institute of Biology of the Southern Seas NASU, Sevastopol, Ukraine*

Benthic diatom algae (Bacillariophyta) are leading among all other groups of microphytobenthos by abundance of population and species richness. They are dwelling in all sublittoral biotopes from the surf zone up to a depth of 50-70 m. They have an important role in self-purification processes, in an oxygen balance of coastal water areas and as trophic basis for larvae stages of many species of necto-benthic and demersal fishes. Benthic diatoms are closely associated with a certain biotope and directly subjected by environmental factors. The latter allows us to consider them as one of the appropriate indicators of anthropogenic impact during the complex monitoring of sublittoral ecosystems.

Benthic diatom taxocene in the western and north-western sectors of the Black Sea have been examined the most, whereas the north-eastern shores (Crimea and Caucasus) are relatively poorly investigated. Information on diatom flora is almost completely lacking for the southern and south-eastern parts of the Black Sea. At present, almost all coastal water areas of the Black Sea are impacted by pollutions of various origin and extent. Therefore, evaluation of anthropogenic impact upon the structure of diatom taxocene can be recognised as one of the important parts of ecological monitoring of the Black Sea coastal water areas by using this group of benthos.

Comparative assessment of the impact of toxicants (heavy metals, chlorine-organic compounds and oil hydrocarbons) on the structure and diversity characteristics of benthic diatom taxocene has been carried out by algorithms of multivariate statistics, based on the database of benthic diatoms.

Two near shore water areas of Crimea were comparatively investigated: Laspi bay (1) is located nearby the borders of a marine reserve and is unaffected by any technogenic pollution, while Sevastopol bay (2) is situated within the industrial zone of Sevastopol port where the average level of toxicant's content in silty bottom sediments was higher compared to area (1) (5-13-fold for heavy metals and 22-270-fold for chlorine-organic compounds and oil hydrocarbons, respectively).

Total species richness, number of mass species and average density of diatom assemblages in the healthy water area, appears to be higher than in the polluted bay: 176 and 53 species and 3.02×106 cells/cm2 (1) and 128, 38 and 2.057×106 cells/cm2 (2), respectively.

Results of the comparative evaluation of the Spearman rank correlation coefficient (Pw) have shown that the combination of the following toxicants: Pb+Mn+Cu+DDT have mostly influenced alteration in structural patterns and diversity of diatom taxocene (Pw=0,73-0,75). "Variability increasing" effect of diatom taxocene structure under conditions of high content of toxicants has also been revealed.

In each of the bay statistically significant taxocenotic complexes and subcomplex groupings of diatoms were distinguished based on results of clusterization and nMDS ordination. Inter-complex differences in quantitative characteristics and species structure of taxoceneare mostly pronounced and caused by different response of discriminating species, determining the highest contribution to dissimilarity between complexes, to a high level of toxicants. Structural differences at subcomplex level are less pronounced and can be conditioned by similar reactions of the respective discriminating forms, defining differences between groupings, upon joint influence of leading environmental factors within a certain bay.

Lists of the most significant discriminating species were determined, which can also be considered as an indicator of the diatom taxocene' condition under anthropogenic load. It is proposed to consider T. *tabulata*, A. *proteus* and *N. reversa* as indicators of conditionally unpolluted biotopes in the Black Sea, whereas *T. punctata* var. *punctata*, *D. smithii* var. *smithii*, *N. sigma* var. *sigma*, *F. forcipata, A. crystallina* and *P. quadratarea* can be marked as indicators of biotopes that are subjected to a persistent technogenic impact.

The ICES North Sea Benthos Project: objectives and data management

H. Rees[1], E. Rachor[2], E. Vanden Berghe[3]

[1] *Center of Environment, Fisheries and Aquaculture Sciences, Burnham Laboratory, United Kingdom*
[2] *Alfred Wegner Institut für Polar- und Meeresforschung, Helgoland, Germany*
[3] *Flanders Marine Institute, Oostende, Belgium*

The ICES Study Group on the North Sea Benthos Project 2000 (an offshoot of the ICES Benthos Ecology Working Group) is integrating recent macrobenthic infaunal data (1999-2001) available from various sources, including national monitoring surveys, in North Sea soft bottom sediments. It is expected to cover most of the North Sea. The main goal is an overall comparison with the ICES North Sea Benthos Survey data of 1986, in order to determine whether there have been any significant changes and, if so, what may be the causal influences (e.g., climate change, fishing impacts). The work will contribute valuable information on several other topics such as habitat classification and the distribution of endangered species.

In addition to physico-chemical measurements of sediment samples alongside the benthic fauna, information on water depths, temperature, water quality and salinity will be incorporated in the analysis of species and community distributions. Also, we will use existing ecological and hydrographical models for currents, bottom shear stress and carbon input, along with information on the distribution of habitat types, to explain the observed distribution patterns.

Data received from the project participants were uploaded into a central database. The database itself is an MS SQL server database, with a front-end developed in MS Access. For the taxonomy/species lists, no separate database was developed, but the 'Aphia', existing species register at the Flanders Marine Institute (VLIZ) was used. The structure for the taxonomic information was adapted from the structure of the ITIS database. The taxonomic hierarchy is implemented as an open-ended hierarchy, where every taxonomic name or name part is stored in a single record, together with a pointer to a 'parent' record. Rather than linking the distribution records directly to Aphia, intermediate tables were used, allowing conservation of the original names, and control over the degree of lumping of doubtful taxa in a single entity for purposes of analysis.

Predicting trophic level for all fishes

J. Rius [1], G. Tolentino-Pablico [1], R. Froese 2, D. Pauly [3]

[1] *Worldfish Center, Laguna, Philippines*
[2] *Leibniz – Institut für Meereswissenschaften, Kiel, Germany*
3 University of British Columbia, Fisheries Center, Canada

The place of species in the aquatic food web is usually expressed as trophic level (troph), with primary producers such as algae at the bottom (troph = 1), herbivorous species that feed on plants with troph = 2, predatory species that hunt herbivores with troph = 3, and top predators that hunt other predators with troph > 4. Trend in mean trophic level of catches has been identified as an indicator for sustainable use of ecosystems. Trophic levels are derived from diet studies, which are available for only a subset of relevant species. Here we present a method for estimating trophic level based on trophic information from closest relatives and maximum size of the respective species. This approach provides preliminary trophic level estimates with standard error for practically all fishes.

Biological data management activities at CENDOC related to research cruises in the Chilean inner southern channels

R.L. Rojas

National Oceanographic Data Center oh Chile, Servicio Hidrográfico y Oceanográfico de la Armada (SHOA), Valparaíso, Chile

The region of the Chilean fiords extends from 41º31'S to 55º00S. It can be divided into three big areas which encompass the following geographic features: the Gulf of Reloncaví (41°31´S) to San Rafael Lagoon (46°32´S), the Gulf of Penas to the northern border of Magellan's Strait (52°30'S) and from Magellan's Strait until Cape Horn (56°S).

This whole large area, one of those fewer populated and more unknown of the Chilean territory, presents multiple queries from the oceanographic, geological and marine biodiversity point of view, and for which a growing demand of use of its coastal area exists, especially for the development of the aquaculture and tourism.

For this reason, the National Oceanographic Committee (CONA) located at SHOA, a committee that includes most of the academic and governmental institutions related to marine sciences in Chile, identified this area as a high-priority one in 1994, and managed to get funds from the government to support a a long-term multidisciplinary research study to the fiords. The first CIMAR Fiordo cruise started in 1995 and up to present a total of 8 research cruises, on board Chilean navy research vessel AGOR Vidal Gormaz, have been performed into this area, collecting data and samples for meteorology, geomorphology, physical and chemical oceanography, geochemistry of sediments, fish, crustacean, plankton and benthos. The results of, at least, the first 6 cruises of this research as a whole has produced a considerable amount of scientific papers, presentation to workshops in national and international forum.

Nevertheless this rapid growth in academic production, the data management was phased until 2000 when the Chilean NODC (CENDOC) was tasked to start gathering, quality control and archiving the data as well as all the information obtained from those cruises that were piling up at CONA. In order to obtain prompt results, due to delayed starting time, CENDOC prepared oceanographic products in the form of Data Reports in CD. So Far, four data reports have been produced (http://www.shoa.cl/cendoc/organizacion/datareport_cimar.htm) and an on-line bibliographic index web page was created (*http://www.shoa.cl/cendoc/ib/ibc/ibc1-7.html*). Besides the data center has handled to archive all the physical metadata an data on its databases.

Regarding the biological management of the CIMAR Fiordo data, besides that the data have been included in the CDs, there is still some uncertainty regarding what is the best way to handle such heterogenic parameter. So far CENDOC has collected the biological metadata of the cruises and archived accordingly and plans called for a new data organisation in order to be better suited to build new oceanographic data products, such as GIS charts. Some contacts have been made with the CONA´s Task Team on Marine Biodiversity in order to be advised on how to apply standards in classifying taxonomic groups to CIMAR data. Improvement is expected to arise in the next months and they will be shown in the poster presentation.

Marine biological data management at Japan Oceanographic Data Center (JODC)

S. Sato [1], M. Nagao [1], N. Baba [2], Y. Tomioka [1]

[1] *Japan Oceanographic Data Center, Japan Coast Guard, Hydrographic Department, Tokyo, Japan*
[2] *Universiti Teknologi Malaysia, Jalan Semarak, Coastal and Offshore Engineering Institute, Malaysia*

Japan Oceanographic Data Center (JODC) has been tackling the development of the marine biological data management system under the supervision of the Marine Biological Data Management Advisory Group, since 1984. This group consists of Japanese researchers, mainly marine biologists, from universities, research institutes and private research companies. The first edition of 'JODC Taxonomic Code of Marine Organisms (Plankton)' was published in 1988, in order to manage and process efficiently the marine plankton data obtained by Japanese research institutions. Recently, JODC revised the taxonomic code and published the results as "JODC Taxonomic Code of Marine Organisms (Plankton) 2001 Edition" with the co-operation of 53 researchers on taxonomy so as to meet the recent demand for knowledge.

The 2001 Edition contains up to 8,088 marine organisms (8,242 including subspecies). Each species is identified by code composed of two sets of numbers, Taxonomic Code and Name Code. A Taxonomic Code is a 14-digit integer. A 2-digit integer is assigned to phylum (division), class, order, family, genus, species, and subspecies. A Name Code is a 5-digit integer. It is a one-to-one correspondence with the name of the organism genus or species. The Taxonomic Code may change in case that some species changes the taxonomic group in future, but on the other hand the Name Code will never change.

JODC has collected the marine plankton data observed by Japanese institutes and classified organisms by the JODC Taxonomic Code. At present, more than 3,800 profiles of plankton data are available on J-DOSS (JODC Data Online Service System: *http://www.jodc.go.jp/service.htm*), which is an online data retrieval system operated by JODC.

In addition, JODC has been developing the new JODC taxonomic code for benthos in the adjacent seas of Japan for several years, in order to expand the marine biological data management system for the benthos data. By April 2005, JODC will complete the development of the new code, in which more than 18,000 benthos species are assigned the codes.

Environmental sensitivity mapping

O. Sergeyeva

Ukraine National Academy of Sciences, Institute of Biology of Southern Seas, Sevastopol, Ukraine

Environmental sensitivity maps become one of the very important tools in ecology. They are widely used in the resource assessments, recreational planning, biodiversity conservation, environmental impact assessments, and many others areas. They are important, for example, in the oil spill contingency planning process. However, the environmental sensitivity maps are usually prepared as a hard copy. They are rather complicated, include many special symbols, a lot of intercrossing coloured areas, etc. That is why they are difficult to read and analyse, and they do not allow to see the processes in dynamics.

A special Environmental Sensitivity Mapping software tool that is based on the raster digital maps overlay, has been developed and tested using data for the Black Sea region. This interactive tool uses the original method of the “waited” maps overlay. It operates an unlimited set of maps. The input maps can be taxon based or non-taxon based. One can also jointly use the maps calculated as a result of modelling and the maps based on the field observations. For each map the user can define an aggregate environmental sensitivity value, which represents the integration of different parameters such as sensitivity, vulnerability etc. In contrast to the paper maps, where the values are set up by developers, this tool allows users to set values themselves according to their needs and interests. After the input maps are chosen and environmental sensitivity values are defined, the several operations can be performed. They are raster Boolean overlay operations and arithmetic overlay operations. From this a great variety of generated maps can be created, combining the various sensitivity parameters. There is a possibility to add your own colour scheme for values for the resulted map. So, user can “play” with data changing weight of different species, the represented season, etc.

There is an urgent need to develop such tools, which will enable proper environmental planning, as well as emergency response to mitigate the impacts of future environmental damage. By mapping the sensitivity of ecosystems it should be easier to identify potentially damaging emissions, analyse the consequences of pollution, display new relationships e.g. between biodiversity and anthropogenic factors, fauna and flora distributions. The most sensitive areas are the first to be affected and this software is a tool for identifying such areas as well as important natural and human-use resources. It is a powerful tool to analyse and present the dependences of ecological factors and to visualise the modelling results.

Using data systems to evaluate seamount biogeography

K.I. Stocks [1], D. Tittensor [2], R.A. Myers [2]

[1] *San Diego Supercomputer Center, University of California , San Diego, USA*
[2] *Department of Biology, Dalhousie University, Halifax, Canada*

This project employed online data resources in an effort to model and predict patterns of seamount community structure. The purpose was two-fold: to evaluate the ability of these data resources to test ecological hypotheses, and to guide future research priorities for an international seamount field program currently under development.

Seamounts are undersea peaks in the ocean's floor. Recently, they have become the focus of scientific interest for two reasons. First, the faunas found on many seamounts are unique and unusual: in several recent studies, 15-40% of species found were endemic, though rates varied greatly by seamount. Secondly, seamounts are experiencing increasing pressure from deep commercial fishing, and there is mounting concern over how to best mitigate and manage the potential impacts of trawling. Recognising the ecological and economic importance of seamounts, the Census of Marine Life (www.coml.org) is considering launching an international field program on seamounts to catalyse and coordinate future research. A primary concern in the planning of this program is how to allocate future sampling. Given that tens of thousands of seamounts exist in the ocean, is it possible to identify a subset that can reasonably be sampled and will significantly increase our knowledge of seamount biogeography?

SeamountsOnline (seamounts.sdsc.edu) is a online database collecting data on the species composition of sampled seamounts. It was used, together with accessory information from FishBase (www.fishbase.org), EarthRef (earthref.org), ETOPO global bathymetry, and other sources, to evaluate whether proxy variables exist that can predict seamount endemism. For example, we examined whether factors highlighted by Island Biogeography Theory (age, distance from like habitat, size) correlate with rates of endemism. The goal was to develop a subset of important variables and then evaluate the current body of seamount sampling to determine where future sampling would most increase knowledge of under-sampled types of seamounts.

In the results, we present our findings on patterns in seamount endemism and gaps in current sampling. In addition, we evaluate the ability of available data systems to address this kind of ecological analysis: i.e. what kinds of questions can existing systems be applied to, given their characteristics (i.e. data incompleteness, heterogeneity, etc.), and what kinds of questions cannot be appropriately addressed?

A dynamic web portal for integrated coastal monitoring

E. Tamiro [1], C. Gitto [1], A. Bergamasco [1]

[1] *Consiglio Nazionale delle Ricerche, Istituto per l'Ambiente Marino Costiero,Messina , Italy*

The main objective of marine environmental data management is the setting of technical and scientific bases for querying and viewing such data for monitoring and investigation purposes. The key element to promote integration among scientist and between scientists and environmental managers is a shared data bank to make the stored data sets easily available via an Internet browser.
In the framework of the SAM Project, funded by the Italian Ministry of Research, two elements have been put together to support Integrated Coastal Monitoring: an operational infrastructure at sea and a software infrastructure in the data management centre.

The former is composed of a network of automatic coastal platforms connected via GSM technology for real time monitoring and a boat for in situ data collection and analysis, the latter, called SAM-BA, is composed of a RDBMS for data storing and querying, a dynamic web portal for worldwide data distribution with graphing capabilities and a computer cluster for data analysis.

The SAM database manages: i) automatic meteoceanographic and water quality data; ii) data from lab analyses on water, sediment and biota; iii) vertical CTD profiles and surface data collected by underway systems; v) marine biodiversity data, in the form of lists of species and abundances.

Automatic data flow along the following path: after being collected by platforms they are sent via SMS in form of encoded e-mails that are automatically forwarded to the data centre. These e-mails are parsed and data are decoded for storing in the database. For loading non-automatic data, different origin-specific clients are currently under development. In the final stage the data are published on the web by an Internet portal, which queries the database with user defined criteria.

The database is designed focusing on environmental data as a key-part of the whole analysis process; the database structure can be seen as a main table storing environmental data, with a set of linked tables, which specialise the first one by adding information about time and space, measurement units, projects, campaigns, etc. A quick data reliability test compliant with the MAST Code on Data Quality is performed during the mail parsing to allow the realtime display of checked data.

A cluster of Linux workstation running LAM/MPI parallelized software is then used to speed up a second and more precise quality check on the overall data set taking into account averages and standard deviations of each parameter. The final quality rate is then available as an attribute for web queries.

A web portal has been developed to explore data using all of the specialising information such as the time-window or the depth-window, the parameter, the producer (project/researcher) etc. according to the interests of the user. Once selected, the suitable data set can be exported as a formatted text file for spreadsheet importing.

It is also possible to graph the queried data using the graph tools provided by the web portal. Graphs can be built using a multiple parameters search and can be plotted in form of time series or vertical profiles with a specific scale unit for every parameters. Graphs can be also customised by specifying line colours and drawing styles. The web portal is built on the HTTPS protocol for secure data share.

Biodiversity and functioning of the sound scattering layers (SSL) myctophidae taxocenosis in the Tropical Ocean

S.A. Tsarin

Institute of the Biology of Southern Seas (IBSS), National Ukrainian Academy of Sciences, Sevastopol, Ukraine

Lanternfishes or Myctophidae are one of the most numerous fishes in oceanic pelagic zones. They are important components of SSL fauna in the upper pelagic zone at night. The Myctophidae population of the night SSL in the epipelagic zone is a special ecological group. They are convenient for examining as taxocenosis.

Differences of biodiversity Myctophidae taxocenosis from the night SSL in epipelagic zone between the Tropical Atlantic and the Tropical Indian Ocean are observed. A greater number of species has been marked in the species composition of Myctophidae taxocenosis in the Tropical Indian Ocean, in comparison with the tropical zone of the Atlantic. (Up to 30 species simultaneously against 10-15). Thus in the Atlantic Ocean mainly taxocenosis of powerful jet currents has been observed. In the Indian Ocean only taxocenosis has been found that is connected with a certain water mass, usually limited to the closed circulations. Probably, this phenomenon is connected with often (semi-annual) changes of current directions in the Indian Ocean. It depends on monsoons. Taxocenosis, which are directly connected with jet currents do not have time to develop.

Presence of myctophids in the upper water layers in structure of SSL fauna during nighttime is connected with peculiarities of their nutrition. In principle, myctophids eat at night. The main components of myctophids food are copepods and euphausiids. Their ratio in diet of lanternfishes depends on the size of fish. Their vertical distribution is also connected with peculiarities of their nutrition in SSL. The majority of myctophids eats whatever food is accessible to them. However some specific selectivity of nutrition exists. For example, Hygophum proximum, having the widest distribution area in the Indian Ocean, also has the widest spectrum of nutrition. And the species Ceratoscopelus warmingi, occuring in the three oceans, proves to be the greater predator in comparison with other species of the same sizes. In addition, this species can switch to eating diatomaceous algae in oligotrophic areas.

Although the majority of myctophids does not belong to gregarious fishes, they form certain aggregations. These aggregations will consist of individuals of a single species of nearby sizes. Thus aggregations of different species are not isolated from each other and can be imposed.

The big influence to myctophids biodiversity in the night SSL also renders lunar light. The species composition of Myctophidae taxocenosis considerably varies depending on the intensity of lunar light. At the maximal lunar light into structure of SSL fauna in epipelagic zone enter superficial myctophids, but species of genus Lampanyctus do not enter, at the minimal and moderate lunar light on the contrary. It is connected with that, these groups of species deepen and thus avoid bright light. It testifies to the big role of sight in the life of myctophids.

Reproduction of myctophids occurs in the upper layers of water. At this process, the concentration of myctophids in local sites increases and is usually marked of sexually mature fishes. The mechanism of this process is not known yet. Mature individuals might actively find each other. And it is possible at a casual congestion, that a significant amount of individuals somehow accelerate each other to enter puberty. Anyway, the process of egg hydration runs very quickly (within some hours).

A web-based fisheries oceanographic information system for the Gulf of Maine

V.M. Tsontos [1], D.A. Kiefer [1], F.J. O'Brien [2]

[1] *Department of Biological Sciences, University of Southern California, Los Angeles, USA*
[2] *System Science Applications*

A prerequisite for better understanding the factors influencing the dynamics of marine populations in space and time is the integration, within a unified spatial modelling framework, of quantitative survey data on species distributions, environmental information, and analytical models. Developed as the informatics component of the Census of Marine Life regional project for the Gulf of Maine (GoM), the GoM electronic fisheries atlas combines these features with standardised protocols for data access within a Web-based oceanographic GIS to support the information needs of scientists and decision makers. Building upon the successes of the GMBIS (GoM Biogeographic Information System) pilot project, the fisheries atlas combines historical groundfish survey databases from both Canadian and US sources (DFO and NMFS) for the GoM and Scotian Shelf with coincident physical data and an extensive series of satellite imagery of diverse types. Integral to the system are also a series of bio-optical algorithms that facilitate the estimation of optical properties at depth using remotely sensed surface observations as inputs, thus providing a full 3D-characterization of primary productivity fields that are relevant to studies of fisheries oceanography. These bio-optical algorithms are implemented via a comprehensive COM modelling interface recently developed within the GIS software (EASy) that supports the coupling of a full range of computational model types, whether they be dynamic process models, statistical models, or data transformation algorithms.

Genetic and morphologic divergence and phylogenetic relationships of Mediterranean Mullidae species (Perciformes)

C. Turan

Fisheries Genetics Laboratory, Faculty of Fisheries and Aquaculture, Mustafa Kemal University, Hatay, Turkey

Allozymic, morphometric and meristic divergence and phylogenetic relationships of four Mullidae species Mullus barbatus, Mullus surmuletus, Upeneus moluccensis, Upeneus pori were investigated. Eight enzymatic systems corresponding to 12 putative loci were assayed, among which, three (AAT-2*, IDHP*, PGI-2*) were found to be polymorphic in at least one species, while the remaining nine were monomorphic in all species. Several loci showed different electrophoretic patterns among species and thus can be used in species taxonomy as diagnostic markers. Especially, PGM* and SOD* loci proved to be species-specific.
Pairwise comparisons of genetic distance were found to be 0.533 between M. barbatus and M. surmuletus species of the genus Mullus, and relatively higher genetic differentiation (D= 0.864) was observed between U. moluccensis and U. pori. For intergeneric comparisons, the highest genetic distance (D=1.401) was detected between M. surmuletus and U. pori, and the lowest (D=1.106) was found between M. surmuletus and U. moluccensis. Remarkably U. pori was genetically the most distinct species from the genus Mullus. UPGMA clustering analysis of genetic data separated the two genera: M. barbatus and M. surmuletus clustered together in one group, while U. moluccensis and U. pori clustered more divergently in the second group.
UPGMA clustering analysis of morphologic data revealed similar pattern of divergence and phylogenetic relationships with the genetic data. However the correlation between the genetic distance and Squared Euclidean distance of morphologic data showed non significant correlation ($r=0.71$, $P>0.05$). The morphological divergence was higher within the genera Mullus than within the genera Upeneus. The highest morphologic divergence was observed between M. surmuletus and U. pori, and the lowest was detected between U. moluccensis and U. pori.

EurOBIS: the European node of the ocean biogeographic information system

B. Vanhoorne [1], S. Claus [1], D. Cuvelier [1], E. Vanden Berghe [1], J. Mees [1]

[1] *Flanders Marine Institute (VLIZ), OOSTENDE, België*

Within the MARBEF (MARine Biodiversity and Ecosystem Functioning) network, the European regional node of the Ocean Biogeographic Information System (EurOBIS) will be developed. This distributed system to present biogeographic information will integrate individual datasets on marine organisms into one large consolidated database, and can therefore provide a better understanding of long-term, large-scale patterns in European marine waters. The system makes use of the Distributed Generic Information Retrieval (DiGIR) protocol and is fully platform independent. It can be installed on any computer that has PHP and a web server running. The DiGIR records are available in a standardised XML format and are transferred via HTTP. A local copy of all remotely-held data is stored at the Flanders Marine Institute (VLIZ) for performance reasons. Regular queries guarantee the data integrity of the system. An interactive layered Scalable Vector Graphics (SVG) map is available for comparing and visualising the biogeographic data.

The ultimate goal of EurOBIS is to provide the end-user with a fully searchable biogeographic database, focused on three main parameters of a distribution record: taxonomy, temporal and geographical cover. The European Register for Marine Species (ERMS) will function as the taxonomic backbone; each species list from every incoming dataset will be matched with this register.

The inventory and integration of marine biological datasets is one of the main tasks of the data management team of the MARBEF project. This data archaeology, facilitated by the creation of the Network, tries to recover as much biogeographic information as possible. Data inventories and datasets of former European projects, such as BIOMARE will be recovered and will contribute to EurOBIS.

The most effective fields of use of XML language

E. Vyazilov [1], S. Sukhonosov [1]

[1] *Hydrometeorological Information, Obninsk, Russia*

In recent years, XML language is becoming more often applied in various domains. The fields of its application are extending very quickly. Unfortunately, XML use is not always effective in the modern development of computer engineering and in telecommunication means.

In this work, the generalisation of progressive experiences of using XML language, including in All-Russia Research Institute of Hydrometeorological Information – World Data Centre (RIHMI-WDC) in 1999-2004, is offered.

The lacks and advantages of using XML language in the following areas are determined as:

- Creation of demonstration variants of systems with dynamic representation of the data on the screen browser;
- Development of means of dynamic access to the interconnected objects metadata (item of information on files and databases, organisations, experts, formats of a storage given, software, used at it, measuring systems, parameters, other);
- The description of information resources in the Internet in various subject domains (education, scientifically technical information, hydrometeorology, etc.);
- Development of formats of an exchange given between the various applications, distributed servers.

Use XML in 1999-2001 for creation of dynamic pages on Web portal has allowed revealing directions of its effective application. So on the basis of language XML in RIHMI-WDC the circuit of the description of information resources is developed in the field of research of sea environment and sea activity, with which help the description of more than 600 information resources is made. Language XML is applied also to the description of metadata objects (items of information on the projects, experts, formats etc.).