



EU Life-project  
DEMOWAD

## Data Management in International Monitoring Programs

Joint Workshop of the  
European Environment Agency (EEA) and the  
Common Wadden Sea Secretariat (CWSS)

Copenhagen 18<sup>th</sup>/19<sup>th</sup> February 1998

---

### Final Draft Minutes

#### 1. OPENING AND INTRODUCTION TO THE WORKSHOP

The meeting was opened by Mr. Evangelos Papathanassiou (EEA). He welcomed all participants in Copenhagen also on behalf of the Common Wadden Sea Secretariat. The agenda of the workshop was adopted as attached in Annex I. The list of participants is in Annex II.

Mr. Papathanassiou informed the workshop about the present activities of the EEA regarding collection of data and information for the Dobø Report.

He stressed that the EEA aims at optimizing and facilitating the flow of data and information from the originators in the different EU countries to the EEA. The LIFE DEMOWAD Project can be regarded as a step forward to a better comparability and accessibility of monitoring data from different countries.

In the Wadden Sea Cooperation, a Trilateral Monitoring and Assessment Program (TMAP) was elaborated by the LIFE DEMOWAD Project in 1996 and 1997. At last year's Trilateral Ministerial Conference (October 1997), the Ministers adopted common TMAP guidelines for a common package of parameters and a trilateral data management system which will be implemented in accordance with the State Declaration as soon as possible and which experiences will be evaluated at the next Wadden Sea Conference.

The aim of this EEA/CWSS workshop was threefold:

- to discuss concepts and implementations of existing data management systems in international monitoring programs (e.g. to evaluate the results of the DEMOWAD project regarding data management in the TMAP);
- to discuss technical, as well as, administrative aspects of data handling in international monitoring programs (e.g. main gaps and obstacles at the technical, administrative and political level);
- to discuss possibilities for future cooperation (e.g., by defining fields of cooperation, common work programs or projects).

## 2. PRESENTATIONS AT THE WORKSHOP

Six data management systems were presented and are summarized below. The complete contributions are attached in Annex III.

Mr. M. Joanny (IFREMER) presented a European-wide inventory of several national data management systems for marine environmental information. The systems differ considerably in their concepts, functionality, contents, quality and purposes. As a main result, harmonization efforts on the international level were strongly recommended in order to improve the output of comparable data and products.

The ICES Data Centre, represented by Mr. J. R. Larsen (ICES), is the international data center for marine environmental data with the most experiences in Europe. The main task of the ICES Data Centre is to compile, store and analyze data on the marine environment and on marine fisheries activities. This involves screening the quality and consistency of reported data and make it available, in preprocessed form, for expert workshops and assessment groups.

G. Luerßen (CWSS) introduced the trilateral DEMOWAD (EU Life project) data management system to the meeting. DEMOWAD makes Wadden Sea monitoring data comparable and exchangeable with decentral and autonomous, but identically structured database units. The up-to-date concept could be used as model and case study for further developments.

The AMAP data management, presented by S. Wilson, utilizes thematic data centers to compile data from various wide-spread monitoring and research data sources, to respond to the general heterogeneous data situation. Different methods such as GIS, modeling and statistics have been used in preparation of the AMAP assessment which has recently been published.

H. Niesing (RWS) presented the REMSSBOT (Regional Environmental Management Support System Based on Telematics) project as an architecture for environmental information service based on an access to databases which are physically separated. The databases can be of different type (Oracle, Ingres, Informix, etc.) or structure and are accessible through the data catalogue (CDS) and telematical extensions. The web-based user interface allows the users to navigate in the meta data of several databases and a simple query making and storing functionality is possible. The system provides environmental managers with comprehensive but not harmonized information and real data of the connected databases.

J. Kroos (RIKZ) introduced the Dutch water data management system DONAR to the meeting. The whole waterbound environmental data is stored in a central database and partly copied in decentral autonomous databases. Data exchange is carried out only inside the Rijkswaterstaat via WAN. Wadden Sea data will be provided on the trilateral level by the DEMOWAD Unit.

Mr. L. Valdés (IEO) presented the IEO Pelagic Ecology Data Base in a written contribution. The data management system attaches importance to data reporting formats and data quality control. It also provides the user, besides with data in common formats, with analyzing tools.

### **3. CONCLUSIONS AND RECOMMENDATIONS**

On the second day, the workshop discussed and listed the main technical and administrative obstacles in the cooperation of national and international data management and elaborated recommendations for selected main issues.

The assessment of the European environment needs a common environmental information base. Therefore, the available and relevant data for the EEA has to be delivered from the European countries through their reporting system to the regional intergovernmental organizations and conventions to the EEA, in order to avoid duplication of work. The AMAP, HELCOM and OSPAR deliver, as an example of successful data management, harmonized data and information of several marine monitoring parameters to ICES and perform regular assessments.

Unfortunately the data and information flow of the most other relevant environmental monitoring parameters to the European level does not work sufficiently, although the EEA needs only aggregated data or the results and summarized output of the national environmental data management systems.

Several administrative and technical gaps and difficulties on the national and international level were identified. Besides the poor existing data flow, the meeting also discussed the problem of missing definitions of data quality and reliability, reporting products, and possibilities for future cooperation to improve the data exchange from the national to the international level.

The following technical and administrative aspects of the insufficient data management on the European level were discussed and recommendations for the EEA and the Trilateral Cooperation were made.

#### **3.1 Information exchange**

The European countries have developed and maintained their own data management models, which focus on national interests but also cater for the needs of reporting to international monitoring programs. Although these different models are in most cases sufficient for the national requirements, the information or the delivered products are, in a number of cases, not comparable on an international level. For a better tuning of the different national data management systems, a continuous information exchange between experts on the international level is an important element.

Experiences show that the European Environmental Agency (EEA) has difficulties to get substantial environmental data from the countries, mainly because the direct contact is insufficient and there is only little response on questionnaires and data inquiries from the countries. Furthermore, if like in several countries autonomous regions and institutions prevent data delivery or send incorrect data to the national level, to forward data to the EEA is not possible. A better information exchange and further development of the EIONET (European Information and Observation Network) could improve the situation.

A sufficient exchange of environmental information between the national and European level needs a suitable and practical management system with a defined network of institutions, responsible persons for coordination and reporting schedules and formats. This approach would give the basic technical and administrative requirements for the information exchange.

## Recommendation (1):

**EEA:**

The EEA should address the definition of a practical environmental data reporting system to appropriate international fora, e.g. the Interregional Forum of the EEA.

**Trilateral Cooperation:**

In parallel, bilateral contacts should be improved and extended, e.g. by establishing a permanent contact between the DEMOWAD System and other international data management systems.

### 3.2 Common standards

For an international base of environmental information and a successful assessment of the available data, at least minimum standards and guidelines for environmental data exchange and reporting are required. The output of the different national data management systems and the compatibility of information exchange of national and international data handling systems is more important than the technical background of the different systems. The national interests in developing common international standards is limited, because the already existing and used national standards are mostly regarded as sufficient for national purposes. The example of marine data collection carried out by ICES, AMAP, HELCOM and OSPAR demonstrates that an international data management system is feasible.

The Interregional Forum could facilitate the process of data and information collecting by establishing international minimum standards of data types, data quality and definition of exchange formats for data reporting.

The international commitments of data reporting are a compelling driving force for standardization of data formats and products. Therefore, the EEA as information collector can improve the standardization of data output and increase the data flow. The appropriate requirements have to be defined by the EEA since it is important for the countries to know the reporting requirements of the EEA in the near future. The national information output can entail harmonized data, products with comparable functionality or simple standardized data assessments. To achieve this minimum aim, the countries are requested to establish national environmental databases with minimum standards of output information. National data has to be aggregated to these minimum standards and the information could then reported to the EEA. To implement this procedure, an action plan for national data handling is demanded. The DGXII supports the development of data management systems with the experiences of MASTIII.

The EEA will publish the state of European Environment by the end of 1998. After that "Yearly Indicator Reports" will be published in order to be able to assess the European environment each year. Detailed guidelines on required environmental information for the "Yearly Indicator Reports" will be available in the first half of 1999.

The DEMOWAD project developed the minimum standard for a trilateral data management. The presented prototype allows the data exchange of two parameters. The remaining parameters will be implemented stepwise in close cooperation with international organizations using international standards.

## RECOMMENDATION (2):

**EEA:**

Common minimum standard outputs for environmental information should be developed. The EEA will define and develop standards at the end of 1998 and discuss them on the Interregional Forum in March 1999.

**Trilateral Cooperation:**

The development of new database structures and data exchange formats for biological monitoring parameters in the TMAP should be carried out in close connection with the work of ICES and other relevant groups.

### 3.3 Data quality

To guarantee a high data quality including a documented and continuous data flow on the European level, common standards and policies on data quality are urgently needed. The EEA will propose to discuss these mentioned above issues on the next Interregional Forum in the beginning of 1999.

The meeting regarded a document describing the quality assurance is necessary when delivering data from the national to an international level. It should describe how the delivered data is collected, processed, published and generally used and how the data originators are involved in this process.

## RECOMMENDATION (3):

**EEA:**

On the international level, the EEA should support and facilitate the harmonization of the data quality in the next years.

**Trilateral Cooperation:**

In the TMAP, a data policy document should be elaborated, which defines the dealing with delivered data and guarantees the data quality.

### 3.4 Data quality/documentation

The comparability of international data also depends on the associated meta data. A standard definition of meta data handling and meta databases to guarantee the quality and reliability of this data is presently not available. On the meta data level ownership and other policy issues, like "who can use what data in which way", and on processing and data flow, documentation must be stored.

The Catalogue of Data Sources (CDS) of the EEA is a first step towards an international harmonization of meta data. OSPAR has used data quality rules for 6 years now.

The usage of the CDS in the REMSSBOT project could serve as a model in this field.

#### **RECOMMENDATION (4):**

**EEA:**

The EEA should support the use of the CDS in close contact with national data handling activities in the European countries.

### **3.5 Data delivery**

The meeting regarded the involvement of the data originators as an important issue to increase the acceptance of a cooperation on an international level.

To increase the acceptance of national and European data management systems by data deliverers and data originators, it is very important to keep continuous contact with them on an international level, to mark their products with their names and add additional links to their work. If this is the case, the data originators can identify their own data on the international level and recognize how the data is used.

On international level widely respected organizations built up databases and perform data exchange and environmental assessments although on the national level, data delivery to international organizations always causes additional effort. The capacity of the national data management is already burdened with the collection of data, with organizing the national data administration and problems with data owner-rights, etc.

For each project or program which intends to bring data on an international level, right from the beginning, a budget for the data handling has to be included or contracts are needed to build separate data products. The close contact between the data originator and the data recipient is important for the quality of the product and prevents that the quality assurance of data don't just take place during the assessment process. Beside the data delivery a documentation of the quality of data and the connected products is always necessary.

#### **RECOMMENDATION (5):**

**EEA and Trilateral Cooperation:**

A reporting format, that can be used for data reporting to different organizations and databases should be developed and established in close contact with the data deliverers to minimize the effort the data compilation for delivery.

### **3.6 Continuation of DEMOWAD**

Because the technical development of data management systems is continuously ongoing, model and case studies on an international level are always requested.

The DEMOWAD prototype started in 1996 and will implement further monitoring parameters and additional functionality during the next years. The trilateral data management system is a good model and case study regarding the concept and realization of an international data management system. Other international cooperation or regions can benefit from experiences made with the DEMOWAD model.

#### **RECOMMENDATION (6):**

##### **Trilateral Cooperation:**

The DEMOWAD data management system should be further developed and extended to a fully functional and operational data management system of the TMAP.

## Agenda

## Annex I

### Wednesday, 18 February 1998

- 14:00 *Mr. E. Papathanassiou* (EEA, Copenhagen)  
Opening and introduction to the workshop
- 14:30 *Mr. M. Joanny* (IFREMER, ETC/MC, France)  
Title: Overview of the marine monitoring data management in Europe  
(Author(s) Michel Joanny and Muriel Relandeau)
- 15:00 *Mr. J. R. Larsen* (ICES, Denmark)  
Title: The management of marine environmental data at the ICES Data Centre
- 15:30 *Mr. G. Lüerßen* (CWSS, Germany)  
Title: DEMOWAD: Data Management of Wadden Sea monitoring data within the Trilateral Monitoring and Assessment Program (TMAP)
- 16:30 *Mr. S. Wilson* (AMAP, Norway)  
Title: Data handling activities and experiences within the Arctic Monitoring and Assessment Programme (AMAP)
- 17:00 *Mr. H. Niesing* (RIKZ, The Netherlands)  
Title: REMSSBOT: An architecture to access environmental information independent of structure and location
- 17:30 *Mr. J. Kroos* (RIKZ, The Netherlands)  
Title: Data management in the Netherlands
- 18:00 End of session
- 19:00 Dinner

### Thursday, 19 February 1998

- 08:45 - 09:45 On-line presentation of DEMOWAD, REMSSBOT, ICES and AMAP.
- 09:45 - 12:30 Plenary discussion:  
Technical and administrative aspects of data management systems
1. Discussion on the presented data management systems.
  2. Identifying the main gaps and obstacles in international data management.
  3. Future cooperation on international and bilateral level.
- 12:30 - 13:00 Conclusions and Recommendations
- 13:00 Closure of the meeting

## Participants of the Workshop

## Annex II

Mr Sigfus Bjarnason European Environment Agency (EEA) Kongens Nytorv 6 DK-1050 Copenhagen K	Tel.: Fax.: e-mail: Sigfus.Bjarnason@eea.eu.int
Mr Stig Carlberg Swedish Meteorological and Hydrological Institute (SMHI) Byggnad 31 Nya Varvet SE-426 71 V. Frölunda	Tel.: +46 31 751 8976 Fax.: +46 31 751 8980 e-mail: stig.carlberg@smhi.se
Mr Michel Joanny IFREMER Centre de Brest Direction environnement et aménagement du littoral BP 70 F-29280 Plouzane	Tel: +33 2 98 22 43 58 Fax: +33 2 98 22 45 48 e-mail: mjoanny@ifremer.fr
Mr Hansjörg Krasemann GKSS-Forschungszentrum Geesthacht GmbH Max-Planck-Str. 1 D-21502 Geesthacht	Tel.: +49 (0)41 52 87 18 77 Fax.: +49 (0)41 52 87 18 88 E-mail: krasemann@gkss.de
Mr Jan Kroos RIKZ P.O. Box 20907 NL-2500 EX Den Haag	Tel.: +31 (0)70 31 14 512 Fax.: +31 (0)70 31 14 500 e-mail: kroos@rikz.rws.minvenw.nl
Mr Jan Rene Larsen International Council for the Exploration of the Sea (ICES) Palægade 2-4 DK-1261 Copenhagen K	Tel: +45 33 15 42 25 Fax: +45 33 93 42 14 e-mail: jrl@ices.dk
Mr Henrik Søren Larsen Environment Protection Agency Chemical Division Strandgade 29 DK-1401 Copenhagen	Tel: +45 32 66 02 92 Fax: +45 32 66 02 61 e-mail: hel@mst.dk
Mr Gerold Lüerßen Common Wadden Sea Secretariat (CWSS) Virchowstr. 1 D-26382 Wilhelmshaven	Tel.: +49 44 21 91 08 16 Fax.: +49 44 21 91 08 30 e-mail: luerssen@cwss.nord.de
Mr Harald Marencic Common Wadden Sea Secretariat (CWSS) Virchowstr. 1 D-26382 Wilhelmshaven	Tel.: +49 44 21 91 08 15 Fax.: +49 44 21 91 08 30 e-mail: marencic@cwss.nord.de
Mr Hugo Niesing RWS Koestraat 30 NL-4331 KX Middelburg	Tel: +31 (0)118 686 401 Fax: +31 (0)118 686 231 e-mail: n.niesing@dzl.rws.minvenw.nl
Ms Graça Noronha Ministerio do Ambiente	Tel: +35 11 472 83 33 Fax: +35 11 471 90 74

Direcção Geral do Ambiente  
Rua da Murgueira - Zambujal  
P-2720 Amadora

e-mail: g.noronha@dga.min-amb.pt

Mr Evangelos Papathanassiou  
European Environment Agency (EEA)  
Kongens Nytorv 6  
DK-1050 Copenhagen K

Tel: +45 33 36 71 55  
Fax: +45 33 36 71 99  
e-mail: vpapath@eea.eu.int

Ms Maren Pommerencke  
Roggenkamp 3  
D-22305 Hamburg

Tel:  
Fax:  
e-mail: marenpommerencke@enator.de

Mr Peter Sandbeck  
Danmarks Miljøundersøgelser-NERI (Roskilde)  
Dept. of Marine Environment and Microbiology  
P.O. Box 358  
DK-4000 Roskilde

Tel.: +45 46 30 12 61 op. 12 00  
Fax.: +45 46 30 12 11  
e-mail: ps@dmu.dk

Mr Luis Valdés  
Instituto Español de Oceanografía (IEO)  
P.O. Box 240  
E-39008 Santander

Tel: +34 42 27 50 62  
Fax: +34 42 27 50 72  
e-mail: luis.valdes@st.ieo.es

Mr Simon Wilson  
Arctic Monitoring and Assessment Programme (AMAP)  
AMAP Secretariat  
P.O. Box 8100 Dep.  
N-0032 Oslo

Tel./Fax: +31 10 466 29 89  
Alternative: AMAP office  
Tel: +47 22 57 36 36  
Fax: +47 22 67 67 06  
e-mail: s.wilson@inter.nl.net

**ANNEX III**

**Summary of Workshop Presentations**  
**Marine monitoring data management workshop (Copenhagen 18-19/02/98)**

## Short summary of the presentation made by Michel Joanny

### National status of marine monitoring data management

The information and recommendations presented hereafter were prepared for a report to the European Environment Agency (EEA) by the European Topic Center on Marine and Coastal Environment (ETC/MCE).

In order to give a quick overview of the national marine monitoring data management in Europe, the following table summarize their main characteristics.

Country	Name of the database	Date of operating	Type	Size
BELGIUM	IDOD	Under project	-	-
DENMARK	-	1987	Local Oracle RDBMS	1200 Mb
FRANCE	QUADRIGE	1996	Client/server Sybase RDBMS	800 Mb
GERMANY	MUDAB	1993	Client/server Ingres RDBMS	2.25 Mrecords
GREECE		Planned project	Local Oracle RDBMS	
IRELAND	-	1997	Local Ingres RDBMS	-
ITALY	-	-	-	-
NETHERLANDS	DONAR	1993	central/decentral Ingres RDBMS	15 Gb
NORWAY	BIOTASYS, SEDIMSYS	1992	Client/server Sybase RDBMS	50 Mb
PORTUGAL	-	-	data managed in different places	-
SPAIN	-	-	data managed in different places	-
SWEDEN	SHARK, BED, BIOMAD	-	-	-
UNITED KINGDOM	-	1996	Local MS Access on PC	10 Mb

## **Data harmonisation**

### Conditions for feasibility

**The quality of the databases**

**The state of the reference data systems**

**The accessibility of the databases**

**The capabilities of the user information system**

### Recommendations

#### **In the short term**

1. Put in contact people of marine monitoring database management for exchange of experience and mutual improvement.
2. Define the needs for harmonized improvements of national data management.  
Recommendations

#### **In the medium term**

3. Harmonize reference data and create a common multilingual data catalogue.
4. Create a common exchange format.
5. Define common products (aggregated data, indicators) to be elaborated in such a way every country could contribute to their achievement with its own system.

#### **In the long term**

5. Define the data management procedures for the EURO-SEAWATERNET, based on national databases for raw data and a european database for aggregated data.

## **The marine monitoring data management system in France**

### Architecture of the information system

The Quadrige database has a client/server architecture. Clients are connected to the server via local Ethernet network, or via Internet, or via a telephone line. The server is located in the Ifremer centre in Brest (Brittany - France). Around 100 users are registered to enter or extract data in real time from 50 PCs in 20 sites along the French coast.

### Description of the database

The Quadrige database (800 Mb) is a SYBASE relational database management system. It was developed using PowerBuilder development tool in 1994-1996, and put in operation in June 1996.

This database is the result of an important project which objective was to merge the content of the databases of three different national monitoring programmes into a unique up-to-date and efficient database. It thus contains the monitoring results of the following national monitoring programmes :

- the "Réseau National d'Observation de la qualité du milieu marin (RNO)", the national monitoring network for pollutants and biological effects, started in 1974,
- the "Réseau de surveillance du phytoplancton (REPHY)", the national monitoring network for

phytoplankton, started in 1984,

- the "Réseau de surveillance microbiologique (REMI)", the national monitoring network for microbiology, started in 1989.

Data from the ecological monitoring of coastal nuclear power plants, started in 1973, are also included in the database, i.e. zooplankton and benthos data.

Data are directly entered in the database from the different sites along the French coastline, with a Windows-based application on PCs. Controls are made at many different steps of the input process. The user has the possibility to check and correct input errors, before validating the set of data entered. After the validation, no modification is allowed without a special procedure including advice from the programme leader and the database administrator.

Data can be retrieved from the database with a multi-criteria Windows-based application tool. The user get a text file he can work on its own.

## Quality assurance procedures

The conception and development of the application were conducted in accordance with the best quality procedures available on every step of the process. The most crucial were the following :

- Close cooperation between thematic and computer people, structured within the project in working groups. Those working groups are in charge of the definition and harmonization when necessary of all the thematic entities, and specify the new system with the computer team.
- Use of a recognised methodology for the definition of the data model and the treatment model.
- Use of a CASE (Computer Aided Software Engineering) in order to improve quality (graphic data modelling, interactive automatic documentation and reporting, tools for model control).
- Use of pre-defined ergonomic and development specifications (Windows norms, for this application).
- Use of extensive test procedures of the application, to check that the norms are fulfilled.

The data model was carefully designed by french working groups of experts from Ifremer and outside, in order to harmonise concepts between user in the different fields of activity (physico-chemistry, chemistry and biology). Common reference data were defined, and the final data model can include every monitoring result coming from a sampling in the field.

### **The French national network on water-related data**

In order to improve the availability of water-related information, French Ministry of the Environment, the six Water Agencies, the Higher Council for Fisheries and the French Institute of Environment established the French national network on water-related data. It formalise the cooperation between the major institutional water-related data producers, and is based on two operational services : The national secretariat on water resources data (SANDRE), and the National water-related data bank (BNDE).

The SANDRE lists the available data on inland surface and ground waters, and seawater. It defines agreements, references and standards necessary for a homogeneous and modern data network and update the catalogue together with data producers and data bank managers. It plans the computerized interfaces to develop between data banks to enable computerized exchanges of data with the network partners.

The BNDE aims at setting up a computerized exchange network on data collected by various professional or regional data banks, already existing or to be created, in order to enable an harmonized and simplified consultation and to implement the necessary synthesis for informing national public authorities and the European Agency for Environment. For the completion of its task, the RNDE is setting up an Internet server called the "Guichet RNDE" that could be a fruitful experiment for the EEA.



# The DEMOWAD data management system

*G. Lüerßen, H. Boer, H.S. Larsen, M. Pommerencke, P. Sandbeck*

## **Introduction**

The Wadden Sea is a shallow sea extending along the North Sea coast of The Netherlands, Germany and Denmark and is the largest wetland in Europe. As a high dynamic ecosystem with a large number of species and high biological productivity, it is a unique area worthy of protection.

In 1978, the Trilateral Cooperation of the three Wadden Sea states held their first Governmental Conference, and in 1987 the Common Wadden Sea Secretariat was established to coordinate and facilitate the work of the cooperation. The Trilateral Wadden Sea Plan (TWP), a common management plan, was adopted in 1997 and will be the basis for the further cooperation.

1991 the Trilateral Governmental Conference in Esbjerg decided to cooperate in scientific research and monitoring and to implement a Common Wadden Sea Monitoring Program. In the period 1992 - 1993, a basic concept for an integrated Trilateral Monitoring and Assessment Program (TMAP) was elaborated, followed by an implementation plan carried out within the DEMOWAD project which is running from April 1995 to March 1998. The TGC 8 1997 in Stade agreed to implement a common package of TMAP parameters and the associated data handling.

## **TMAP**

The TMAP combines a comprehensive set of monitoring parameters with concomitant ecosystem research. The trilateral data management has to make the derived monitoring data available for the scientific assessment of the status of the ecosystem and for the assessment of the implementation status of ecological targets as laid down in the trilateral management plan. The objectives of the DEMOWAD data handling project was to fulfill these requirements with the development of a prototype of a data management system.

### **DEMOWAD data handling project (organizational matters)**

The DEMOWAD data handling project started in the beginning of 1996. For the development and support of the project, three national databases were appointed. The Danish Wadden Sea Database (DAWAD) at the National Environmental Research Institute (NERI) in Roskilde, the German "Wattenmeerinformationssystem" (WATiS) at the GKSS in Geesthacht and the Dutch database (DONAR) at the National Institute for Coastal and Marine Management (RIKZ) in Haren. At each of these data centers, data handling experts were employed. The coordination of the project was carried out by the CWSS in Wilhelmshaven.

The tasks of the project was to define and implement a prototype of a data management system, which makes TMAP data available for trilateral monitoring experts and gives an overview of running projects and their status of data collection. At the beginning, three main tasks were identified according to the outlines and principal requirements of the project. The Netherlands was responsible for a harmonized database structure and data exchange format, Denmark for a data catalogue to give an overview of the data, and Germany for the networking respective the data access via network.

### **Definition phase**

The project started with a detailed definition of the product, which had be able to solve the requirements of the trilateral data management system. For each of the three main tasks an inventory, a definition and an analysis of the most international standardized data exchange formats, data catalogues and data transfer systems were made. An analysis of the existing technical and political possibilities of the participating data centers including a trilateral questionnaire to get the requirements of the later users was carried out. Together with criteria like costs, efforts, user-friendliness, interactivity and security matters, different and detailed options for the implementation of the three tasks were elaborated.

For example, the networking task options for different data transfer possibilities, user interfaces and server solutions were elaborated. These options were presented to the Trilateral Monitoring and Assessment Group (TMAG) and the Trilateral Working Group (TWG) for decision to guarantee, besides the technical feasibility, also the administrative acceptance of the data management system.

## The difficulties

At the beginning of the project, it turned out that three main issues had to be solved to make monitoring data available in a complete, consistent and correct form to safeguard a sufficient assessment of the data on a trilateral level.

**Structure:** The data was stored in different not comparable structures with different co-variables, without or insufficient meta information and with different storage techniques.

**Contents:** The data of a specific monitoring parameter was available in one data center, but not in another. Beside different contents, the data was qualitatively or quantitatively incomplete, inconsistent and not equally or not at all quality checked .

**Location:** The data was stored either at the national Wadden Sea data centers, different institutes or private persons' and other locations.

Technical obstacles like different databases and operating systems at the Wadden Sea data centers, administrative problems such as data security demanding firewalls, copyrights on data, financial constraints such as the demand to develop a cost-neutral system, the minimization of implementation efforts and maintenance of the future data management system completed the complexity of the task to develop a prototype of a trilateral data management system.

## The solution

The existing preconditions of the data centers in the three countries resulted in a data management system with the following attributes:

- de-central,* to be close to the data originators, to have TMAP data in definite administrative boundaries and to get more acceptance by cooperation with local authorities
- flexible,* to give attention to local conditions and to use already existing data management systems
- simple and open,* to allow for an installation and operating the system with low costs and to be open for further development

Harmonizing of data structures:

To make data comparable and calculable in a standardized way it is necessary to develop homogeneous data structures. The project decided to implement this requirement in the data model and data structure of a trilateral harmonized so-called TMAP database. Because of the very different monitoring parameters, which had already to be taken into account in the implementation of the prototype (data of breeding birds and heavy metals in mussels are very different), a subject-orientated database model was chosen. A subject was defined as one parameter or in the view of data experts as a class of very similar parameters. The prototype subject "contaminants in biota" includes the parameter "heavy metals in mussels", but will also cover the parameter "contaminants in breeding bird eggs" in future.

Data catalogue for overview of contents:

To provide continuously an overview of the stored raw or real data of the TMAP, an integrated data catalogue was implemented on top of the TMAP database. The

design of the catalogue meets international standards and was adapted to the subject orientated structure of the database. The catalogue provides the public with an overview of existing TMAP data and allows trilateral experts a selection and download of the stored raw data.

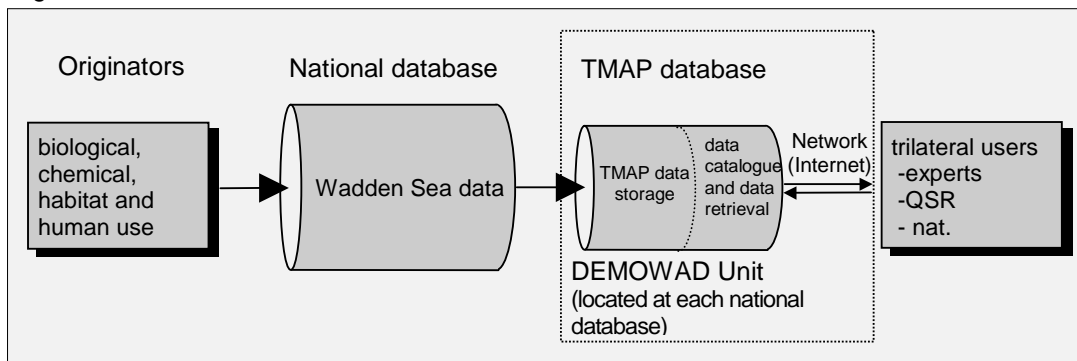
Link to different data locations:

For information transfer, the Internet as the most general, available and convenient network has been chosen. The services of the Internet allow for a close contact and data exchange between different data centers and the trilateral users. Three different databases used as TMAP database were successfully connected: MS-Access in DK, Oracle in Germany and Ingres in the NL.

### The DEMOWAD Unit

Because the participating countries insisted, from the beginning of the project, on an autonomous data storage, the idea of a trilateral commonly but de-centrally orientated data management concept was favored. Together with the defined requirements of a modern data management, the concept of the DEMOWAD Unit with three main components, a commonly structured TMAP database, an integrated data catalogue and a network access was developed and installed at each participating national Wadden Sea database.

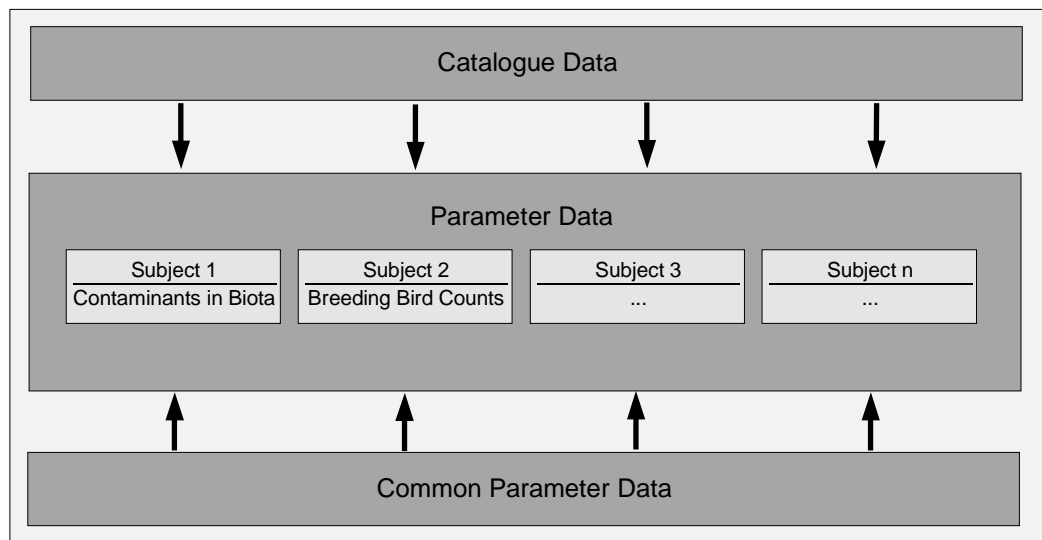
Fig.1: Data flow from national to trilateral level



The core of the DEMOWAD Unit is the at each location similarly structured TMAP database, which gets the monitoring data (biological, chemical, habitat and human use data) by regularly updating from the national database. Therefore, a relational database with quality checked monitoring data is requested for the Unit.

The database contains, in its parameter data part, the real or raw data of the different subjects attended by common parameter tables, which store common data such as species, units, and areas for more than one subject. The who, what, where and when of the measured raw data is recorded in an integrated data catalogue, which allows the access to the real data.

Fig. 2: Prototype of the TMAP database



A Web server in connection with a CGI program establishes the link of the TMAP database to the Internet. When the users enter the DEMOWAD Unit via an Internet browser, they get a HTML page as search form. The selection items within the different menus are retrieved via SQL from the database and therefore are always up-to-date with the contents of the database. After the selection of certain items, the CGI program creates with these selected items a new SQL-query and creates dynamically with these results an updated new search form. This functionality allows the users to navigate through the database with the guarantee to see the actual contents of the TMAP database.

Trilateral users are also allowed to retrieve the real data from the Unit. After selection of certain data in the mentioned way, the data is retrieved from the database and transformed into the trilateral data exchange format, which is a set of well defined ASCII formatted tables in different files. These files can be easily downloaded as an archived and compressed file from a FTP server to the local computer of the user. After decompressing and splitting the downloaded file into the original exchange files, the further processing of the data with standard word processors and spread sheet programs is possible.

## Evaluation of the prototype

### Technical issues

The used UNIX operating systems in Germany and The Netherlands caused less problems and efforts during the installation of the DEMOWAD Unit than the Windows NT system in Denmark. The reason is that UNIX has the better process management, is more flexible and open and the Perl language used for the CGI program is better integrated in UNIX operating systems.

## Costs

The costs for Hard- and Software were low. Only the database software had to be paid for, Web server and the Perl language were free of charge. In case of high security demands like in NL and DK, money had to be invested in server hardware outside the firewalls. The development and installation of the prototype had been done with approximately 1 person per center in 18 months and 1 coordinator. The maintenance effort of the system is low. Half a person is needed to maintain a DEMOWAD Unit.

## Trilateral user test

After installation of the prototype at the Wadden Sea databases, the trilateral monitoring experts tested the data management system. The evaluation of this test showed that the users were satisfied with the system. The most technical problems occurred during downloading and decompressing the selected data, which could be caused by the different platforms used and the operating systems and missing experience of the users.

## The prototype

The DEMOWAD Unit, based on proofed data handling techniques, uses an advanced technology with even more options in the future, allows for a fast and reliable data exchange, generates relatively low costs for installation and maintenance, turns out as a user friendly with an easy to handle and always up-to-date data catalogue. Furthermore, the overall structure and design of the Unit allows for an implementation at each data center which intends to join the trilateral data exchange.

## Maintenance

The main difficulty of the operating DEMOWAD Units will be the TMAP data update from the underlying Wadden Sea databases. Within the DEMOWAD project, the entering of the prototype parameter data into the TMAP database caused the most troubles and efforts, because the available data was incomplete, inconsistent, not quality checked and stored in different formats and media.

To solve these problems agreements on data delivery, data copyrights and data reporting formats have to be defined and organized on administrative and political level.

## Future work

The further implementation of the TMAP parameter package, as agreed by the ministers of the TGC 8 in 1997 has to be done in close contact with the monitoring experts in the next years. The data handling has to be an integrated part of the different trilateral monitoring programs to smooth the data flow from the data originators to the TMAP databases. That means that database structures for biological, chemical, habitat and human use data have to be developed in close cooperation with other international programs and experts.

Like the already installed map applet, which allows for the geographical selection of data via a map and the news applet, which will be implemented in the near future

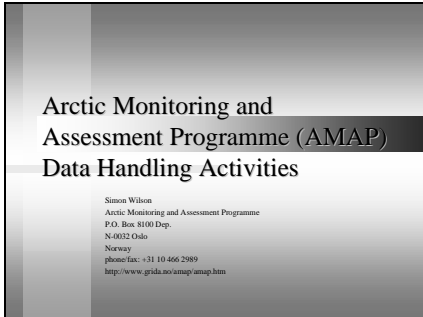
and allows an easy and fast message transfer between the monitoring experts, the DEMOWAD Unit will have in future only a single Java Interface between the Web server and the TMAP database. This allows for more platform independent development, the usage of advanced Web-database technologies and an opening of the system to integrated and interactive data processing, data presentation and data assessment.

A future work is also the implementation of GIS data in the DEMOWAD Unit to allow for a sophisticated geographical assessment of the TMAP data.

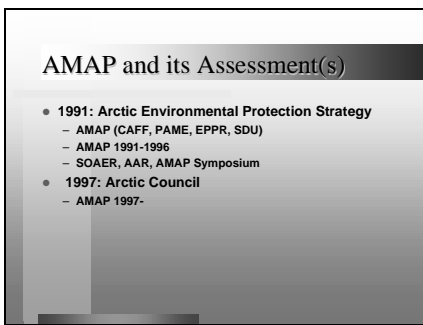
With advanced data processing tools, the publication of assessment results and as discussion platform, the presented data management system will be the basic component for a future trilateral Wadden Sea information, management and presentation system.

# Presentation of the Arctic Monitoring and Assessment Programme.

Simon Wilson, AMAP

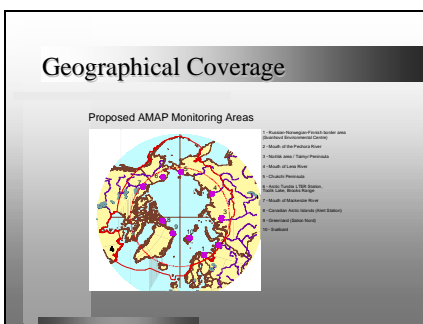


The Arctic Monitoring and Assessment Programme (AMAP) is one of 5 programmes established to implement the Arctic Environmental Protection Strategy (AEPS). Other programmes deal with Conservation of Arctic Flora and Fauna (CAFF), Protection of the Marine Environment (PAME), Emergency Prevention, Preparedness and Response (EPPR) and Sustainable Development and Utilisation (SDU).

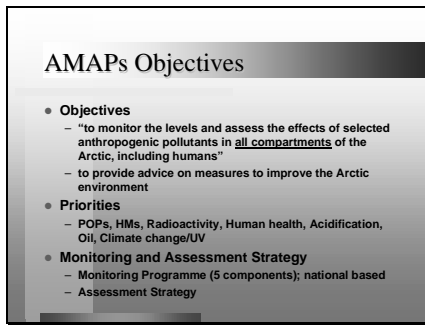


The AEPS was adopted by the eight Arctic countries (Canada, Denmark/Greenland, Finland, Iceland, Norway, Russia, Sweden and the USA) in 1991. In 1997, the Arctic Council was established, and the AEPS programmes and activities are now part of the Arctic Council initiative.

Between 1991 and 1997, AMAP developed and implemented its monitoring programme and prepared its first assessment.

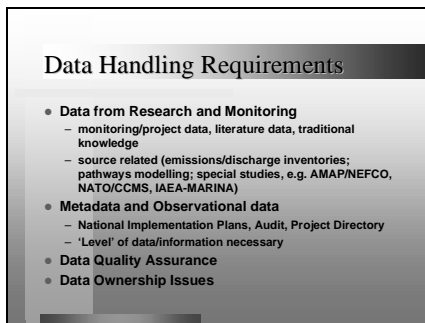


This assessment took the form of a comprehensive 'State of the Arctic Environment Report' delivered to Ministers of the Arctic countries for their consideration at a meeting in June 1997. The full scientific documentation is contained in a separate (AAR) report. In conjunction with the publication of its assessment, AMAP also arranged a major symposium (ca. 450 participants) in June 1997 to facilitate a discussion of the conclusions of the AMAP assessment and to encourage the necessary dialogue between scientists, representatives of northern inhabitants and environmental managers and decision makers, etc.



AMAP’s principle objectives are: ‘to monitor the levels and assess the effects of selected anthropogenic pollutants in all compartments of the Arctic, including humans’, and ‘to provide advice on measures to improve the Arctic environment’. AMAP priority issues included persistent organic pollutants, heavy metals, radioactivity; acidification and petroleum hydrocarbons (on a sub-regional basis); and climate change and UV; and their effects on Arctic ecosystems and humans, including human health.

To avoid duplication of existing and ongoing work, the AMAP monitoring programme was based largely on existing national and international activities, adapted or expanded where necessary to meet AMAP requirements. Countries developed their National Implementation Plans (NIPs) to fulfil AMAP goals, and the resulting ‘programme’ was subjected to an audit process, the outcome of which resulted in further adjustments. The AMAP monitoring programme includes both ‘traditional’ monitoring and research activities and covers the atmospheric, marine, terrestrial and freshwater environments, as well as a human health component.



In conducting its assessment, AMAP faced the challenges of compiling and incorporating data from a wide range of sources, both observational data (i.e., monitoring and research data from sources in all 8 AMAP countries, co-operating international programmes, and others), information published in the scientific literature and in institute/project reports, etc., and also traditional knowledge. Data quality assurance was a major consideration as were ‘data ownership’ issues. Development of an appropriate ‘data ownership’ agreement was critical, in particular in facilitating access to the most recent research data.

Metadata was compiled in the form of a project directory describing some 500 projects and programmes to enable the assessment process to locate relevant sources of information.

### Data Handling Solutions

- Thematic Data Centres
  - Atmospheric: NILU
  - Marine: ICES
  - Freshwater: FWI-DFO
  - Radioactivity: NRPA
  - Human Health: (Denmark 1997-)
- Assessment Methods and Data Products
  - Comparative analysis/Criteria, Modelling, GIS, Statistics
- Strengths and Weaknesses
  - 1-stop data supply for assesment & future availability (+)
  - harmonised/consistent data treatment (+)
  - organizational aspects (+/-), cost (-?)

A core 'solution' to the data handling problem adopted by AMAP was the establishment and use of 'thematic data centres' (TDCs) to handle 'observational' data arising from AMAP affiliated monitoring programmes or research projects, and to serve the AMAP assessment process. These centres (a) ensure a future accessibility to data, much of which is never otherwise properly archived, and (b) ensure that all necessary data components including supporting information for QA purposes, etc. are available to allow (c) a consistent treatment of data. Published (e.g.) literature data were also a major component of the information on which the AMAP assessment is based, but these often lack the complete descriptions of the data including the desired range of statistical parameters that is often necessary for a valid assessment process.

### International Harmonisation

- Organizational Linkages
  - OSPARCOM, HELCOM, EMEP, ICES, International Projects/Programmes (UNEP, ACSYS, IAEA, etc.)
- Monitoring/Research
  - Monitoring Guidelines and Components
- Data Centres and Data Reporting
  - harmonised formats
  - one-off workload for data originators
  - cost efficiency/cost sharing
- Assessment Methods
  - Tools, Criteria, Products

In both its monitoring and assessment activities, AMAP attempts to develop appropriate co-operation with other organisations engaged in compatible work. In some cases, AMAP relied on other organisations to contribute parts of its assessment. Harmonisation of activities between AMAP and other similar programmes engaged in, e.g., environmental monitoring, is vital to ensure cost effective information supply and avoid duplication.

### Future Developments

- AMAP 1997-
  - Targetted Assessments vs. SOAER
  - Monitoring Strategy
- Future directions
  - Environmental impact assessment support for sustainable development, Climate change/UV effects, Contaminant 'focussing' areas, Combined effects, Filling gaps

Harmonisation was also a consideration in AMAP data handling work, where, e.g., the decision by AMAP to use TDCs which are also used by other international organisations greatly facilitated possibilities for data sharing between programmes, and also minimises the data reporting work by institutes engaged in several different international programmes.

AMAP is currently developing its future activities, to build on the results of its first assessment.

## **Remssbot**

### **Access towards different environmental data sources through an integrated environmental management system.**

H. Niesing.

#### ***abstract***

The physical and administrative areas related to environmental topics may vary from a part of a region up to several countries, especially in watersystems.

Managing these areas, accessing the information available about these areas, can result in difficulties. This can cost a significant amount of time or it is unknown what information is available at the neighbouring administration.

Challenging these problems, Remssbot, (Regional Environmental Management Support System Based on Telematics) benefits from access to different information sources. Remssbot is a co-operation project among the regions Piemonte (Italy), Attica (Greece) and Scheldt (Netherlands and Flanders). Remssbot aims to improve environmental information services at the level of the water system, taking benefit of an innovative design.

The innovative aspects of the Remssbot system are the techniques used which allows the administrations and environmental managers to have access to each others data on an independent base while maintaining control over their own data. The data provider also determines what access should be given to a certain user. This concept results in several connected databases which are physically separated, but appear as one.

During the project a so-called demonstrator has been developed, this demonstrator is a tool which shows the functioning of the system with limited contents and applications which are not yet completely operational.

The developed "demonstrator" is a tool which is build up on roughly two elements, the Catalogue of Data Sources (CDS) which provides the metadata search and the telematical extensions to provide the actual information (weather documents, measured parameters etc) to the user.

#### ***Use of Remssbot in the Scheldt region.***

The high number of habitants, the high degree of industrialisation and the agricultural use of a big part of the area resulted in a considerable pressure on the river ecosystem. The ecological problems in the Scheldt basin are partly due to the enclosure of wetlands for agricultural use and more recently for a wider variety of industrial and port related, urban, safety and recreational purposes. As a result, unfortunately the river Scheldt is still one of the most polluted river systems in Western Europe.

The project in the Scheldt region concerns the needs of the water managers within the region to have access to the right information on the right moment. The main activities of the participating administrations are related to the water quality and the ecological functioning of the river Scheldt.

#### ***Project objectives :***

- Developing an environmental management support system which can be used for a wide variety of environmental topics and management levels.
- Facilitate the search for and access to specific environmental data by several well developed navigation systems.
- Increasing the quantity of environmental information and velocity of availability.
- Wider dissemination of all kind of publicly available environmental information concerning the water system using the internet access.
- Using the Catalogue of Data Sources in the data communication between regions and countries will stimulate the standardisation of terms used for a certain object, activity or location within a system. This is essential in using one European standard thesaurus for environmental subjects. The Remssbot project supports the EEA bringing this multilingual and hierarchical thesaurus into the daily practice of environmental managers.
- Stimulating integrated water management development which places the river basin as the central entity within policy plans, instead of geographical determined boundaries, such as land frontiers.

- Sharing experiences, knowledge and measured data on water management topics between administrations and their water managers concerned with an ecosystem. This can result in closer co-operation and understanding of public administrations concerning environmental information.

#### **Actual situation in the Scheldt basin on water management and information sharing**

- Each administration has his own database
- Few knowledge about activities of the different department in the catchment area
- Few use of each others information (policy, research monitoring etc.)
- No compatibility to share the information
- No creation of a central database, due to financial and managing problems.
- Differences in measuring, units used, co-ordinates etc.
- International Commission for the protection of the river Scheldt started to make policy.

#### **Main points of Remssbot Functionality**

- Knowledge of “what” is “where”
- Knowledge of “who” does “what”
- User can search in the metadata from a remote data “provider”
- User can access (and download) the environmental information from a remote “provider”
- At the “provider” database site no changes are occurring
- The complete control for data owner is guaranteed
- User cannot change anything until the information is downloaded
- Easy query making and saving and recalling possibilities
- Easy broadening to more data “providers”
- The accessed information is offered in an easy way
- For the use no manual is needed containing more as 2 pages
- Creation of 1 “virtual” database for the user
- Stimulate European standardisation process in environmental management

#### **Discussion statements:**

1. The data provider should not adapt his information (databases) for the external user, in order to make co-operation as easy as possible and reduces the costs. Secondly this will guarantee an optimal updating of the information. The data provider will do this for the users of the information at the provider’s site itself.
2. The same can be said for the metadata, because the knowledge and interest to do so is present at the data provider site. This will result in different spread information sources and metadata bases.
3. The interface and use of the system should be completely self explaining, so for the use no manual is needed at all or not containing more as 2 pages.
4. Users have to be confronted with the raw information of the connected administration in order to start talking about standardising methods, units, co-ordinates etc. First between the bordering countries, later on an European level.

# Data Management in the Netherlands

by Jan Kroos

## 1 Introduction

In the Netherlands sampling of data in case of water management is very important. For the Dutch people water is sometimes a friend and sometimes an enemy.

At one hand large areas of the country lie below sea level and their safety is at stake. Many people still have vivid memories of the storm surge disaster of 1 February 1953, when 135,000 hectares of land were inundated and 1835 people lost their lives in a single night. Tens of thousands of farm animals perished and there was vast damage to infrastructure, some of it irreparable. Since then, large parts of the sea defences along the Dutch coast have been strengthened. Implementation of the so called Delta Plan has led to the construction of dams and the barriers in the Eastern Scheldt and the New Waterway (port of Rotterdam). For the developing and the construction of these dams and barriers actual and historical data was needed. Nowadays when these barriers are operational, data is needed for maintaining and operating these barriers.

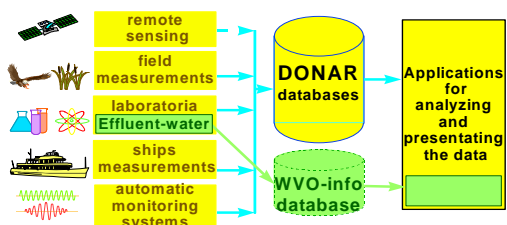
At the other hand the Netherlands are dependent of the flow of fresh water. Fresh water is used for agriculture, shipping, fishery, industry, nature and recreation. The quality and the condition of the water is also a matter of dead or alive for the people and the biological living in the country.

## 2 Datamanagement in Rijkswaterstaat

As part of the Ministry of Transport, Public Works and Water Management, in the Netherlands the Rijkswaterstaat is among other things responsible for the planning and implementation of the policy on water management and for the operational water management itself. The National Institute for Coastal and Marine Management/RIKZ provides the Rijkswaterstaat with advice and information on subjects as the sustainable use of estuaries, coasts and seas and on coastal flood protection.

To carry out these tasks Rijkswaterstaat needs a lot of information, for example:

- the current environmental condition of the water systems;
- the economic use of the waterbound environment;
- the expected trends;
- information about the relevant physical, chemical and biological processes;
- for the policy makers aggregated information.



Data and information about the waterquantity, waterquality, biological conditions, the heights of the dikes and dunes, and the depths of the shipping-roads are very important. As the Dutch people samples different "water-data" since more than two centuries, for some parameters datamanagement is implemented a long time ago.

## 2.1 Definition Data Management

The definition of datamanagement we use in the Netherlands is as follows:

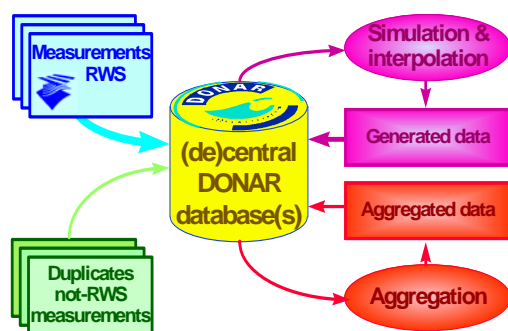
*The planning, the measuring, the documentation and by electronic way making accessible of data by a way that an optimum quality and availability of data will be succeeded, during or at the end of a (monitoring, research or advice) project.*

## 2.2 Measures taken by Rijkswaterstaat

From the seventies Rijkswaterstaat has taken different rigorous measures for implementation of datamanagement:

- Good Measuring Practise (GMP)  
Different standardized sensors, sampling methods and prescriptions are taken in use by the measuring services of the Rijkswaterstaat;
- Good Laboratory Practise (GLP)  
Different standardized analysing methods and prescriptions are taken in use by the Laboratories of Rijkswaterstaat (certifying of the Laboratories);
- Standardization of check and validation of the measurements  
Very sophisticated methods based on the newest mathematic possibilities are being developed and used by checking and validation the data;
- Unambiguous storage of data;
- Standardized and uniform datamodel for datastorage.

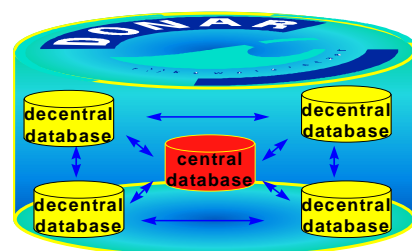
## 2.3 DONAR, the database for various data for Rijkswaterstaat



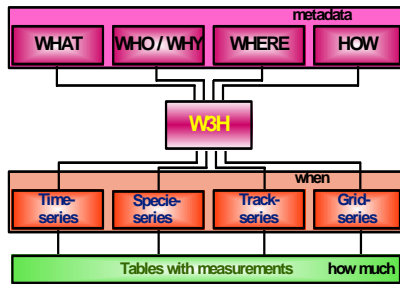
At the end of the eighties Rijkswaterstaat starts the development of a database-system for storage their data. Per year Rijkswaterstaat samples about  $10^9$  measurements. After operational use the data will be reduced to about  $75 \cdot 10^6$  measurements per year. In the beginning of the year 1998 the total size of measured data archived in databases is about  $1.4 \cdot 10^9$  measurements. Most of them are physical data. At the end of 1999 99% of these measurements are stored in the DONAR databases. It is also possible to store simulated or generated and aggregated data in the

DONAR database. *Until now no simulated or generated data are archived in the DONAR-databases.*

The DONAR-system is a set of databases based on a common datamodel a common set of procedures, rules and metadata. The software for all the DONAR databases is equal. The functionality of the databases is comparable. The metadata for all the databases are managed in the central database. The contents of the different databases is different. In the central database the user has an overview of all the dataserie in all the decentral databases. If necessary the measurements can be copied from one database into another. The copied measurements are marked as a copy.



The DONAR-datamodel is an unique datamodel. This datamodel gives a very high performance and a relatively small size of datastorage. The metadata are stored in the about 30 What, Who/Why, Where and How tables. These metadata are joined into the W3H table. Below the W3H relation table the datamodel splits up into 7 different tables for different kind of dataseries. There are 2 time-related series tables (equidistant and non-equidistant); 4 location-related series tables (for track and grid series, both equidistant and non-equidistant); and 1 specimen-related series table. At the bottom of the datamodel there are about 10,000 data tables, with the about  $1,4 \cdot 10^9$  measurements over all the databases.



Data-exchange between the different DONAR-databases, the DONAR-applications and other systems has been done by a standardized file(-format).

### 3 Data Management in the Netherlands

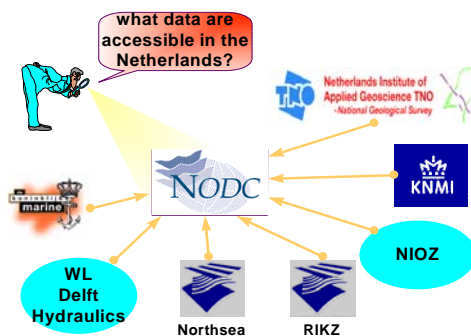
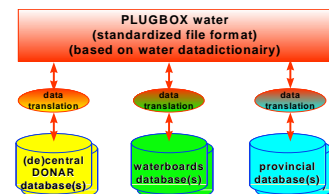
Just like the Rijkswaterstaat all the other institutes have taken different measures for delivering a higher quality of data. In cooperation with other institutes (national and international) we started developments for the exchange of data. Some of these developments are:

- National
  - ◆ Standard water datadictionary for 'water' information systems for Rijkswaterstaat, Waterboards and Provinces
    - Standard set of metadata (based on the DONAR-metadata)
      - one set of standardized "what" - codes;
      - one set of standardized "how"- codes;
      - one set of standardized codes for biological species.
    - Translation standard set of metadata to metadata from different information systems.

- ◆ 'Plugbox'
  - standardized data and file format based on the water datadictionary.
- ◆ Start National Oceanographic Data Committee (NODC)
  - Exchange of national oceanographic data.



water



- International
  - ◆ DEMOWAD
    - In cooperation with the Danish and German institutes development of a database conglomeration for the exchange of trilateral Waddensea data.
  - ◆ REMSSBOT
    - In cooperation with Greek, Italic and Belgium institutes development of a advanced coupling between the several national databases for the exchange of data;
    - Based on the CDS of the European Environmental Agency (EEA).

# IEO Marine Pelagic Ecosystem Studies and Data Management System

Luis Valdés

Instituto Español de Oceanografía, Centro Oceanográfico de Santander, P.O. Box 240, 39080 Santander, Spain [tel: (+34)42 27 50 62; fax: (+34)42 27 50 72; E-mail: luis.valdes@st.ieo.es].

## 1. Introduction

The Marine Environment Department of the Instituto Español de Oceanografía (IEO) is conducting several research projects based on the systematic and continuous study of the Ocean. The principal goal of the core project (titled "Studies on time series of oceanographic data") is to understand the underlying causes of temporal variability of the physical and biological properties and processes in the pelagic ecosystem in the neritic and oceanic waters surrounding the Spanish coast. The research effort involves 1) time series measurements in several transects along the Spanish coast, on both the Atlantic and the Mediterranean sides, 2) large-scale measurements of physical and biological properties of the marine pelagic ecosystem from ichthyoplanktonic, acoustic, bottom trawls surveys, etc, the spatial resolution in this case being more important than the temporal resolution, and 3) retrospective analysis of existing data sets from landings of pelagic fisheries, sea level measurements, climatological data, etc.

In response to the need to deal effectively with the large and varied volumes of data that have been accumulated as a result of the above activities 1 and 2, the IEO has developed a relational database that acts as an archive of the data and as a tool for data analysis and elaboration of reports.

## 2. Data Management System

The IEO Pelagic Ecology Data Base is an ORACLE relational database. The database is accessed from PC's running Windows NT. It is prepared to store and manage data of Phytoplankton, Microzooplankton, Mesozooplankton, Ichthyoplankton and environmental data. Various routines are available for coding and entering data via keyboard or directly from data files (e.g. for CTD data). The database includes more than 30 Master-tables that codify the type of data to be introduced but also give help to the user and minimise the effort of introducing the data. Data can be extracted from defined areas, periods or selected entities (surveys, species, etc), and all the results can be easily exported to EXCEL with provision for users without prior experience of the ORACLE system. A collection of selected outputs is included in the package.

The structure is divided into several levels which constitute a hierarchy and it is the major axis of the database. The data conforming to the hierarchy described in Table I is pyramidal in structure, thus, for a given cruise only a single input is required at level 1. Each cruise has a number of stations associated. At each station several samples with different devices and purposes are recorded, and at each sample a different assemblage of species are counted, this level being the base of the pyramid and by far the most numerous in number of records.

Table I. Summary of the data hierarchy

Level	Code	Contents
1	BaseCamp	Information about the organization of a cruise
2	BaseEst	Information about an individual station
3	Mues*. *	Details about a Phyto, Micro, Mesozoo, etc sample taken at a given station
4	Rec*. *	Details of the assemblage of Phyto, Micro, Mesozoo, etc. in a given sample, detailed by species and numbers

## 3.DataEntry

All information associated with cruises, stations and sampling procedures is entered from specially designed forms (Data Reporting Formats), which were defined prior to the developing of the database. In order to avoid the risk of errors when typing data at some critical records, the system derives the user to a master table when some input is unknown. If the input is really new, then it is added to the stored list. More than 25 master tables are appended at these levels of hierarchy, including: Vessel names, Harbours, Chief scientist, Sampling Area, Periodicity of sampling, another four tables to identify the cruise and sampling station; Type of device, Mesh microns, Sampling preservation, Kind of tow, another five tables to identify the sampling details; Douglas scale, Beaufort scale, another six climatological and environmental scales to define the conditions of the sampling. At level 1, SST images, depth contour maps, station maps, etc., can be added as PIF, TIF, as well as other format files.

The assemblages of species counted at each sample and their numbers are entered by means of a multi-line stream. A three to six letters code is required to enter the species and taxa (*Acartia clausi* is ACACLA; Copepoda is COP, and so on), this code is checked against the species Master-list, where every single species or taxa have been assigned an NODC code, the complete Linnaean system name, ecological information (e.g. cosmopolitan, neritic, value as indicator species, etc.), biological (e.g. holoplankton, meroplankton, etc), physiological (Developmental stage, Weight in Carbon, etc.) and other data of interest.

If a wrong species is typed or it is a new species not included in the master list, then a pick-list is displayed on the screen to help the user to find the correct name, or to add and store the new species in the database.

Once that the complete assemblage of a given sample has been entered in the database, several crossed sums are computed (e.g. Holoplankton, Meroplankton, Copepods, information that the database obtains directly from the master list) to verify that the data have been typed correctly. Thus, at this level of hierarchy more than 5 types of data are codified in master tables.

#### 4. Logistic considerations:

As the data entry process is critical to the accuracy and proper use of the stored data, some logistic considerations are under continuous revision to improve the database.

Data reporting formats: With the objective of unifying the type of information associated with cruises, stations and marine and climatological conditions, a Data Reporting Format was created, which must be used by those scientists using the database. This form is independent of the discipline, and the main purpose is to give the essential space-time information to allow every observation to be assigned and located in a specific project and cruise. With respect to the cruise this form includes the following items: Project, Cruise identifier, Vessel, Start of Cruise ddmmyyyy, End of cruise, Sampling area, Number of sampling stations, Operator (Chief Scientist). With respect to the station the form includes among others the following items:

Project, Cruise identifier, Station number, Date, Station Start Longitude, Station Start Latitude, Station Start Depth, Station Start Time, Station End Time, Station End Longitude, Station End Latitude, Station End Depth, Number of sampling devices, 14 different records of sea and climatological conditions, and Operator.

With respect to the sampling procedures, the report format is different depending on the discipline, but basically they contain the information necessary to: i) connect the sampling with a specific station and cruise, and ii) define the quality of sampling, e.g. sampling gear, type of tow, speed of the tow, depth of sampling, volume of water filtered, existence of clogging, etc (i.e. metadata).

Data quality control and Metadata: In the case of hydrographical parameters, an indication on quality control can be added to the database if calibration has been applied to the data, this is to ensure that the data are of high quality, and to make sure that all data are processed in a similar manner. However, this is not completely solved in the case of biological data, and the best approach to a data quality control is that obtained by means of the metadata (data that inform about the data).

Thus, our database contains two types of records, those that must necessarily be filled (they allow us to

connect different levels of the database, and are the only data necessary to operate the database), and those that can be left blank (but they don't). Most of the latter refer to the data that inform about the data (Metadata).

The problem that arise with the indicator is that, according to the situation it may be necessary to fill more than 20 items to inform properly about a given measure (e.g. ICC, 1996, suggests 23 questions to inform about the methodological details for Primary Production measurement). We have tried to find a meeting point between metadata requirements and the operative procedures of the scientist when working at sea, thus for primary production we have reduced the metadata to 12 questions.

In recognition that there is an essential need to provide metadata, at every level of the hierarchy, the database has been provided with a series of metadata requirements that make the data useful for future users of the data and when submitting data to a data centre.

Taxonomic code: Capacity of banking and handling data is not a real limit in terms of computing time and space, and the interest of using a numerical code to classify the species is not based on these objections but on the interest of the hierarchy of a numeric taxonomic code. For example, NODC v.7 code for Acartia clausi is 6118290101 which means that Acartia clausi belongs to the Family Acartiidae, Order Calanoida, Subclass Copepoda etc., which allows us to operate when summing up all the species that share one given character, e.g. those that are Copepoda, or Decapoda, or Clupeida, etc.

The election of a single taxonomic code for the ICES area was extensively discussed during 1997 in several Working Groups, and a general agreement was reached about the NODC code on the basis that i) this taxonomic code contains logical hierarchical taxonomic information, ii) it is not limited to pelagic species, iii) it is flexible, giving the opportunity to incorporate new codes, and iv) it is well established (WGZE, 1997). (Subsequent to the 1997 Working Group meetings, a new version of NODC code has been released, which differs significantly from v.7).

The WO Data Management System has incorporated the NODC taxonomic code in its version 7.0 as the standard to operate in all the planktonic communities managed in the database.

## 5. Data Outputs: Reports, ASCII Tables, EXCEL Tables, and Figures

The database not only acts as an archive of the data, but also as a tool for data analysis and elaboration of reports. Data can be extracted from defined areas, periods or selected entities (surveys, species, etc), and all the results can be easily exported to EXCEL and ASCII tables with provision for users without prior experience of the ORACLE system.

In addition, a collection of selected outputs is included in the package, which allows the user to obtain Reports, Tables and Figures. For example it is possible to select, from a pre-defined menu, the sampling area, dates, depth and parameters to be plotted in a graph or to be transferred to a Table, or produce the list of species for a particular survey, etc. Options allow the user to choose any combination among more than 20 abiotic and biotic parameters (Depth, Temperature, Salinity, Nutrients, PAR, Fluorescence, Chlorophyll, Biomass, Abundance, etc.).

Also two different types of cruise reports are pre-defined in the outputs menu and a dialogue menu is being developed to edit and feed the ROSCOP format directly from the database.