Vol. 54 – SUPPLEMENT

Bollettino di Geofisica
teorica ed applicata

An International Journal of Earth Sciences

IMDIS 2013
International Conference on Marine Data and Information Systems

23-25 September, 2013 - Lucca (Italy)

Organised by
ENEA, jointly with
EU SeaDataNet project, IFREMER and IOC/IODE.

Organizing committee
Giuseppe Manzella, chair - ENEA, Italy; Michele Fichaut - Ifremer, France
Dick Schaap - MARIS, The Netherlands; Lesley Rickards - NERC, UK
Vanessa Tosello - Ifremer, France

International Steering Committee
Giuseppe Manzella, chair - Italy; Steve Miller - USA
Robert Keeley - Canada; Margarita Greggs - USA
Roger Proctor - Australia; Gilbert Maudire - France
Michele Fichaut - France; Dick Schaap - The Netherlands
Lesley Rickards - UK; Peter Pissierssens, IODE
Serge Scory - Belgium; Jean-Marie Beckers - Belgium
Reiner Schlitzer - Germany; Friedrich Nast - Germany
Athanasia Iona - Greece; Alessandra Giorgetti - Italy
Simona Simoncelli - Italy; Vittorio Barale - EU

Istituto Nazionale di Oceanografia
e di Geofisica Sperimentale

ISSN 0006-6729

TABLE OF CONTENTS

Session 1
Marine information and data management 13

Oral presentations

M. Fichaut, D. Schaap and G. Maudire
SeaDataNet - Second phase achievements and technical highlights 15

A. Leadbetter, R. Arko, C. Chandler and R. Lowry
Linked Data: An Oceanographic Perspective 18

H. Glaves, S. Miller, R. Proctor and D. Schaap
Ocean Data Interoperability Platform (ODIP): developing a common framework for marine data management on a global scale 21

C. Borremans
Management of events, samples, videos and data generated by deep-sea submersibles 22

S. Simoncelli, C. Coatanaon, Ö. Bäck, H. Sagen, S. Scory, D. Tezcan, D.M.A. Shaap, R. Schlitzer, S. Iona, M. Fichaut and M. Tonani
Temperature and salinity historical data collections for the european marginal seas: aggregation and quality assessment procedures 23

S. Pouliquen, T. Carval, D. Guillotin, C. Coatanoan, T. Loubrieu, A. Grouazel, K. Von Schuckmann, H. Wedhe, L. Sjur Ringheim, T. Hammarklint, A. Hartman, K. Soetje, T. Gies, M. De Alfonso, L. Perivoliotis, D. Kassis, A. Chalkiopoulos, V. Marinova, P. Jaccard, A.B. Ledang, K. Sorensen, G. Notarstefano, J. Tintore, Seppo.Kaitala, P. Roiha, L. Rickards and G. Manzella
MyOcean In-Situ Thematic Center: A service for operational Oceanography 25

L. Vandepitte, F. Waumans, L. Tyberghein, B. Vanhoorne and F. Hernandez
Marine biogeographic data in EurOBIS: assessing their quality, completeness and fitness for use 26

P.C. Brauer and W. Hasselbring
PubFlow: a scientific data publication framework for marine science 29

E. Boldrini, S. Nativi and D.M.A. Schaap
INSPIRE compliant international standards for the SeaDataNet marine metadata 32

R. Proctor, K. Finney and S. Mancini
Vocabulary enhancements for the Australian Ocean Data Network 34

J. Hodge and N. Car
From silos to systems: using standardised services and Linked Data methodologies to augment government and research information systemse 36

S. Carbotte, R. Arko, V. Ferrini, K. Stocks, S. Miller, A. Maffei, C. Chandler, S. Smith and M. Bourassa
Rolling Deck to Repository (R2R): Working with International Partners to Improve Data Discovery and Access 38

S.J. Worley, S.D. Woodruff, J.E. Freeman and S.R. Smith
ICOADS the Global Marine Surface Reference Dataset 40

K. Deneudt, H. M. Jensen, B. Vanhoorne, L. Vandepitte, S. Claus and F. Hernandez
Making SeaDataNet more fit for handling biological data 42

T. Loubrieu, V. Claverie, R. de Dianous, T. Jolibois, L. Romero, J. Sala, O. Cases, A. Bon, M. Treguer, J. Blower and G. Griffiths
MyOcean Central Information System 44

M. Tonani, S. Simoncelli, A. Grandi and N. Pinardi
New gridded climatologies, from in-situ observations, for the Mediterranean Sea 47

Posters

R. Lowry, A. Leadbetter and M. Vinci
Semantically-enhanced Aggregation of SeaDataNet Data 49

A. Kozyr
CDIAC Data Management Support for Ocean Carbon Dioxide Measurements 52

R. Keeley and R. Schlitzer
Modular Information Content for Ocean Data Systems 55

M. Pagnani, R. Huber and R. Lampitt
FixO^3: The next phase of Open Ocean observatory Data Management Harmonisation 56

T. Jolibois, V. Claverie, T. Loubrieu, L. Romero, J. Blower, R. De Dianous and G. Griffiths
Data access services in MyOcean: needs, technical challenges and implementation 58

D. Cotton, T. Allen and L. Rickards
MEDIN - A UK Partnership for Managing Marine Data and Information 60

J. Olive, A. Vetrano, M. Borghini, J. Sorribas and M. Nokin
CTD calibration description using SensorML files 63

C. Pinto and N. Holdsworth
ICES Data portal, a single entry point to marine knowledge 66

G. Coro, A. Gioia, P. Pagano and L. Candela
A Service for Statistical Analysis of Marine Data in a Distributed e-Infrastructure 68

S. Iona, S. Balopoulou, P. Karagevrekis and A. Lykiardopoulos
The HNODC Data & Information Management Services: Description & Recent Upgrades 71

C.M. McCann and J. Harvey
STOQS: The Spatial Temporal Oceanographic Query System 73

J. Gatti, J. Arroqui, S. Trigo, A. Baldoni, G. Molinari and R. Guerrero
ALBATROSS: Retrieval, processing and visualization of INIDEP's oceanographic data and products 76

M. Moreno-Báez, C. López-Sagastegui, J. Esquivel and O. Aburto-Oropeza
A participatory framework to collect data for understanding the linkages of coastal and marine resources and human use 79

G. Reed
A Quality Management Approach to the Delivery of Marine Data and Services 81

D. Ivankovic, I. Vucic and V. Dadic
Using GeoServer and spatial database for receiving and analyzing of georeferenced data 82

R. Kottmann, A. Pop Ristov, J. Schnetzer, A. Fernandez-Guerra and F. Oliver Glöckner
Bridging the gap between marine environmental and molecular data: The Micro B3 Information System 84

A. Pannocchi, L. Arranz, A. Hernández, E. Tel and A. Viloria
Sub-surface Underway Data Validation at IEO Data Center 87

A. Behnken, R. Huber, C. Waldmann, M. Kononets, A. Tengberg, P. Hall and H. Stahl
Using OGC standards (SOS, O&M) to standardize data flow from cabled marine observatories to the PANGAEA data archive 90

A. Lykiardopoulos, K. Kalkavouras, F. Gheorghe and A. Iona
Monitoring services in the framework of SeaDataNet II Project 92

L. Rickards, J.-B. Minster and M. Mokrane
Progress in developing the ICSU World Data System 94

V. Grande, F. Foglini, F. Lisi, F. De Leo, A. Chassanite, O. Neprokin and R. Lisovsky
COCONET PROJECT (Towards COast to COast NETworks of Marine Protected Areas) - the architecture of a common Geodatabase for Marine Data Management and synthesis 96

Session 2
Marine environmental data bases: infrastructures and data access systems 99

Oral presentations

H. Parner
ICES Report on Ocean Climate (IROC) - Online 101

S. Belov, N. Mikhailov and K. Belova
Ocean Data Portal: from data access to integration platform 103

D. Castelli, P. Pagano, L. Candela and G. Coro
The iMarine Data Bonanza: Improving Data Discovery and Management through an Hybrid Data Infrastructure 105

P. Diviacco, A. Busato, M.-P. Corre, K. De Cawer, M. Nokin, S. Scory, J.-M. Sinquin, J. Sorribas and Y. Stojanov
Bridging the gap between data and metadata (Part 2): looking backward at what happened during a survey with the Eurofleets Automatic Reporting System (EARS) 108

A. Che-Bohnenstengel and F. Nast
Modernised Cruise Summary Reports Management, in particular with a Link to Data 111

V. Forneris, C. Tronconi, G. Volpe, S. Colella, B. B. Nardelli, A. Pisano and R. Santoleri
The CNR-ISAC Informatics Infrastructure for the Satellite Climatological and Oceanographic data: production, harmonization and dissemination in Interoperability Frameworks 113

S. Tarot, S. Guibert, Y. Kerr and C. Larigauderie
Data management at the CATDS-PDC, the operational processing center for high level data of the SMOS satellite 115

H. Glaves and D. Schaap
Geo-Seas: delivering harmonised marine geoscience data on a European scale 117

A. Novellino, G. Manzella, D. Schaap, P. Gorringe, L. Rickards and S. Pouliquen
EMODNet Physical Parameters.. 118

M. Vinci, A. Giorgetti and A. Brosich
New EU efforts to assess the state of the marine environment: the EMODnet Chemistry pilot project.. 121

M. Moreno-Báez, V. Obregón-Noriega V., A. Sánchez-Rodríguez and O. Aburto-Oropeza
DataMARES: Marine Access Repository for the Gulf of California, Mexico.................................... 123

D. March, S. Lora, J. P. Beltrán, K. Sebastian, S. Gómara and J. Tintoré
Data Management and applications at SOCIB 125

T. Suzuki, M. Ishii, T. Nakano, H. Yoritaka, A. Seta, S. Kizu, K. Sato, Y. Shimizu, K.-Amaike, T. Suga, and Y. Michida
XBT Data Management and Quality Control in Japan 127

J. Melles and R. Lehfeldt
The German Marine Data Infrastructure (MDI-DE).. 129

M. Wichorowski and S. Sagan
Integrated Ocean Data and Information Management System Oceanographic Data Center of IO PAN 131

L. Petit de la Villéon, T. Delcroix, G. Alory, F. Gaillard and M. Krieger
GOSUD: Global Ocean Surface Underway Data From gathering underway data to providing high quality delayed mode data 134

I. Gertman, Y. Krivenko, T. Ozer, V. Belokopytov, A. Ingerov and A. Khaliulin
Regional project (PERSEUS) oriented system for storage and operative exploration of Mediterranean and Black Sea cast data 136

D. Tezcan, V. Myroshnychenko, E. Yılmaz, B. Salihoglu, Z. Uysal and A.E. Kıdeys
Monitoring Long Term Variability in the Eastern Mediterranean 138

Posters

G. Breitbach and H. Krasemann
A new Approach in Data Management for an Integrated Coastal Observation System................ 141

C. Coatanoan, C. Cabanes, A. Grouzael and V. Turpin
CORA: in-situ re-qualified dataset at the Coriolis Data Center... 143

D.F. Morales, J.E. Fuentes, J. Bohorquez and D.M. Rozo
An Information System for Monitoring Coastal Erosion at Colombia - COSTERO........................... 144

O. Lavrova, E. Loupian, M. Mityagina, I. Uvarova and T. Bocharova
See the Sea - Multi-User Information System Ocean Processes Investigations Based on Satellite Remote.. 146

P. Diviacco, A. Busato, H. Glaves and D. Schaap
Bridging the gap between data and metadata (Part 1): a novel paradigm and technologies for data discovery, evaluation and access, emerging from addressing the issues of seismic data within the Geo-Seas project 148

E. Moussat, B. Loubrieu and T. Schmitt
Harmonization of Digital Terrain Model production in EU distributed infrastructures 151

I. Zananiri, D. Mitropoulos, V. Zimianitis, C. Ioakim, V. Papadopoulos and G. Efthimiou
Marine Geology data accessibility in the european framework: the I.G.M.E. Database and access services after its participation in the geo-seas project 153

T. Alcorn and E. O'Grady
Opening up Irish marine data 156

G. Hendriksen, G. van der Kolff and K. Borst
Towards Open Monitoring Data within marine monitoring projects 157

A. Stefanov and A. Palazov
Querying structured and unstructured data - combined approach 159

T. Gies, K.-C. Soetje and M. Luecker
BSH contribution to the Copernicus Marine Core Service "MyOcean": Northwest-Shelf in-situ observations 161

S. Almeida, R. Baptista, I. Golde, A. Lopes and A. Saramago
A contribution to the Marine Observatory network 163

D. Dezeljin and B. Cermelj
Consolidation of marine datasets at Marine Biology Station, NODC for Slovenia 165

V. Solovyev, A. Ostrovskii and A. Zatsepin
The Aqualog profiler is measuring platform for an ocean observatory 167

P. Kulesza and M. Wójcik
Open Source Software as an Efficient Distributed Infrastructure Stack for GIS Data Processing and Dissemination 169

R. Koppe and A. Macario
MaNIDA: An integrated national approach to scientific information management and innovative use of information assets' relationships 172

K. Malde, T. Furmanek and E.M. Hassen
A data storage for generic and heterogenous scientific data 174

G. Silvoni and P.A. Azrilevich
National Sea Data System: The Argentine strategy 176

B. Meola and C. Webster
MAPAMED, the database on Marine Protected Areas in the Mediterranean Sea 177

A. Spinu, L. Buga, R. Mateescu, L. Alexandrov, and M. Emanuela
Integration of oceanographic databases into the Geographic Information System (GIS) as support for Marine Spatial Planning. Case of Romanian Littoral 179

H. Glaves and A. Stevenson
EMODNET-Geology: assembling and disseminating marine data for European seabed mapping .. 180

T. de Bruin, R. de Koster and P. Thijsse
The Wadden Sea Long Term Ecosystem Research (WaLTER) project: Using the SeaDataNet infrastructure to provide access to integrated environmental and socio-economic data from the Dutch Wadden Sea 181

Y. Sagarminaga, I. Galparsoro, I. Orbe and J.A. Sánchez
Development of ITSASGIS-5D: seeking interoperability between Marine GIS layers and scientific multidimensional data using open source tools and OGC services for multidisciplinary research 182

D. Allison, R. Groman, C. Chandler, S. Rauch, D. Kinkade, P. Wiebe, D. Glover and C. Galvarino
Are the Data Really Available? 184

K. Tikka, G. Hällfors, S. Hällfors, M. Huttunen, R. Olsonen and M. Virkkunen
Management of taxonomic species and observational data - case Baltic Sea phytoplankton 186

P. Carrara, A. Sarretta, A. Giorgetti, M. Ribera D'Alcalà, A. Oggioni and E. Partescano
An interoperable infrastructure for the Italian Marine Research 188

V. Dadic, D. Ivankovic, J. Bačic and I. Vucic
Information system of marine environment indicators in Croatia 191

S. Diggs and B. Sloyan
GO-SHIP: building a time series of a suite of ocean properties 193

Session 3
Data Services in ocean science 195

Oral presentations

C.L. Chandler, S. Konovalov, H. Garcia, R. Schlitzer, L. Devine, G. Moncoiffé, T. Suzuki, A. Kozyr and G. Reed
IODE Data Quality Flag Scheme 197

S. Claus, F. Souza Dias, B. Vanhoorne, N. De Hauwere and F. Hernandez
Marine Regions: towards a standard for georeferenced marine names 200

W. Appeltans, P. Pissierssens, G. Coro, A. Italiano, P. Pagano, A. Ellenbroek and T. Webb
Trendylyzer: a Long-Term Trend Analysis on Biogeographic Data 203

A. Barth, C. Troupin, A. Alvera-Azcárate and J.-M. Beckers
Variational data analysis for generating ocean climatologies (DIVA) and web-based distribution of data products (OceanBrowser) 206

R. Schlitzer, M. Menzel and S. Heckendorff
A new API for accessing ODV data collections from C++ and Java 208

J.-M. Sinquin, J. Sorribas, P. Diviacco, M. Baeye, G. Quemener and B. Loubrieu
GLOBE 3D: an new software for Geosciences data processing and 3D/4D viewing 209

C. Satra Le Bris, E. Quimbert, M. Treguer and A. Louarit
Sextant: the French Spatial Data Infrastructure for Marine Environment 211

C. Chandler, R. Groman, D. Allison, P. Wiebe and D. Glover
One Approach to Ensuring Long-term Access to Ocean Research Data 213

G. Santinelli, F. Baart, G. Ramaekers and E.-J. Vos
Coastviewer: a tool to enable the visualization of marine and coastal data 216

L. Crosnier, J. Dorandeu, R. De Dianous, V. Claverie and T. Loubrieu
The MyOcean Product Database Management 219

S. Stylianou, G. Zodiatis, G. Galanis and A. Nikolaidis
Implementation of a live access server for the management of wave energy data in the Eastern Mediterranean 221

M. Tampucci, M. Martinelli, D. Moroni, G. Pieri, O. Salvetti and P. Villa
A Proactive System for Oil Spills and Marine Environment Monitoring 222

L. Adams, D. Lear, B. Seeley and C. Griffiths
Marine Social and Economic Data - UK Perspective 224

Posters

E. Vyazilov, D. Melnikov and N. Chunyaev
Analytical tools for ESIMO 227

N. Viazilova
Storm activity over North Atlantic with ESIMO analytical tools 229

M. Sinerchia, A. Cucco, L. Fazioli, A. Olita, A. Ribotti and R. Sorgente
A numerical based DSS as tool for local planning of Oil spill emergencies 230

L. Raymond, C. Chandler, A. Leadbetter, E. Urban, P. Pissierssens and L. Pikula
Publishing and Citing Ocean Data 233

S.-D. Kim, H.-M. Park, J.-H. Lee, S.-H. Baek and S.-Y. Park
Establishment of an Operational Data Management System for Korea Ocean Prediction System 235

E. Zhuk, G. Zodiatis, S. Stylianou and R. Lardner
A web-based GIS for oil spill prediction 238

G. Hendriksen, S. Tatman, W. Stolte and F. Baart
Enabling connectivity between data sources within the VECTORS project 239

A. Alyautdinov, I. Kalinkin and V. Semin
"Up-Grade BS-Scene" and "CaspInfo" projects as examples of implementing standards provided by SeaDataNet and supported by INSPIRE 243

M. Osborne
The Application of Marine Data Management in Ports using Ocean Database and GIS 245

J. Madelaine, G. Domalain, W. Marin Soto
ThemaMap: from data to maps 248

P. Weatherall and R. Lowry
General Bathymetric Chart of the Oceans (GEBCO) - Global Bathymetric Data Sets 251

K. De Cauwer, Y. Stojanov, J. Sorribas, P. Diviacco, M.-P. Corre, A. Busato, S. Scory, J.-M. Sinquin and M. Nokin
Semantic interoperability on board: ongoing developments for EARS event logging 253

J. Foden, D. Mills, K. Keeble and S. Keeble
The EMECO Datatool: an online assessment and reporting system for co-production of environmental assessments of shared marine waters 255

P. Springer, H. Mehrtens, D. Fleischer, C. Schirnick and A. Czerniak
Kiel Data Management Infrastructure for Marine Sciences 257

P. Alenius, K. Tikka and R. Hietala
Near real-time quality data from ships to land 259

T. Kõuts, M.-J. Lilover and K. Vahter
On-line ice resistance survey for the ships supporting winter navigation in ice channels of the Baltic Sea 261

C. Troupin, A. Barth, J.-M. Beckers and A. Pascual
Application of the Data-Interpolating Variational Analysis (DIVA) to sea-level anomaly measurements in the Mediterranean Sea 263

E. Partescano, A. Brosich and A. Giorgetti
A system for managing metadata using XML format 265

A. Vetrano, A. Bechini, C. Chiappini and K. Schroeder
Browsing and Selection of CTD/XBT Data through Open-Source GIS Clients 267

E. Godin, A. Khaliulin, A. Ingerov and E. Zhurk
On-line quality control service of MHI NASU 270

M. Pantazi, V. Vassilopoulou, Y. Issaris, A. Kokkali, M. Salomidi, A. Frantzis, P. Cetacean, J. Fric and A. Panou
Addressing uncertainty through data quality issues in environmental assessments to support improved decision making 271

M. Picone, G. Nardone and F. Lagona
Data validation in italian wave metric datasets 274

C. Chandler, A. Maffei, L. Stolp, C. Sellers, E. Coburn, T. Thiel and R. Arko
R2R Science Eventlogger: Supporting Data Interoperability with Controlled Vocabularies 276

Session 4
Services for Users and Education 279

Oral presentations

B. Pfeil, A. Olsen, D. Bakker, S. Hankin, K. O'Brien, K. Smith, A. Kozyr, C. Sabine, M. Telszewski and M. Diepenbroek
Surface Ocean CO2 Atlas (SOCAT) Version 2 - a showcase for transparent data management and international collaboration 281

V. Barale
European Atlas of the Seas: Marine & Maritime Information for Science and Society 283

L. Rickards, T. de Bruin, N. Holdsworth, H. Parner, S. Ringheim Lid, G. Dawson
The ICES Data Type Guidelines and their place and role in the profusion of guidelines, manuals, standards, cookbooks and best practice ... 284

G. Matishov, S. Berdnikov, V. Kulygin and N. Yaitskaya
Geoinformation Systems, Databases, and Climatic Atlases of the Sea of Azov, Black Sea, and the Caspian Sea ... 286

Posters

A.-M. Eversley
Caribbean Marine Atlas ... 289

L. Rickards, E. Bradshaw and J. Andrew
Rescue of historical UK sea level charts and ledgers - creation of Open Educational Resources (OERs) ... 291

G. Manzella and OTTIMA
Collaborative Knowledge Building ... 293

C. Delgado and P. Pissierssens
OceanTeacher goes global: The OceanTeacher Global Classroom ... 295

D. Bayley and A. Ridgeway
Joint Nature Conservation Committee Marine Protected Areas Interactive Online Mapper ... 298

R. Lagring
Back to the sea of the past to evaluate the status of today: Integration and valorisation of historical data ... 299

B. Karlson, Ö. Bäck, L.J. Hansson, A. Andreasson, M. Johansen, M. Mohlin, A.-T. Skjevik, A. Loo and L. Fyrberg
Nordic Microalgae - an information system for aquatic microalgae ... 301

K. Shein, J. Stachniewicz, L. DeWayne Cecil and S. Ansari
A Climate Data Access and Visualization Tool for Marine Ecosystem Managers ... 302

R. Proctor, P. Oke and U. Rosebrock
Australia's marine virtual laboratory ... 305

E. Zhuk, A. Khaliulin, A. Ingerov and E. Godin
Web interface for the Black Sea database with GIS functionality ... 307

K. Gordon, A. Matthews and L. Rickards
Celebrating 80 years of the Permanent Service for Mean Sea Level (PSMSL) ... 309

M. Belhassen, N. Barraj, S. Lahbib and S. Messaoudi
Tunisian Coastal Marine Atlas ... 310

D. Hayes, G. Zodiatis, D. Schaap, V. Vassiliou, C. Vrioni, C. Coman, A. Karaolia and I. Constantinou
SeaMobile apps - marine data and forecast services available on mobile devices ... 311

SESSION 1

Marine information and data management

- Exchange, processing and interactive work with marine data sets from highly heterogeneous sources
- Federation and integration
- Network services and technologies

ORAL PRESENTATIONS

SeaDataNet – Second phase achievements and technical highlights

Michèle Fichaut, IFREMER, michele.fichaut@ifremer.fr (France)
Dick Schaap, MARIS, dick@maris.nl (Netherlands)
Gilbert Maudire, IFREMER, gilbert.maudire@ifremer.fr (France)
and the SeaDataNet consortium

SeaDataNet is the leading network in Europe actively operating and developing a pan-European infrastructure for managing, indexing and providing access to ocean and marine metadata, data and products originating from data acquisition activities by all engaged coastal states. SeaDataNet aims at serving users from science, environmental management, policy making and economic sectors. The second phase of the project SeaDataNet started on October 2011 for a 4 years' duration; it aims to upgrade and improve the SeaDataNet infrastructure built during the previous phase of the project (FP6, 2006-2011) on the foundations established during the Sea-Search FP5 project (2002-2005).

Objectives of the Second Phase of SeaDataNet

At the end of the successful first phase of SeaDataNet the distributed data system connecting more than 50 data providers distributing 850,000 datasets was operational, common data management procedures and tools, common data policies, common data license and common vocabularies and standards were adopted by all the connected data centres. Nevertheless many improvements were still to implement in order to move towards an operationally robust and state of the art Pan-European and the main objective of SeaDataNet II, that will be detailed in the presentation, is to improve operations and to progress towards an efficient data management infrastructure able to handle the diversity and large volume of data collected via the Pan-European oceanographic fleet and the new observation systems, both in real-time and delayed mode.

Achievements at mid-term of the project

Since the beginning of SeaDataNet phase II, a lot of progress has been made in order to become 'bigger and better':

- monitoring of the infrastructure has been developed in order to improve its robustness and availability for users;
- compliancy to INSPIRE has been implemented by moving the most visible and used SeaDataNet metadata catalogues to ISO-19139 and by modifying all tools related to these catalogues;
- SeaDataNet NetCDF format definition and machine-to-machine interface definition and first tests to be able to serve operational oceanography an ocean atmosphere communities;
- Ocean Data View (ODV) format has been adapted to be able to manage biological data exchange;

- IOC-IODE Ocean Data Portal is now linked to SeaDataNet portal and SeaDataNet can be searched through ODP;
- improvement of the quality of the delivered data (duplicate checks, format checks, quality check loops in cooperation with the regional coordinators in strong relationships with MyOcean intu-tac and regional GOOS organisations);
- More data and metadata are available in all SeaDataNet catalogues and more data providers are connected to the infrastructure.

Technical activities

The technical work plan for SeaDataNet II is organised as a cycle of activities that pass from operation to development to operation. A very important aspect is that new services, components and standards must be implemented over the whole network and without causing disturbances in the operational functioning of the infrastructure. This is achieved by versioning of services, parallel installation and testing before moving to production, and careful coordination of upgrades implementation.

An important service in SeaDataNet is the Common Data Index (CDI) data discovery & access service. This gives users a highly detailed insight in the availability and geographical spreading of a large variety of marine and ocean data sets, that are managed by data centres, that are connected to the SeaDataNet infrastructure. Moreover it provides a unique interface for requesting access, and if granted, for downloading data sets from these distributed data centres across Europe.

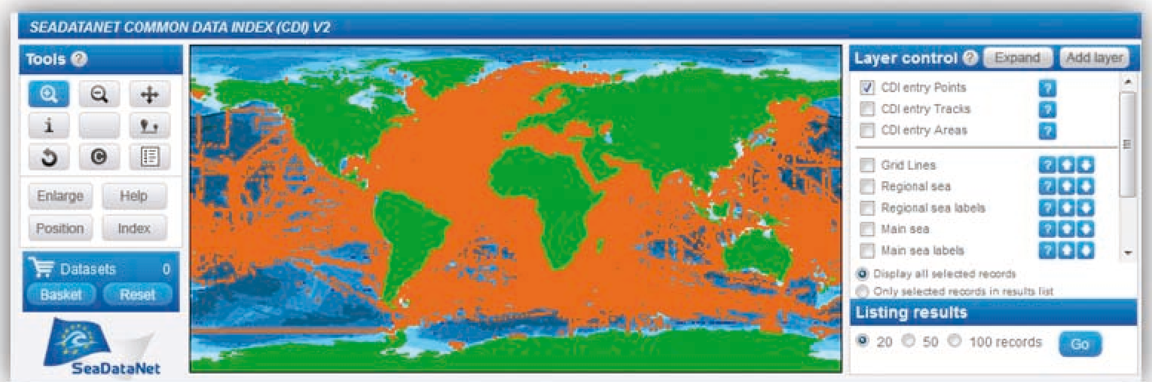


Fig. 1 - CDI content – location map of the vertical profiles.

The CDI service gives access to a vast and rapidly increasing resource of marine and ocean datasets, managed by an increasing number of distributed data centres. At present it provides metadata and access to more than 1.300.000 data sets, originating from more than 400 organisations in Europe, covering physical, geological, chemical, biological and geophysical data, and acquired in European waters and global oceans. Already more than 80 data centres from 29 countries are connected. As part of SeaDataNet II and associated projects more data centres are connecting while also the volume and range of types of data sets is expanding.

Future of SeaDataNet

SeaDataNet is now one of the major players in informatics in oceanography and has created strong relationships with other EU or international projects. In particular, SeaDataNet plays major roles in the continuous serving of controlled vocabularies widely used, the provision of tools for data management as well as providing access to oceanographic data, metadata and products of importance for society.

Within the European framework of Marine Knowledge 2020, SeaDataNet must continue on standards and protocols development and must provide infrastructure for data management for marine research community within a legal framework.

SeaDataNet Consortium

44 partners:

IFREMER (France), MARIS (Netherlands), NERC-BODC (United Kingdom), BSH-DOD (Germany), SMHI (Sweden), IEO (Spain), HCMR-HNODC (Greece), OGS (Italy), RIHMI-WDC (Russian Federation), ENEA (Italy), INGV (Italy), METU-IMS (Turkey), CLS (France), AWI (Germany), ULG (Belgium), IMF (Norway), NERI (Denmark), ICES (Denmark), IES-JRC (Italy), MI (Ireland), IHPT (Portugal), NOIZ (Netherlands), RBINS-MUMM (Belgium), VLIZ (Belgium), MRI (Iceland), FMI (Finland), IMGW-PIB (Poland), MSI (Estonia), LHEI (Latvia), EPA (Lithuania), SIO-RAS (Russian Federation), MHI-DMIST (Ukraine), IO-BAS (Bulgaria), NIMRD (Romania), TSU-DNA (Georgia), IOF (Croatia), NIB (Slovenai), UOM (Malta), OC-UCY (Cyprus), IOLR (Israel), CNR (Italy), IBSS (Ukraine), UniHB (Gremany), TUBITAK (Turkey).

13 subcontractors:

CluWeb (Italy), Eu-Consult (Netherlands), IMBK (Montenegro), INEWI (Albania), INRH (Morocco), INSTM (Tunisia), IOC-IODE (Belgium), IOC-JCOMMOPS (France), RSHU (Russain Federation), SHODB (Turkey), SHOM (France), STFC (United Kingdom), UTM-CSIC (Spain).

1 associated partner:

IOPAN (Poland).

Linked Data: An Oceanographic Perspective

Adam Leadbetter, British Oceanographic Data Centre, alead@bodc.ac.uk (United Kingdom)
Robert Arko, Lamont-Doherty Earth Observatory, Columbia University, arko@ldeo.columbia.edu
Cynthia Chandler, Woods Hole Oceanographic Institution, cchandler@whoi.edu
Roy Lowry, British Oceanographic Data Centre, rkl@bodc.ac.uk

'Linked Data' is a term used in Computer Science to encapsulate a methodology for publishing data and metadata in a structured format so that links may be created and exploited between objects. Berners-Lee (2006) outlines the following four design principles of a Linked Data system:

- Use Uniform Resource Identifiers (URIs) as names for things.
- Use HyperText Transfer Protocol (HTTP) URIs so that people can look up those names.
- When someone looks up a URI, provide useful information, using the standards (Resource Description Framework [RDF] and the RDF query language [SPARQL]).
- Include links to other URIs so that they can discover more things.

In this paper we present Linked Data best practices as outlined by Berners-Lee and demonstrate advances in the application of Linked Data principles in the oceanographic domain.

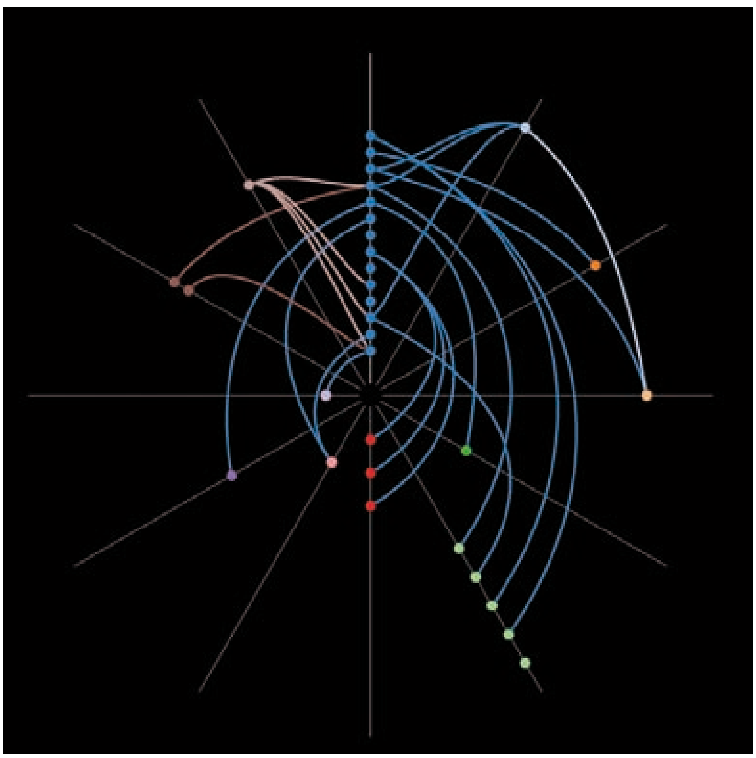


Fig. 1 - Hive plot of links between NVS2.0 (dark blue); R2R (dark brown); BCO-DMO (light brown) and other Linked Data resources.

Five-Star Linked Data

In 2010, Berners-Lee revisited his original design plan for Linked Data to encourage data owners along a path to 'good Linked Data'. This revision involved the creation of a five star rating system for Linked Data outlined below.

- One star: Available on the web (in any format).
- Two stars: Available as machine-readable structured data (e.g. An Excel spreadsheet instead of an image scan of a table).
- Three stars: As two stars plus the use of a non-proprietary format (e.g. Comma Separated Values instead of Excel).
- Four stars: As three stars plus the use of open standards from the World Wide Web Commission (W3C) (i.e. RDF and SPARQL) to identify things, so that people can point to your data and metadata.
- Five stars: All the above plus link your data to other people's data to provide context.

Five-Star Linked Oceanographic Data

Taking this ranking scheme into consideration, many oceanographic data providers will publish one-, two- or three-star Linked Data, perhaps without realising it. For example, the SeaDataNet project publishes its Common Data Index metadata online, with Uniform Resource Identifiers in an XML format, with links to the data files in non-proprietary formats. Therefore, this project already publishes three-star Linked Data.

However, there have been recent efforts in both Europe and the United States to publish connected oceanographic data resources which conform to the five-star standard. The first of these resources is the Natural Environment Research Council (NERC) Vocabulary Server (NVS2.0) – see http://vocab.nerc.ac.uk.

This resource serves lists of standard terms for populating fields in oceanographic metadata including descriptions of data, platforms, instruments and geographic locations. The design principles, described in detail by Leadbetter et al., 2013, and mapping to other semantic resources (such as the Marine Metadata Interoperability Ontology Registry and Repository) show that this is a five-star Linked Data resource using the W3C Simple Knowledge Organisation System standard which may be queried through a SPARQL endpoint.

Building on these vocabularies, both the Rolling Deck to Repository (R2R) and the Biological and Chemical Oceanography Data Management Office (BCO-DMO) projects have used terms published on NVS2.0 to mark up metadata (Arko et al., 2013). Publishing their metadata in standard RDF XML and exposing SPARQL endpoints renders them five-star Linked Data repositories.

The benefits of this approach include:

- increased interoperability between the metadata created by R2R and BCO-DMO;
- improved data discovery as users of SeaDataNet, R2R and BCO-DMO terms can find data using labels with which they are familiar;
- both standard tools and newly developed custom tools may be used to explore the data; and using standards means the custom tools are easier to develop.

Conclusions

Linked Data is a concept which has been in existence for nearly a decade, and has a simple set of formal best practices associated with it. Linked Data is increasingly being seen as a driver of the next generation of "community science" activities (McGuinness, 2012). While many data providers in the oceanographic domain may be unaware of Linked Data, they may also be providing it at one of its lower levels. Here we have shown that it is possible tio deliver the highest standard of Linked Oceanographic Data, and some of the benefits of the approach.

Acknowledgements. In the European Union this work has been funded through the Framework Programme 7 projects NETMAR (2010-3), SeaDataNet-II (2011-5) and Ocean Data Interoperability Platform (2012-5). In the United States the National Science Foundation funded projects R2R and BCO-DMO have contributed to the work. The title of this paper is borrowed from Hart and Dolbear (2013).

References

Arko, Robert, Cynthia Chandler, Paul Clark, Adam Shepherd, Carla Moore and Stace Beaulieu, 2013. *Rolling Deck to Repository (R2R): Collaborative Development of Linked Data for Oceanographic Research.* Geophysical Research Abstracts 15: EGU2013-9564

Berners-Lee, Tim, 2006. *Design Issues : Linked Data.* World Wide Web Consortium, 27th July. Revised 18th August 2009. http://www.w3.org/DesignIssues/LinkedData.html

Hart, Glen and Catherine Dolbear, 2013. *Linked Data: A Geographic Perspective.* CRC Press, 289pp. ISBN-101439869952

Leadbetter, Adam, Roy Lowry and Oliver Clements, 2013. *Putting meaning into NETMAR - the Open Service Network for Marine Environmental Data.* International Journal of Digital Earth. DOI:10.1080 /17538947.2013.781243.

McGuinness, Deborah. 2012. *Community Science – The Next Frontier.* In Proceedings of AGU Fall Meeting 2012 (December 3-7 2012, San Francisco, CA).

Ocean Data Interoperability Platform (ODIP): developing a common framework for marine data management on a global scale

Helen Glaves, British Geological Survey, hmg@bgs.ac.uk (UK)
Stephen Miller, Scripps Institution of Oceanography, spmiller@ucsd.edu (USA)
Roger Proctor, University of Tasmania, roger.proctor@utas.edu.au (Australia)
Dick Schaap, MARIS, dick@maris.nl (Netherlands)

As marine research becomes increasingly multidisciplinary in its approach there has been a corresponding rise in the demand for large quantities of high quality interoperable data. A number of regional initiatives are already addressing this requirement through the establishment of e-infrastructures to improve the discovery and access of marine data. Projects such as Geo-Seas and SeaDataNet in Europe, Rolling Deck to Repository (R2R) in the USA and IMOS in Australia have implemented local infrastructures to facilitate the exchange of standardised marine datasets. However, so far each of these regional initiatives has been developed to address their own requirements and independently of other regions. To establish a common framework for marine data management on a global scale these is a need to develop interoperability solutions that can be implemented across these initiatives.

Through a series of workshops attended by the relevant domain specialists, the Ocean Data Interoperability Platform (ODIP) project will identify areas of commonality between these regional infrastructures and use these as the foundation for the development of a number of prototypes to demonstrate interoperability between existing e-infrastructures. These prototypes will be used to underpin the development of a common approach to the management of marine data which can be promoted to the wider marine research community with a view to expanding this framework to include other regional marine data infrastructures.

ODIP is a community lead project that is currently focussed on regional initiatives in Europe, the USA and Australia. It is supported by parallel funding from the responsible agencies from each region. The European component of ODIP includes 10 partners from 6 European countries and is funded by the EU Framework 7 programme. The US participation in the project is being supported through a supplement from the NSF for the R2R project, and the Australian contribution is being sponsored by the Australian government.

Further project details are available on the ODIP website: http://www.odip.org

Management of events, samples, videos and data generated by deep-sea submersibles

Catherine Borremans, Ifremer, catherine.borremans@ifremer.fr (France)

In the context of the renewing of Biocean -the Ifremer database for deep-sea benthic ecological data- evolutions are planned in order to open the system to other disciplines and to adapt the related tools to new technological constraints. Entire submersible datasets must be managed in a standardised way, from the event generating the data to the final data banking ashore. However it was not the case until recently for certain types of events (*e.g.* geology) or data (*e.g.* video).

The project aims to develop the onboard systems to record and report the full range of events (called "operations") taking place both onboard the research vessel and from submersible and to track samples obtained from them.

On land, operations are intended to be integrated in the central Cruise dabase (SISMER), they constitute the link to the data and the samples. Geological and biological samples will be managed and followed until analysis or final identification thanks to specialised softwares.

Final data resulting from submersible dives are diverse : bathymetry, ADCP, CTD, seismic, videos, pictures, navigation and biological, hydrological and geological samples analyses. They will be archived in the dedicated data banks (geophysics, videos, biology, geology, hydrology).

Although appropriate softwares and databases covering numerous oceanographic fields already exist in Ifremer, some of them have to be developed. Particularly, submersible videos have been centrally safeguarded since 2010 but have not been published yet. Another crucial part of the project is then to set up the whole archiving and diffusing system for these data.

This work to organise an integrated submersible data management system is tightly linked to European and International projects and initiatives such as OBIS, SeadataNet, Eurofleets and ODIP. Making references to standards established by those consortiums but also participating in the development and implementation of standards for the formats of metadata, data and common vocabularies are both important tasks necessarily included in the project. This especially allows interoperability and connexion of Ifremer biological data through the national, European and international portals in order to feed the Biodiversity programs (*e.g.* Census of Marine Life, WoRMS) and to comply with the European directives (*e.g.* INSPIRE, MSFD).

The communication presented at IMDIS 2013 will give an overview of the developments for the onboard events and samples recording tools. The resulting data flow and the submersible data management system on land will be described, with a focus on the new scientific videos library. The relations with global projects will be shown.

Temperature and salinity historical data collections for the european marginal seas: aggregation and quality assessment procedures

Simona Simoncelli, INGV, simona.simoncelli@bo.ingv.it (Italy)
Christine Coatanaon, Ifremer, Christine.Coatanoan@ifremer.fr
Örjan Bäck, SMHI, orjan.back@smhi.se
Helge Sagen, IMR, helge@imr.no
Serge Scory, MUMM, Serge.Scory@scarlet.be
Devrim Tezcan, METU, dtezcan@gmail.com
Dick M.A. Shaap, MARIS, dick@maris.nl
Reiner Schlitzer, AWI, Reiner.Schlitzer@awi.de
Sissy Iona, HCMR, sissy@hnodc.hcmr.gr
Michèle Fichaut, Ifremer, Michele.Fichaut@ifremer.fr
Marina Tonani, INGV, marina.tonani@bo.ingv.it

Temperature and Salinity (TS) historical data collections from 1900 were created for each European marginal sea within the Framework of SeaDataNet2 (SDN) EU-Project. V1 collections of data were created to meet operational oceanography and climate change community requirements that need longer and longer time series of in situ observations to study long term ocean phenomena and their implications in the surrounding environment. This work has been developed in synergy with MyOcean In-Situ Thematic Assemble Centre (INS-TAC) to support and promote monitoring, modeling or downstream service development.

The harvesting procedure of temperature and salinity files was performed by a new CDI Robot that used the CDI Data Discovery and Access Service to query, shop and retrieve data sets from the distributed data centers (NODC) in an automatic way. Firstly the shopping mechanism has been tuned through a series of “massive” requests of data. Then the robot retrieved the whole dataset as ODV files including the full CDI metadata, which were then aggregated into a single TS Data Collection using SDN Importer of ODV 4.5.3. It followed the creation of regional and 1900-2012 subsets and the distribution to SDN regional groups, responsible of SDN products, in order to perform quality assessment analysis.

Before and during the harvesting and aggregation procedures many efforts have been done to assure the best quality of the V1 collections of data. A screening procedure was applied to identify duplicates and clean SDN infrastructure from redundancies. Numerous files were found to not comply with the ODV/SDN format specification and were rejected. These files have been corrected at the NODCs level.

Six TS data collections, one per each European marginal sea (Arctic Sea, Baltic Sea, North Sea, North Atlantic Ocean, Mediterranean Sea, Black Sea), were then analyzed at regional level to assess and certify the quality of these products. The objective was twofold to report to the NODCs about further and necessary improvements for next data collection release and to report to the users all the procedures utilized. Basic quality check (QC) procedures were applied in a coordinated way to all data collections to assure a progressive harmonization of product quality.

SDN V1 TS data collections will be presented for the first time in occurrence to their official release. All basic QC procedures will be described and the results summarized.

MyOcean In-Situ Thematic Center: A service for operational Oceanography

Sylvie Pouliquen, Ifremer, Sylvie.pouliquen@ifremer.fr ,Brest, France
Thierry Carval, David Guillotin, Christine Coatanoan, Thomas Loubrieu, IFREMER, Brest, France
Antoine Grouazel, Karina Von Schuckmann, CNRS, Brest, France
Henning Wedhe, Lid Sjur Ringheim, IMR, Bergen, Norway
Thomas Hammarklint, Anders Hartman, SHMI, Stockholm, Sweden
Kai Soetje, Tobias Gies, BSH, Hamburg, Germany
Marta De Alfonso, Puertos Del Estado, Madrid, Spain
Leonidas Perivoliotis, Dimitris Kassis, Antonis Chalkiopoulos, HCMR, Athens, Greece
Veselka Marinova, IOBAS, Varna, Bulgaria
Pierre Jaccard, AnnaBirgitta Ledang, Kai Sorensen, Niva, Bergen, Norway
Giulio Notarstefano, OGS, Trieste, Italy
Joaquin Tintore, IMEDEA/SOCIB, Mallorca ,Spain
Seppo.Kaitala, SYKE, Helsinki, Finland
Petra Roiha, FMI,Helsinki, Finland
Lesley Rickards, BODC, Liverpool, UK
Giuseppe Manzella, ENEA, La Spezia, Italy

MyOcean is the implementation project of the GMES Marine Core Service to develop the first concerted and integrated pan-European capacity for Ocean Monitoring and Forecasting. Within this project, the in-situ Thematic Assembly Centre (in-situ TAC, INS-TAC) of MyOcean is a distributed service integrating data from different sources for operational oceanography needs. The MyOcean in-situ TAC is collecting and carrying out quality control in a homogeneous manner on data from outside MyOcean data providers, especially EuroGOOS partners in Europe, to fit the needs of internal and external users. It provides access to integrated datasets of core parameters (temperature, salinity, current, sea level, chlorophyll, oxygen and nutrient) to characterise ocean state and ocean variability, by this contributing to initialization, forcing, assimilation and validation of ocean numerical models. Since the primary objective of MyOcean is to forecast ocean state, the initial focus is on observations from automatic observatories at sea (e.g. floats, buoys, gliders, ferrybox, drifters, SOOP) which are transmitting to the shore in real-time. The second objective is to set up a system for re-analysis purposes that integrate data over the past 20 years. The global and regional portals set up by the INS-TAC have been extended by the EuroGOOS ROOSes (Arctic ROOS, BOOS, NOOS, IBI-ROOS, MOON and Black Sea GOOS) to integrate additional parameters (wind, waves,..) important for downstream and national applications.

The product and services provided by the MyOcean in situ thematic assembly centre will be presented as well as how it developed in partnership with EuroGOOS, JCOMM and SeaDataNet to provide products useful for operational oceanography needs both for Forecasting and reanalysis activities but also useful to the research communities.

Marine biogeographic data in EurOBIS: assessing their quality, completeness and fitness for use

Leen Vandepitte, Flanders Marine Institute, leen.vandepitte@vliz.be (Belgium)
Filip Waumans, Flanders Marine Institute, filip.waumans@vliz.be
Lennert Tyberghein, Flanders Marine Institute, lennert.tyberghein@vliz.be
Bart Vanhoorne, Flanders Marine Institute, bart.vanhoorne@vliz.be
Francisco Hernandez, Flanders Marine Institute, francisco.hernandez@vliz.be

The European Ocean Biogeographic Information System - EurOBIS - is an online marine biogeographic database compiling data on all living marine creatures (www.eurobis.org). The principle aims of EurOBIS are to centralize the largely scattered biogeographic data on marine species collected by European institutions and to make these data freely available and easily accessible.

The available data are either collected within European marine waters or by European researchers and institutes outside Europe. The database focuses on taxonomy and distribution records in space and time; all data can be searched and visualized through a set of online mapping tools. All data are freely available online and easily accessible, without requiring a login or password.

Given the very diverse nature of the data - going from museum collection data to literature data and research and monitoring data -, the standardization of both the data and the data format and evaluating the quality is not always evident. To simplify this task, a set of quality control procedures have been developed, encompassing taxonomic, geographic, outlier and data format checks.

The aim of these quality control procedures is two-fold. First of all, it helps the data management team and data providers to check the quality and completeness of the submitted data and detect (possible) errors. Records not meeting the assumed quality standards are sent back to the provider for a secondary check-up, clarification and/or corrections. This back-and-forth communication between the provider and the data management can greatly improve the quality of the submitted dataset, thereby also enhancing the quality of the data available in the integrated data system. Secondly, the assigned quality flags can help users in selecting data from EurOBIS that are fit for their use and purpose.

On the data management level, each individual distribution record is submitted to 20 quality control (QC) steps, generating 20 quality control flags. These steps include e.g. a check on the completeness of the required fields of the OBIS data scheme, a verification of the filled in values for the date related fields (e.g. is the month value between 1 and 12 or does the start date precede the end date) or checking whether the given sampling depth is a possible value compared to existing depth profiles. One of the most important checks is related to the taxon name of a record:

any taxon name within EurOBIS should be matched to the World Register of Marine Species (WoRMS, www.marinespecies.org), the most authoritative and comprehensive list of names of marine organisms, including information on synonymy. Only be linking to WoRMS, it is possible to rule out spelling variations and synonyms. This allows grouping of distribution records in a reliable way for further analyses. If a taxon name does not appear in WoRMS, this is further investigated: If it would consider a marine taxon not yet present in WoRMS, the appropriate taxonomic editors are contacted, so the taxon can be added and linked. If the name is from a non-marine taxon or does not make sense, it is added to an annotated list, explaining why it is not documented in WoRMS.

Next to these 20 QC steps on record level, additional geographic and taxonomic outlier analyses are run on respectively the dataset and the entire database, generating 2 more quality control flags. The geographic outlier analysis compares the observation points within a dataset and identifies possible outliers.

This kind of check can reveal possible errors in the latitude and longitude values, e.g. because of switched latitude and longitude values or a missing minus sign to indicate south or west. The taxonomic outlier analysis runs on the entire EurOBIS data system and will identify species that are documented outside their normal occurrence range, which is based on the available data within EurOBIS. These species outliers will need further investigation and verification, as they can be actual outliers or they can be new occurrences of the species in a previously undocumented geographic area. Both these outlier analyses will thus help in assessing the validity of a record compared to all available distribution records within one dataset and within the entire EurOBIS data system.

All these automated quality control steps will also be implemented on the international OBIS data system, greatly improving the 'fitness for use' of marine species distribution data on an international level.

The assigned quality control flags can be combined according to the required 'fitness for use' for the users, thereby creating specific filters on the available data within EurOBIS. The European Marine Data and Observation Network (EMODnet) Biology Portal (http://bio.emodnet.eu/portal) is currently applying such a filter. It only makes available those distribution records that comply with the following QC steps: the required fields - according to the OBIS data schema - are completed, the taxon name relates to a genus or species and is listed in the World Register of Marine Species (WoRMS), and the coordinates are – format wise – correct. On the dataset level, users can see how many records have passed the postulated quality control procedures and are thus available through the portal.

For the individual data providers, a number of the developed quality control procedures are offered as a web service through the LifeWatch Portal (www.lifewatch.be), enabling researchers to run certain checks on their data themselves. These web services currently encompass a data format check, a geographical check, a taxon name check and a check for (possible) duplicate records. The results of each check are directly available in an output file, so the provider can immediately check and – where necessary – adapt the uploaded file. The taxon check not only includes a match with the World Register of Marine Species (WoRMS), but also has look-up functionalities for other taxonomic databases such as e.g. the Integrated Taxonomic Information System (ITIS), the Catalogue of Life (CoL) and the Integrated Register of Marine and Non-Marine Genera (IRMNG).

Checking taxon names against several taxonomic registers will already help the provider in assessing the validity of the name and indicate possible errors (e.g. a terrestrial animal in a deep-sea dataset), before submitting it to EurOBIS.

The development of these QC procedures is part of the VLIZ contribution to LifeWatch, and funded by the Hercules Foundation. The main goal is to facilitate the fitness for use of individual and integrated biogeographic data for scientists, by offering several tools that help in the assessment of the completeness and validity of distribution records.

PubFlow: a scientific data publication framework for marine science

Peer C. Brauer, Kiel University, peer.brauer@informatik.uni-kiel.de (Germany)
Wilhelm Hasselbring, Kiel University, wilhelm.hasselbring@informatik.uni-kiel.de

We are facing a deluge of data. Improvements in technology and research methods have considerable increased the amount of data collected by scientific experiments. An example for this evolution of scientific instruments and methods is the development in the domain of depth measurement. For years this was a data poor science. The depth of the sea was calculated by using a hand lead or later a simple echo sounder. Today things have changed. Modern multi-beam echo sounders produce a detailed and also data rich height profile of the sea floor.

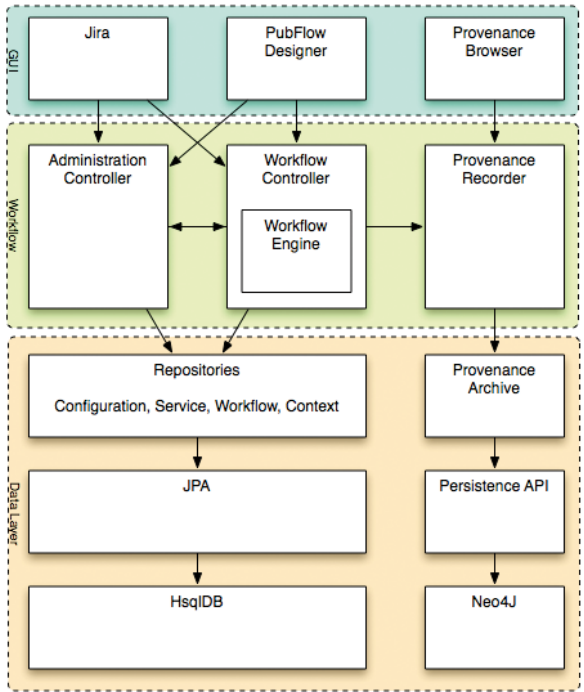


Fig. 1 - The architecture of the PubFlow system.

But these developments rise a new question, how to deal with the bulk of data created by all these experiments? Scientific data is too valuable to just be used once. On the other hand, new research methods, which developed during the last years and which Jim Gray [Bell et al., 2009] referred to as the fourth paradigm of science, base on the reuse of huge amounts of data from other experiments. So new ways and methods have to be discovered to alleviate the reuse of data and to get it out of the local repositories into public available archives [Fleischer et al., 2011]. PubFlow [Brauer at al., 2013] is our suggestion for such a tool.

PubFlow

PubFlow is a data publication framework for scientific data, build on top of proven business workflow technologies like BPMN 2.0[1], Apache ODE[2] and JBoss JBPM[3]. It brings automation and the division of work to the domain of scientific data management. Pubflow is based on the assumption, that data managers know best about the processes and guidelines, which have to be followed to publish research data to a public available archive. Unfortunately the amount of scientific data is so overwhelming, that data managers alone can not curate each dataset and upload it to the archives. Researchers, institutes and funding agencies on the other hand want their research data to be published. This factor was considered, when the PubFlow system was planed. In PubFlow the role of the data managers is to define the publication workflows and to take care for complex tasks. The publication process for a specific datset on the other hand is started by a scientist.

He chooses a predefined workflow meeting his requirements from a list and starts it through a normal ticket system like i.e. Jira[4] providing his datset as input. After this the PubFlow system runs the selected workflow, which was predefined by the data managers, on the dataset the scientist uploaded. Every time a problem occurs and the workflow can not be continued, PubFlow creates a new ticket in the ticket system and assigns it to a datamanager or to the researcher, who uploaded the dataset to the PubFlow system. If the problem described by the ticket is marked to be solved, PubFlow continues the workflow execution.

Fig. 1 depicts the architecture of the PubFlow system. On the GUI-Layer at the top one can find three user interface components. This reflects the division of work in the PubFlow system. Jira is the interface for the researchers to start the workflows, the PubFlow designer is the interface for the datamanagers to create and maintain the workflows. The third interface provides access to pubflows provenance system [Brauer at al., 2012; Brauer at al., 2013]. During the exection of each workflow PubFlow collects provenance information for the data currently processed by the workflow. This information is available through the PubFlow Provenance Browser.

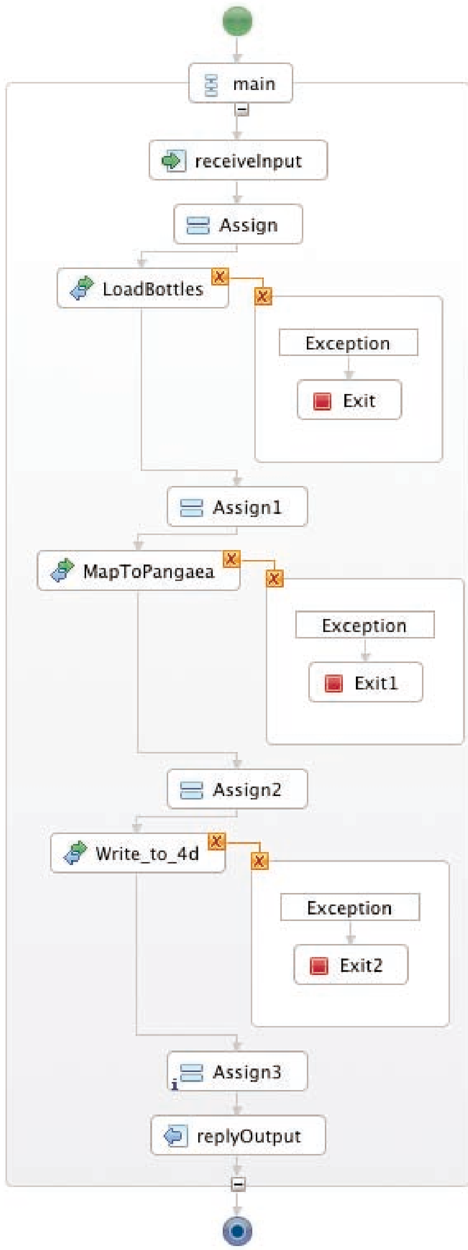


Fig. 2 - PubFlow workflow.

The evaluation scenario

The PubFlow system is evaluated in cooperation with the Kiel Datamanagement Team[5] located at the Geomar Helmholtz Centre for Ocean Research Kiel. The first publication workflow we implemented for PubFlow is a workflow for publishing data collected by a CTD probe. This data is stored in an institutional repository and has to be transfered to the world data center mare – Pangaea[6]. Fig. 2 shows this workflow. It mainly consists of five steps. In a first phase the information, the scientist provided, when he started the task, is loaded from the ticketsystem into the workflow engine.The next step is to fetch the data from the institutional repository. Now PubFlow performs predefined mapping and conversion operations on the dataset, so the output data format is compatible to the one used by the world data center. At last the data is written to the specified output format and exported. Although it is still under active development, PubFlow has already proven to be a very helpful tool for data managers by automating simple or periodic data management tasks.

References

BELL, G., HEY, T., AND SZALAY, A. *Computer science. beyond the data deluge.* Science 323, 5919 (2009), 1297–8.

BRAUER, P. C., AND HASSELBRING, W. *Capturing provenance information with a workflow monitoring extension for the kieker framework. In Proceedings of the 3rd International Workshop on Semantic Web in Provenance Management,* vol. 856 of CEUR Workshop Proceedings, 2012, CEUR-WS.

BRAUER, P. C., AND HASSELBRING, W. *PubFlow: provenance-aware workflows for research data publication In: 5th USENIX Workshop on the Theory and Practice of Provenance (TaPP '13),* April 2-3, 2013, Lombard .

FLEISCHER, D., AND JANNASCHK, K. *A path to filled archives. Nature Geoscience 4 (2011)*, 575–576.

[1 Mark] omg.org/spec/BPMN/2.0/

[2] ode.apache.org

[3] jboss.org/jbpm

[4] atlassian.com/jira

[5] geomar.de/en/service/data-management/

[6] pangaea.de

INSPIRE compliant international standards for the SeaDataNet marine metadata

Enrico Boldrini, CNR – Institute of Atmospheric Pollution Research, enrico.boldrini@cnr.it (Italy)
Stefano Nativi, CNR – Institute of Atmospheric Pollution Research, stefano.nativi@cnr.it
Dick M. A. Schaap, Marine Information Service (MARIS), dick@maris.nl

SeaDataNet implements a distributed pan-European infrastructure for Ocean and Marine Data Management whose nodes are maintained by 80 national oceanographic and marine data centers from 35 countries riparian to all European seas. Marine metadata play an important role in such an infrastructure, by providing an efficient description of the sea resources of interest (e.g. dataset products, cruise reports). As an example, different tools and services currently deployed in SeaDataNet makes use of marine metadata in order to enable the inventory, discovery, access and use of the underlying marine data:

- The Mikado metadata editor is used to create and edit both products and cruise metadata at a member node;
- A unique web portal makes possible distributed discovery, visualization and access of the available sea data products (described by their metadata) across all the member nodes;
- The CSR inventory web portal is used to search for Cruise Summary Reports (described by their metadata);
- Catalogue services such as GI-cat implementing international standard protocols and data models (e.g. CSW ISO Application Profile) are used to enable discovery of the available sea data products by different clients (other than the SeaDataNet web portal).

In the context of the second phase of SeaDataNet (SeaDataNet 2 EU FP7 project, grant agreement 283607, started on October 1st, 2011 for a duration of 4 years) a major target is the setting, adoption and promotion of common international standards, to the benefit of outreach and interoperability with the international initiatives and communities (e.g. OGC, INSPIRE, GEOSS, ...).

A standardization effort conducted by CNR with the support of MARIS, IFREMER, BSH, STFC, BODC and ENEA has led to the creation and adoption of ISO 19115 metadata profiles (and ISO 19139 based XML encoding) for SeaDataNet Common Data Index (CDI) and Cruise Summary Report (CSR):

- CDI metadata describe the products datasets, including identification information (e.g. product title, interested area), evaluation information (e.g. data resolution, constraints) and distribution information (e.g. download endpoint, download protocol);
- CSR metadata describe cruises and field experiments at sea, including identification information (e.g. cruise title, name of the ship), acquisition information (e.g. utilized instruments, number of samples taken).

The CDI and CSR metadata models have been mapped to the ISO 19115 and ISO 19115-2 metadata models, finding relations with correspondent elements or creating extended elements

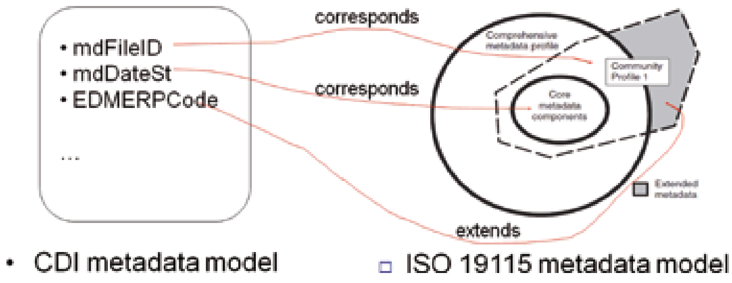


Fig. 1 - Mapping of the CDI metadata model to the ISO Geographic metadata model.

(Fig. 1). The resulting sets of metadata elements, together with a list of constraints demanded by INSPIRE guidelines for metadata constitute the metadata profiles.

The correspondent XML encoding is based on ISO 19139 schema, with additional elements defined in the SeaDataNet namespace. Schematron rules have been also drafted to accompany the schema, in order to formulate and enforce constraints out of the XML schema scope.

The SeaDataNet metadata profiles have been published on the SeaDataNet site, along with XML schema, Schematron rules, XML encoding samples and documentation.

The adoption of well-known ISO standards for metadata, in particular 19115 "Geographic metadata", its second part 19115-2 and its XML encoding 19139 (as also indicated by INSPIRE directive implementing guidelines), eases interoperability with international communities. Indeed, a general user of SeaDataNet services is able to understand the metadata content he or she has acquired (even if with a custom client), being it composed at the core by well-known elements. Extended elements, specifics of the SeaDataNet community, can also be understood, being them created (and documented) according to a precise, standardized extension methodology.

CDI and CSR records amount to about one million records in SeaDataNet and the new formats are being also adopted in the context of different international projects and initiatives, such as the Ocean Data Interoperability Platform (ODIP), EuroFleets, EMODNet …

The easy integration of SeaDataNet metadata into the Global Earth Observation System of Systems is a successful case of interoperability with international initiatives. The GEOSS Discovery and Access Broker (GeoDAB) is currently able to discover and access SeaDataNet resources amongst all the international sources pertaining to the different areas.

Fig. 2 - SeaDataNet data products discovered through the GEOSS Discovery and Access Broker.

Vocabulary enhancements for the Australian Ocean Data Network

Roger Proctor, Integrated Marine Observing System, roger.proctor@utas.edu.au (Australia)
Kim Finney, Australian Antarctic Division, kim.finney@aad.gov.au
Sebastien Mancini, Integrated Marine Observing System, sebastien.mancini@utas.edu.au

The Australian Ocean Data Network (AODN) is a rapidly growing distributed data network bringing together marine data collections from Commonwealth Agencies, Universities, State Governments, national programs and private industry. These data are made publicly available through the AODN portal (http://portal.aodn.org.au), an open source information infrastructure itself downloadable from https://github.com/aodn/aodn-portal. Increasingly, the data collections are multi-disciplinary requiring access to multiple layers of information from different sources.

The functional diagram of user interaction with the AODN portal is shown in the figure. A user either selects a 'search' query to discover what data is available, or uses the map to display data sources directly from a menu of contributing organisations/programs. The portal includes an option to save searches and map displays for future use.

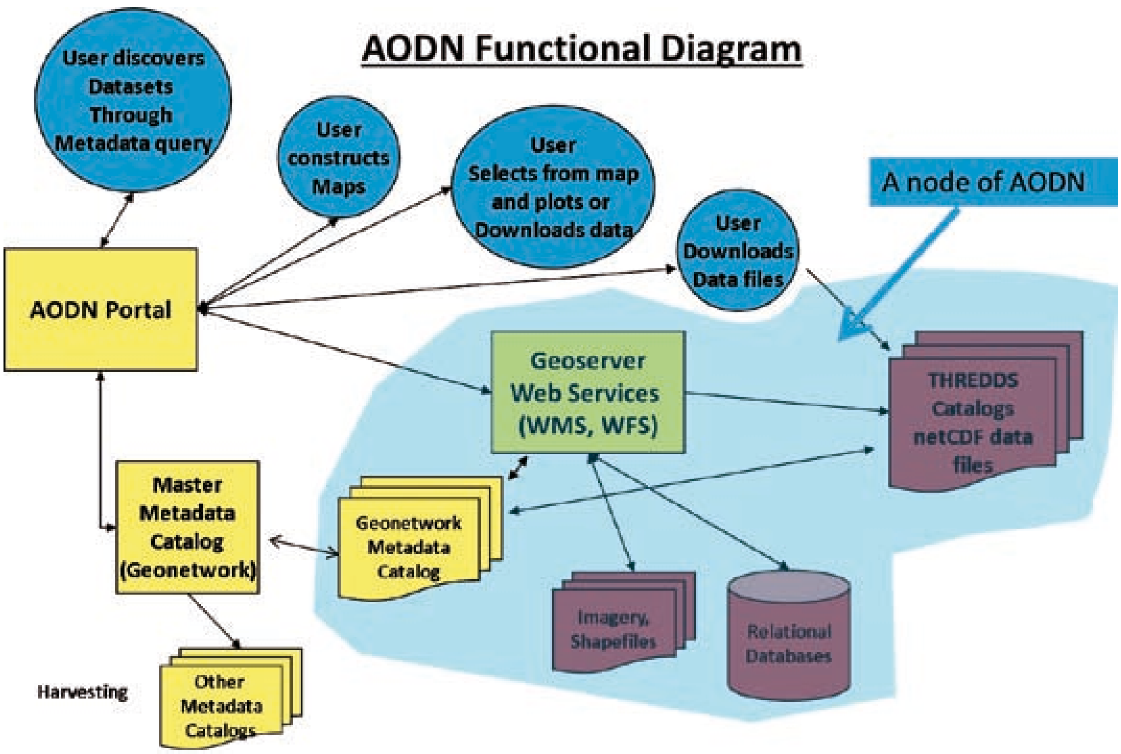


Fig. 1 - AODN Functional Diagram.

The AODN uses the Marine Community Profile (MCP), both a subset and an extension of the ISO19115/19139 geospatial metadata standard, and the Geonetwork metadata catalogue as its search engine. Nodes of the AODN comprise a set of data services, including a Geonetwork catalogue instance, and these catalogues are routinely harvested into a master AODN catalogue. To refine searches the AODN employs a facetted search option. This requires rich metadata to enable the appropriate layers to be discovered and integrated. To date *'subject keywords'* have been a primary source of search terms but recently the need has arisen to identify the role that 'subject keywords' are intended to play in a metadata record, given the existence of other notionally similar elements (e.g., *'parameters'*). Currently proposed amendments to the MCP would see a number of controlled vocabularies introduced to underpin 'parameter' identification, thus paving the way for improved faceted searching.

The existing MCP *'parameter'* name element can be used to record 'controlled and moderated' names. But in order for the metadata record to contain all of the semantics necessary to delineate different or similar observations it is taken for granted that the semantic elements not in a *'parameter'* name, but which are required to delineate different types of observations, must also be present in the record (for example, information associated with platforms, instruments, statistical and analytical methods).

AODN is currently compiling a vocabulary database of *'parameter'* names that match its business rules, which favour granular and discrete semantic descriptions. Where-ever possible AODN is re-using names from existing sources, primarily the BODC Parameter Usage Dictionary (P011, see https://www.bodc.ac.uk/data/codes_and_formats/parameter_codes/). New terms are being created where they don't already exist and where BODC, or other term names, are too complex. In which case, only the semantic elements of interest are being borrowed. The relationships between any compiled AODN vocabulary terms and those that are similar from other sources are also being recorded. Governance will be needed to ensure any new terms are fed back into, or mapped to, existing vocabularies.

The intent is to publish the parameter vocabulary online through at least two mechanisms: in GeoNetwork as a register of AODN *'parameter'* names that can then be used within Geonetwork as a 'parameter' thesauri and through REST and SPARQL-based AODN vocabulary service end-points (most likely using SISSVoc - https://www.seegrid.csiro.au/wiki/Siss/SISSvoc3Overview). Data providers wishing to mark-up their datasets inline using *'Parameter'* names could then include 'Xlinks' to reference 'http://addressed' AODN *'Parameter'* names from the Vocab Service. The GeoNetwork-based vocabulary source will need to be kept in synchronisation with the content from any other type of service end-point.

From silos to systems: using standardised services and Linked Data methodologies to augment government and research information systemse

Jonathan Hodge, CSIRO, jonathan.hodge@csiro.au (Australia)
Nicholas Car, CSIRO, nicholas.car@csiro.au (Australia)

Over the last decade in Australia, government agencies, research institutions and other environmental organisations have made significant progress in publishing, sharing and exposing environmental information. The use of standardised services and open licensing has enabled improved discovery and re-use. This gradual evolution in data availability has, however, occurred in an ad-hoc manner with minimal strategic coordination in terms of agreed standards or definitions around data sharing, services or supporting information. This has resulted in a diversity of technical and architectural solutions which has promoted the existence of data silos. Many of these solutions, however, do support a collection of geospatial standards, albeit using different implementations or technology stacks.

The challenge remains to connect these data silos and increase interoperability at various levels. Data services must be discoverable, self-describing and cross-linked to enable levels of use from general public to managers and policy advisors to researchers and data specialists. This requires more than just the standard collection of geospatial services, which often remain disconnected and incompatible.

The next evolution in information systems in Australia is creating the glue between these systems. Rather than developing a fully customised information suite which needs to be deployed in participating organisations, projects such as eReefs are developing frameworks which will allow existing systems and services to be better described, discovered and accessed in a consistent manner.

The eReefs project is developing a model of distributed Data Provider Nodes to integrate collections of marine and coastal data including hydrodynamic and biogeochemical ocean modelling, satellite remote sensing water quality products, time series monitoring data and other geospatial data layers from various organisations. The Data Provider Node model describes to contributing organisations how to supplement their existing geospatial data delivery platforms with supporting services such as vocabulary, persistent identifier and provenance services. This is facilitated partly by some of the new services themselves and partly by the Linked Data paradigm which extends the use of well-known Internet standards and protocols originally designed for web-based documents (information resources) to data, metadata and other resources.

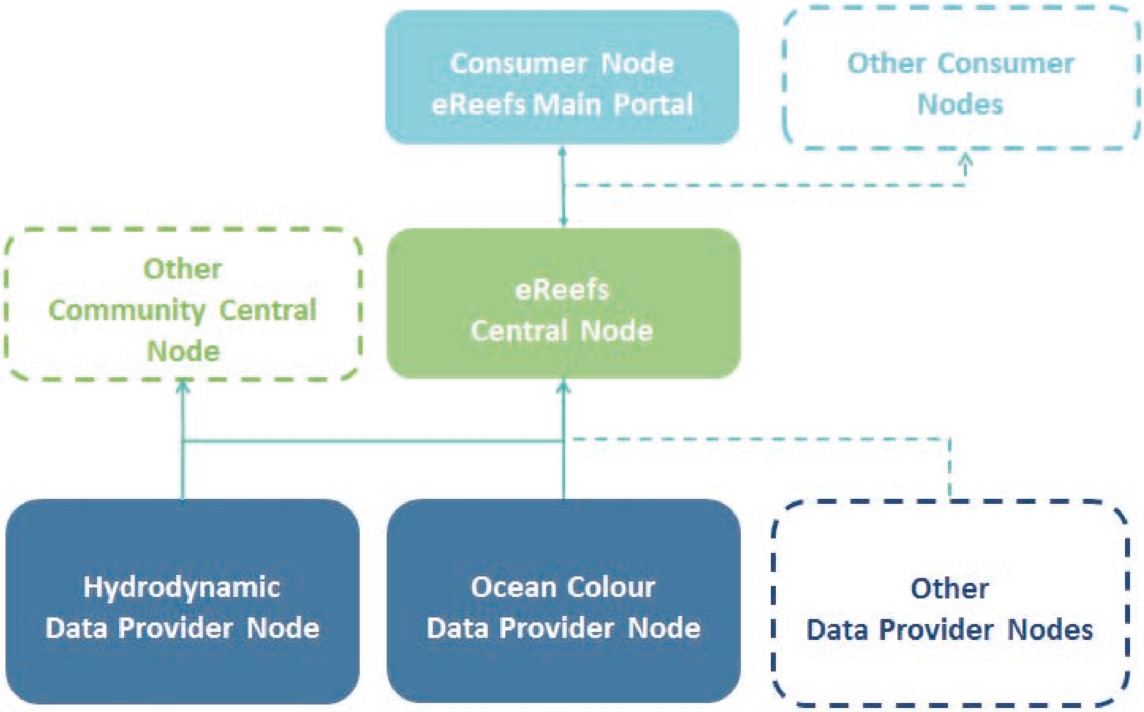


Fig. 1 - A high-level view of the eReefs architecture showing implementations of the three main architectural conceptual entities: Data Provider Nodes (DPNs), a Central Node and Consumer Nodes.

The new supporting services include:

- vocabulary services, defining the terms used in metadata;
- feature type catalogs, defining what 'things' the data is about;
- persistent identification services which allow features and terms to be located over time despite system re-implementation;
- and provenance services which detail the processes carried out in the production of the data.

These supporting services can be implemented in-house, alongside a contributor's existing data systems, or as a collection of distributed 'brokering' services in third-party locations using the Linked Data paradigm. Data Provider Nodes then expose data, metadata and other supporting information which any particular community – through a Central Node of its own – may integrate, add value to and then use for particular purposes. The eReefs project will support the development of an eReefs Central Node, however the Data Provider Nodes used by eReefs may also be used by any other community through their own Community Node implementation.

This approach will allow for the next generation of data discovery, integration and interoperability by augmenting and value adding, rather than replacing, existing systems thereby helping government agencies, research institutions and other environmental organisations to realise the true benefits of open data.

Rolling Deck to Repository (R2R): Working with International Partners to Improve Data Discovery and Access

Suzanne Carbotte, Robert Arko, Vicki Ferrini, Lamont-Doherty Earth Observatory, info@rvdata.us (United States)

Karen Stocks, Stephen Miller, Scripps Institution of Oceanography

Andrew Maffei, Cynthia Chandler, Woods Hole Oceanographic Institution

Shawn Smith, Mark Bourassa, Florida State University

The Rolling Deck to Repository (R2R) program was initiated in 2009 to provide uniform stewardship of routinely-collected environmental sensor data from the U.S. academic research fleet. R2R works with 25 vessels operating year-round/world-wide, and has cataloged over 20 terabytes of data and documentation from 3,188 cruises to date. The program is funded by the U.S. National Science Foundation and works closely with the University-National Oceanographic Laboratory System (UNOLS) Office, the U.S. NOAA National Data Centers, and many specialized/disciplinary data systems.

Fig. 1 - R2R (http://rvdata.us/) and ODIP (http://odip.org/) are working toward greater interoperability among ocean data providers.

R2R delivers services that include: inventorying cruises in a central catalog; organizing, documenting, and depositing sensor data from vessels in long-term archives; assessing the quality of selected datasets; creating quality-controlled products including shiptrack navigation, near real-time met/ocean data, and geophysical profiles; and providing tools for at-sea event logging.

The program provides a reference catalog of vessels, sensors, cruises, datasets, reports, etc, that aims to be authoritative (complete and correct, with persistent/unique identifiers), easily accessible (via lightweight, standard interfaces), and interoperable (mapped to international/community vocabularies). Catalog content is published in multiple formats according to user demand including ISO Metadata records, OGC Web Services, and most recently as W3C Linked Data. R2R has adopted the Natural Environment Research Council (NERC) Vocabulary Server (NVS; *http://vocab.nerc.ac.uk/*) as its reference thesaurus.

R2R participates in the Ocean Data Interoperability Platform (ODIP), a forum for harmonizing standards and developing best practices among data providers in Europe, Australia, and the United States in collaboration with the International Oceanographic Data and Information Exchange

(IOC/IODE). Planned work in ODIP includes publishing R2R content in the SeaDataNet Cruise Summary Report (CSR) profile; deployment of a GeoNetwork node; expansion of R2R Linked Data to include near real-time products; and deeper integration with vocabularies published by SeaDataNet, the International Council for the Exploration of the Sea (ICES), and EUROFLEETS.

References

Carbotte, S., V. Ferrini, R. Arko, C. Chandler, A. Maffei, S. Miller, K. Stocks, S. Smith, M. Bourassa (2009): *Rolling Deck to Repository (R2R): Transforming the Academic Fleet Into an Integrated Global Observing System,* EOS Trans. AGU, 90 (52), Fall Meet. Suppl., IN21B-1045.

Arko, R., C. Chandler, K. Stocks, S. Smith, P. Clark, A. Shepherd, C. Moore, and S. Beaulieu (2013): *Rolling Deck to Repository (R2R): Collaborative Development of Linked Data for Oceanographic Research,* Geophys. Res. Abstr. Vol. 15; EGU2013-9564.

Leadbetter, A., R. Lowry, and D. Clements (2013): *Putting meaning into NETMAR – the open service network for marine environmental data,* Intl. J. Digit. Earth, doi: 10.1080/17538947.2013.781243.

Schaap, D. and R. Lowry (2010): *SeaDataNet – Pan-European Infrastructure for Marine and Ocean Data Management: Unified Access to Distributed Data Sets,* Intl. J. Digit. Earth, doi:10.1080/17538 941003660974.

Glaves, H., D. Schaap, S. Miller, R. Proctor (2013): *Ocean Data Interoperability Platform (ODIP): Developing an International Collaboration in Marine Data Management,* Geophys. Res. Abstr. Vol. 15; EGU2013-12831.

ICOADS the Global Marine Surface Reference Dataset

Steven J. Worley, National Center for Atmospheric Research, worley@ucar.edu (USA)
Scott D. Woodruff, Cooperative Institute for Research in Environmental Sciences, Scott..D.Woodruff@noaa.gov
J. Eric Freeman, National Climatic Data Center/STG, Inc., Eric.Freeman@noaa.gov
Shawn R. Smith, Center for Ocean-Atmospheric Prediction Studies, Florida State University, smith@coaps.fsu.edu

The International Comprehensive Ocean-Atmosphere Data Set (ICOADS) is a unique and widely used archive of historical global *in situ* marine surface observations. It has been foundational data, for over 25 years, for many small and large-scale observational studies, global analyses of observed and computed quantities, and the major global atmospheric reanalyses from NOAA, NASA, ECMWF, and JMA.

The temporal coverage is remarkably long, 1662-2013, and as such holds data from many different observing systems. During early centuries the data are from sailing ships and explorers, where the geographical coverage is sparse and few parameters are recorded. Beginning in the 1850s, stimulated by international agreements, the geographical coverage improves through merchant vessel logbooks and the number of observed parameters increases, e.g. wind, temperatures, atmospheric pressure, and more.

This improving trend continues with notable coverage disruptions during the World Wars. Research vessels began adding data to ICOADS in the 20th Century and many new observing systems emerged starting in the 1970s, including moored and drifting buoys. Automated observing systems, with the purpose to improve operational weather prediction and ocean research, have become a significantly large influence on the total record count, presently representing over 80% of the marine surface observations.

A highly valued strength of ICOADS is the embedded metadata. Every record is traceable to its source collection and is documented with the type of observing system. The voluntary observing ships for 1973-2007 are also augmented with ship characteristic metadata from World Meteorological Organization holdings.

ICOADS has been open and freely available worldwide, since 1985, and has irregularly published new Releases, dictated by the resources available to collect and ingest new data, perform the multiple-pass data processing to merge the many sources that can span the full period of record, and quality check the new output. The collection keeps pace with near real-time by ingesting Global Telecommunication System (GTS) data and extending the time series on a monthly basis.

The current status and characteristics of ICOADS will be reviewed along with the plans for content development. In its primary form the data are coarsely quality controlled and not bias adjusted. To address this issue the ICOADS Value-Added Database (IVAD), which is under development, will be described. IVAD is designed to receive and redistribute bias adjustments that will be contributed by collaborating international partners. The value-added data will be distributed to the public in conjunction with the original observed values. The primary goal is to homogenize data across observing systems, and make these broadly available to support more climate and earth system research.

Making SeaDataNet more fit for handling biological data

Klaas Deneudt, Flanders Marine Institute, klaas.deneudt@vliz.be (Belgium)
Hans Mose Jensen, International Council for the Exploration of the Seas, hans.jensen@ices.dk
Bart Vanhoorne, Flanders Marine Institute, bart.vanhoorne@vliz.be
Leen Vandepitte, Flanders Marine Institute, leen.vandepitte@vliz.be
Simon Claus, Flanders Marine Institute, simon.claus@vliz.be
Francisco Hernandez, Flanders Marine Institute, francisco.hernandez@vliz.be

So far, SeaDataNet has focused on data management and access for physical oceanography, marine chemistry (to support also the EMODNet Chemistry pilot), bathymetry (to support the EMODNet Hydrography and Seabed Mapping pilots), and geology and geophysics (to support the Geo-Seas project and the EMODNet Geology pilot). Many partners in SeaDataNet are also involved in data acquisition and management for marine biology. A number are member of the Marine Biodiversity and Ecosystem Function (MarBEF) network of excellence and contributing to EurOBIS (European Ocean Biogeographic Information System), managed by the Flanders Marine Institute (VLIZ).

One of the objectives of SeaDataNet II is to undertake actions to make SeaDataNet better fit for handling marine biological data sets and establishing interoperability with biology infrastructure developments (Fig. 1). Therefore an analysis is undertaken in SeaDataNet II together with actors from the initiatives mentioned above as to how SeaDataNet can be best adapted for also handling marine biological data sets.

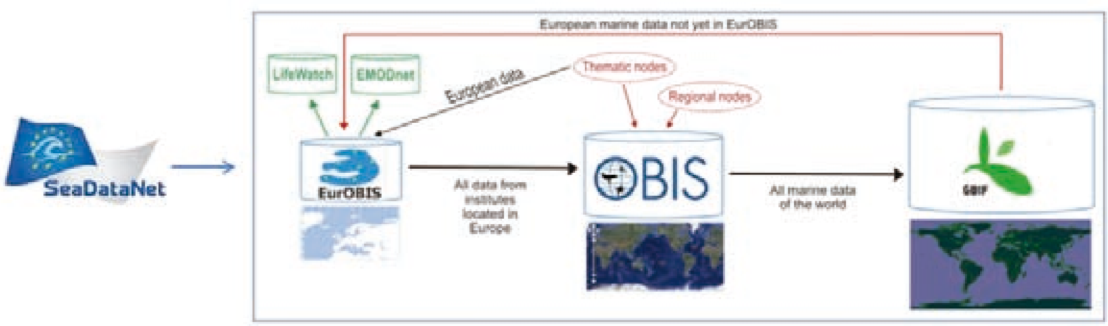


Fig. 1 - Data flow of SeaDataNet contributing to biology data infrastructure developments.

Based on an analysis of the present situation and currently existing biology data standards and initiatives, such as the Ocean Biogeographic Information System (OBIS), Global Biodiversity

Information Facility (GBIF), Working Group on Biodiversity Standards (TDWG) and World Register of Marine Species (WoRMS) standards, a recommended format for data exchange of biological data is being developed.

Key issues that steer the format development are:

- Requirements posed by the intended use and application of the data format (data flows, density calculations, biodiversity index calculations, community analysis, etc…)
- Availability of suitable vocabularies (World Register of Marine Species, SDN Parameter list, SDN Unit list, etc…)
- Requirements for compatibility with existing tools and software (WoRMS taxon match services, EurOBIS QC services, Lifewatch workflows, Ocean Data View, etc…)

This presentation describes the performed analysis and proposes format templates for several types of marine biological data.

MyOcean Central Information System

Thomas Loubrieu, ifremer, tloubrie@ifremer.fr (France)
Vincent Claverie, CLS, vclaverie@cls.fr (France)
Rémi de Dianous, CLS, rde-dianous@cls.fr (France)
Tony Jolibois, CLS, tjolibois@cls.fr (France)
Laia Romero, Altamira, laia.romero@altamira-information.com (Spain)
Joan Sala, Altamira, joan.sala@altamira.information.com (Spain)
Oscar Cases, Altamira, oscar.cases@altamira.information.com (Spain)
Antoine Bon, ifremer, abon@ifremer.fr (France)
Mickaël Treguer, ifremer, mtreguer@ifremer.fr (France)
Jon Blower, ESSC, University of Reading, j.d.blower@reading.ac.uk (United Kingdom)
Guy Griffiths, ESSC, University of Reading, guy.griffiths@reading.ac.uk (United Kingdom)

Since 2009, MyOcean project (http://www.myocean.eu) provides an operational service, for forecasts, analysis and expertise on ocean currents, temperature, salinity, sea level, primary ecosystems and ice coverage. The production of observation and forecasting data is done by **42 Production Units** (PU). Product download and visualisation are hosted by **25 Dissemination Units** (DU). All these products and associated services are gathered in a **single catalogue** hiding the intricate distributed organization of PUs and DUs.

Besides, applying INSPIRE directive and OGC recommendations, MyOcean overcomes technical choices and challenges so that end-user requirements and internal operation constraints are satisfied. This presentation overviews the architecture of the MyOcean Central Information System and focuses on 4 key components: 1) metadata management and catalogue, 2) web portal, 3) system and dataset availability monitoring, 4) user management.

Overview

Relying on a highly distributed organization, MyOcean system requires well defined components scope and thoroughly coordinated interfaces. This optimizes operation costs and system reliability: The data access (view, download) functions are implemented at Dissemination Unit level. The servers are optimally deployed near the datasets and almost autonomous. These components called Dissemination Unit Gateways are detailed in “Data access services in the highly distributed context of MyOcean: needs, technical challenges and implementation” (Jolibois et all, submitted at IMDIS2013). The central catalogue gets product description (e.g. forecasting model at resolution 1/12° on European Northwest Shelves) from the Production Unit managers. The datasets (e.g. data files on web services or FTP for view or download) description is automatically harvested from the dissemination servers operated at Dissemination Unit level. Then the collected information is used for presentation purpose (web portal) and automatic system and production monitoring. At the other end of the system, the users are registered and their activity is monitored for an efficient support activity.

Technically speaking, the usage of on-the-shelf open-source software is preferred as much as possible. This reduces the maintenance costs and increases the reliability of the service thanks to the communities supporting these applications (thredds data server, vftpd, geonetwork, nagios…).

Catalogue

Since April 2013, the catalogue is implemented by Geonetwork which provides both flexibility and reliability. Dedicated metadata editing GUI are proposed to Production Unit managers to ease product descriptions documentation in the system. The quality of metadata is enhanced by avoiding as much as possible manual edition. The usage of controlled vocabularies is preferred. The non specific vocabularies are synchronized with SeaDataNet Vocabulary Services (e.g. CF standard names).

As previously said, the datasets descriptions are automatically harvested from THREDDS Data Servers. The tool will also support the MyOcean Product Manager for preparing and coordinating the future releases of the catalogue (new products, new extents, new satellites…).

Web portal

The web portal front-end for discovery, viewing, soon for download is configured from the information collected in the catalogue. For doing so, it uses the OGC/CS-W v2.0.2 standard interface transporting products and datasets descriptions in ISO19115 standard and encoded in ISO19139 XML.. The OGC/WMS v1.3.0 protocol is also used as viewing back-end. For V4 in April 2014, the web portal will implement a faceted search. Thanks to these new advanced indexes cached at web portal level, the operational constraint on OGC/CS-W will be relaxed and the overall reliability of the system will be enhanced.

Fig. 1 - MyOcean product catalogue on the web.

Monitoring of systems and datasets availability

To match the operational requirements, a system, based on NAGIOS and customized for MyOcean services, checks and reports failures on any distributed component. This system is automatically configured from the metadata in the catalogue (service URL, dataset update schedule).

User registration

Together with SeaDataNet, MyOcean is looking to open its system to users who come from beyond its own centrally managed user directory. Not to mention the organizational issues, this may be technically done by either sharing a user directory in a wider marine community or setting up a federation with systems.

Candidate to become the future Copernicus marine monitoring service infrastructure, MyOcean Central Information System is now reaching a high level of maturity. Systems and procedures have proven their reliability and sustainability during the last 4 years (since myOcean1 kick-off, april 2009) to provide the expected level of service with stable operations and maintenance costs (around 2.5 full-time equivalent currently required to run the whole central systems: infrastructure, configuration). However, the willingness to continue along the same lines implies to remain at the cutting edge of technology and standards.

New gridded climatologies, from in-situ observations, for the Mediterranean Sea

Marina Tonani, Istituto Nazionale di Geofisica e Vulcanologia, marina.tonani@bo.ingv.it (Italy)
Simona Simoncelli, Istituto Nazionale di Geofisica e Vulcanologia, simona.simoncelli@bo.ingv.it
Alessandro Grandi, Istituto Nazionale di Geofisica e Vulcanologia, alessandro.grandi@bo.ingv.it
Nadia Pinardi, University of Bologna, Corso di Laurea in Scienze Ambientali, n.pinardi@sincem.unibo.it

New gridded climatologies for the Mediterranean Sea have been computed in the frame of the EU–project Seadatanet. The horizontal grid resolution is of 1/8 x 1/8 and 32 vertical levels. DIVA (Data Interpolating Variational Analysis) software, developed by GHER, University of Liege, has been used to compute the climatology fields and the associated error.

These climatologies rely on the Seadanaet CDI observations integrated with other dataset available for the Mediterranean Sea. Argo float profiles have been taken into account together with all the other observations like XBT, CTD, bottle and MBT.

The data distribution has been examined for each month and sensitivity experiments have been performed in order to define the best background field for the Variational Analysis. A different choice has been done for temperature and salinity fields. A three months moving window has been considered for temperature background computation while an annual time window has been set for salinity.

A Vertical extrapolation technique has been implemented in order to avoid unrealistic values on the bottom due to the decrease of the number of observations with depth. The impact of this technique has been assessed and justified.

Results have been evaluated against Medar/Medatlas climatology, observations and the existing literature.

The major change during the last century in the Mediterranean Sea in term of water masses has been the so-called “East Mediterranean Transient” (EMT). Dedicated studies have been performed in order to be able to reproduce climatologies for the pre-transient and the post-transient conditions. These climatologies have been tested in the initialization of the Mediterranean ocean Forecasting System (MFS) model.

This study will describe all steps performed for the production of the monthly mean climatology and it is focused on the validation of this product.

POSTERS

Semantically-enhanced Aggregation of SeaDataNet Data

Roy Lowry, British Oceanographic Data Centre, rkl@bodc.ac.uk (UK)
Adam Leadbetter, British Oceanographic Data Centre, alead@bodc.ac.uk (UK)
Matteo Vinci, Instituto Nazionale di Oceanografia e di Geofisica Sperimentale, mvinci@ogs.triestse.it (Italy)

The Problem

SeaDataNet is now a fully operational pan-European data system populated by millions of fine-grained data objects, typically containing a single CTD cast or instrument record. Individually they have limited scientific value and to deliver their full scientific potential they need to be aggregated into related groups, such as the profiles in a section or all the data from a geographic area of interest.

During the first phase of SeaDataNet the Ocean Data View (ODV)[1] visualisation software was developed into an extremely powerful tool for the aggregation of SeaDataNet data objects. File collections are simply 'dragged and dropped' onto the ODV icon and the software analyses their contents and organises them into aggregated collections of profiles, time series and trajectories. The parameters in the aggregated collections are defined using content from the BODC Parameter Markup Vocabulary (P01)[2] included in the SeaDataNet data.

This results in a parameter in the aggregated collection for every different P01 concept included in the all the input files included.

The P01 vocabulary is designed for data models that have limited usage metadata elements linked to the measurements. Consequently, the concepts can become heavily laden with information and the difference between some concepts can be extremely subtle. Furthermore, a long history (the vocabulary was founded in 1979), human frailty and a governance policy that strictly prohibits concept deletion has inevitably resulted in some cases where multiple P01 concepts are totally synonymous.

This means that in some cases the ODV aggregation process parses data into separate parameters that should have been merged into a single channel. Therefore, those using ODV or similar tools to operationally aggregate SeaDataNet, such as the pilot phase of the EMODNET chemical lot, were presented with an unnecessary workload as parameters needed be merged after the data had been aggregated.

The parameter merging process raises the immediate question 'Which parameters should be merged?' The answer is not necessarily straightforward and needs to be provided by a domain expert who has a thorough understanding of both the P01 concept definitions and the use to which the aggregated dataset is to be put. For example, can 'Concentration of chlorophyll-a per unit volume of the water body [particulate <0.2um phase]' be merged with 'Concentration of chlorophyll-a per unit volume of the water body [particulate <20um phase]'? The former is a measure of the pigment concentration in all types of phytoplankton, whereas the latter is a measure of the pigment in large phytoplankton cells, predominantly diatoms. Merging these might be OK were the aggregation a time series used to measure bloom duration, but could be misleading were the aggregation to be a spatial distribution.

Once the decision has been made, two further questions need to be answered:

- What should the merged parameter combination be called? This is important because it is the label that will be seen by users of the aggregated data products.
- How can the composition of the merged parameter be documented as readily accessible usage metadata? This is essential if aggregated products are not to be misinterpreted.

A Possible Solution

A very similar problem to this was encountered during the FP7 NETMAR project[3]. Data streams and service outputs were marked up using the same vocabulary (P01) as the SeaDataNet data. What the NETMAR service chain editor needed was a way of identifying which streams could sensibly be connected to a given service input. For example, how could any of the tens of P01 water temperature concepts be validated for a water temperature service input, whilst other P01 concepts were declared invalid?

The solution developed was to create a new vocabulary (P25)[4] in the NERC Vocabulary Server (NVS)[5] comprising coarse-grained parameter concepts, such as 'Water temperature' and 'Nitrate concentration'.

Domain expertise was used to map each of these P25 concepts to a group of P01 concepts and the resulting knowledge was captured and stored in the NVS. The result was a collection of concepts describing service inputs. Each of these has a Uniform Resource Name (URN) that translates to a Uniform Resource Locator (URL) conforming to Linked Data principles, a preferred label, an alternative label (which could be an abbreviation), and a definition that allows the semantics of merged parameter to be fully described in an understandable form. The URL resolves to a Resource Description Framework (RDF) XML document that conforms to the Rorld Wide Web Consortium (W3C) Simple Knowledge Organisation System (SKOS) reccomendation which includes all the concept labels plus the URLs of all the concepts to which it is mapped. For example, the URL:

http://vocab.nerc.ac.uk/collection/P25/current/NO3/

returns an XML document that includes mappings between the P25 concept and P01. In other words, six P01 concepts that may be merged to produce a parameter we can label 'Nitrate concentration'. This information is permanently available on the internet to either inform a software application of the P01 concepts that may be merged together or to document the P01 concepts that have been merged.

A Way Forward for EMODNET?

The process of building a semantically-enabled data aggregation system involves the following steps:

- Mark up the source data using a suitable parameter usage vocabulary
- Build a collection (i.e. controlled vocabulary) of merged parameter concepts

- Map these concepts to the usage vocabulary concepts
- Enable aggregation tools to use the SKOS documents served by NVS for merged parameter concepts to decide upon the P01 concepts to combine and the labels to use for the result.

By far the most labour-intensive stage in this process is the first which SeaDataNet has already completed. Building and mapping the merged concept controlled vocabulary requires input from experts, but a comparatively small amount of effort. Enabling the software, either through direct internet query or cached copies of the SKOS documents, is by far the smallest undertaking. It is proposed that this is the way forward for the operational phase of most of the EMODNET lots.

Acknowledgements. In the European Union this work has been funded through the Framework Programme 7 projects NETMAR (2010-3), SeaDataNet-II (2011-5) and Ocean Data Interoperability Platform (2012-5).

References

[1] http://odv.awi.de/

[2] http://seadatanet.maris2.nl/v_bodc_vocab/search.asp?name=(P011)%20BODC+Parameter+Usage+Vocabulary&l=P011

[3] http://netmar.nersc.no/

[4] http://vocab.nerc.ac.uk/collection/P25/current/

[5] http://vocab.nerc.ac.uk/

CDIAC Data Management Support for Ocean Carbon Dioxide Measurements

Alex Kozyr, CDIAC, Oak Ridge National Laboratory, kozyra@ornl.gov (USA)

Rising atmospheric CO_2 and climate change are increasing ocean temperatures and affecting ocean chemistry (e.g., ocean acidification). Monitoring these important changes using ships and other platforms generates large amounts of data from heterogenous sources. Since its inception in 1993, when it became a member of the DOE/NOAA Ocean Carbon Science Team engaged in the World Ocean Circulation Experiment (WOCE), the Carbon Dioxide Information Analysis Center (CDIAC) Ocean Carbon Data Management Project has been organizing, quality assuring, documenting, archiving and distributing ocean carbon-related data collected via a number of U.S. and international ocean-observing programs.

CDIAC's ocean carbon data collection includes discrete and underway measurements from a variety of platforms (e.g., research ships, commercial ships, buoys). The measurements come from deep and shallow waters from all oceans. Technological advances make it possible to deliver ocean carbon data real-time but questions about instrument reliability and data quality limit this practice at this moment. All ocean carbon data CDIAC receives come from individual investigators and groups following initial data review.

World Ocean Circulation Experiment (WOCE). The CDIAC Ocean Carbon Data Management Project started in 1993 when CDIAC became a member of the DOE/NOAA Ocean Carbon Science Team with data management and permanent archive responsibilities for the CO_2 measurements during WOCE cruises.

The resulting WOCE carbon database is available from the CDIAC Ocean Web site (http://cdiac.ornl.gov/oceans/CDIACmap.html). WOCE was a major component of the World Climate Research Program with the overall goal of better understanding the oceans role in climate and climatic changes resulting from both natural and anthropogenic causes. The CO_2 survey took advantage of the sampling opportunities provided by the WHP cruises during this period between 1990 and 1998. The final data set covers approximately 23,000 stations from 94 WOCE cruises.

CDIAC provides data management support for the International **Global Ocean Carbon and Repeat Hydrography Program**. The Global Ocean Carbon and Repeat Hydrography Program carries out a systematic and global re-occupation of select WOCE/JGOFS hydrographic sections to quantify changes in storage and transport of heat, fresh water, carbon dioxide (CO_2), and related parameters. The high-quality discrete measurements of carbon-related parameters are available via CDIAC Repeat Hydrography web site (http://cdiac.ornl.gov/oceans/RepeatSections/), and the *Mercury* metadata search engine.

CDIAC provides data management support for the **Global Volunteer Observing Ship (VOS)** Program. The VOS project is coordinated by the International Ocean Carbon Coordination Project (IOCCP). The international groups from 14 countries have been outfitting research ships and commercial vessels with automated CO_2 sampling equipment to analyze the carbon exchange between the ocean and atmosphere. The high-quality surface (underway) measurements of carbon-related parameters are available via CDIAC VOS web site (http://cdiac.ornl.gov/oceans/VOS_Program/) the WAVES Search engine, and the *Mercury* metadata search engine.

CDIAC provides data management support for the **Global CO_2 Time-series and Moorings Project**. The international groups from 18 countries have mounted sensors on moored buoys to provide high-resolution time-series measurements of atmospheric boundary layer and surface ocean CO_2 partial pressure (pCO2). The CO_2 Time-series and Moorings Project is also coordinated by IOCCP. The high-quality measurements of carbon-related parameters from the Moorings are available via CDIAC Time-series and Moorings Project Web site (http://cdiac.ornl.gov/oceans/Moorings/).

CDIAC provides data management support for the **Global Coastal Carbon Data Project**. The coastal regions data are very important for the understanding of carbon cycle on the continental margins. The Coastal Project data include the bottle (discrete) and surface (underway) carbon-related measurements from coastal research cruises, the data from time series cruises and coastal moorings. The data from US East Coast, US West Coast, and European Coastal area are available from CDIAC Global Coastal Carbon Data Project Web site (http://cdiac.ornl.gov/oceans/Coastal/).

GLobal Ocean Data Analysis Project (GLODAP - http://cdiac.ornl.gov/oceans/glodap/) published at CDIAC in 2004. GLODAP is a cooperative effort of investigators funded for synthesis and modeling projects through the National Oceanic and Atmospheric Administration (NOAA), DOE, and the National Science Foundation (NSF). Cruises conducted as part of the WOCE, JGOFS, and the NOAA Ocean-Atmosphere Carbon Exchange Study (OACES) over the decade of the 90s have generated oceanographic data of unparalleled quality and quantity. As of today, the GLODAP database consists of data from 122 WOCE, JGOFS, and other International and Historical Cruises. The **GLODAPv2** database will be published at CDIAC in 2014 and will consist of data from GLODAP+CARINA+PACIFICA+NEW CRUISES with total of ~ 650 cruise data.

CARbon IN the Atlantic ocean (CARINA) database was published at CDIAC in 2009 and consists of data from 188 cruises. CARINA started as the International Project, which emerged from a workshop on “CO_2 in the North Atlantic Ocean,” held in June 1999 in Delmenhorst, Germany and continued throughout of the EU CARBOOCEAN Project. The data and metadata from the CARINA individual cruises and database are now available from the CDIAC CARINA web site (http://cdiac.ornl.gov/oceans/CARINA/) and from the WAVES Search engine.

PACIFICA Database. The work on the Pacific Ocean Carbon data assembly, data analysis and data synthesis started in 2007 as a part of North Pacific marine Science Organization (PICES)

Carbon & Climate Group (C&CG) project. We planed to work on the PACIFICA database for at least 3 years, as this is completely volunteer work for all members of PICES C&CG. The final database will include all cruises in the Pacific Ocean that were not included in the GLODAP database. The PACIFICA Database will be published at CDIAC in May 2013 and will be available to public via WAVES and Mercury.

LDEO Data Base. Approximately 6.4 million measurements of surface water partial pressure of CO2 obtained over the global oceans during 1957-2011 are listed in the Lamont-Doherty Earth Observatory (LDEO) database, which includes open ocean and coastal water measurements. The data presented in this database include the analyses of partial pressure of CO_2 (pCO_2), sea surface temperature (SST), sea surface salinity (SSS), pressure of the equilibration, and barometric pressure in the outside air from the ship's observation system. The data and metadata from the LDEO Database are now available from the Global Surface pCO_2 (LDEO) Database web page (http://cdiac.ornl.gov/oceans/LDEO_Underway_Database/) and from the WAVES Search System

In 2003, CDIAC implemented an instance of the ***Mercury*** metadata system to standardize and inventory CDIAC's ocean data holdings. The Mercury metadata system was developed by staff in ORNL's Computational Physics and Engineering Division. This catalog of CDIAC ocean holdings may be queried at http://mercury.ornl.gov/ocean/.

In 2007, after two years of internal development, CDIAC implemented the **Web-Accessible Visualization and Extraction System (WAVES)**. This data interface permits users to search CDIAC GLODAP and CARINA discrete data and LDEO Database underway data and couples all standardized metadata from the Mercury system to each individual data set. The interface is available at http://cdiac3.ornl.gov/waves/discrete/ for discrete GLODAP and CARINA data search and at http://cdiac3.ornl.gov/waves/underway/ for LDEO underway data search.

Modular Information Content for Ocean Data Systems

Robert Keeley, none, robertkeeley@rogers.com (Canada)
Reiner Schlitzer, AWI, Reiner.Schlitzer@awi.de (Germany)

Data collectors and archive centres all use different hardware and software to assemble and maintain their data. This is appropriate since operations, including hardware, software and human resources, in the many different agencies are different. But the demand of what information and data to keep has an over arching perspective. Preservation and dissemination of an ocean measurement should be accompanied by certain information about that measurement. In the past, it was common to archive simply the observed values of a variable. As technology advanced there was an increased capability to include archiving of some metadata, such as the instrument used for the measurement. But experience in examining very long time series for climate trends demonstrate that the amount of metadata preserved with past measurements has been lacking fundamental information to separate real from apparent anomalies. As well, the different groups making or archiving the information were storing different information. This paper argues that standardization of information content, applying to both measurements and metadata can be realized through modular constructs. Such constructs can be introduced incrementally, can be implemented using appropriate methods in each agency, can allow flexibility to adapt to new measurement technology, and will vastly improve the understanding of exchanged data files. Some preliminary steps along this path that support these assertions will be discussed.

Initial implementations using the Ocean Data View software will be described. This includes the inclusion of structured or free-text documents as integral part of ODV data collections and the usage of metadata links to these documents (local file reference or Internet URL). Bringing up the cruise report or a document containing calibration or data processing information for a given station is as simple as clicking the respective metadata item. In addition to support for external documents, the new ODV data collection format now supports unlimited numbers of « per station » metadata records. These records may contain manually created free text comments, such as descriptions of data processing procedures, or may contain automatically generated structured logs of data and/or quality flag modifications. The complete set of « per station » metadata records will be easily accessible thereby providing essential provenence and data history information for the data of the station.

FixO³: The next phase of Open Ocean observatory Data Management Harmonisation

Maureen Pagnani, BODC, M.Pagnani@bodc.ac.uk (UK)
Robert Huber, WDC-MARE / PANGAEA, rhuber@uni-bremen.de
Richard Lampitt, NOC, R.Lampitt@noc.ac.uk

Since 2002 there has been a sustained effort, supported as European framework projects, to harmonise both the technology and the data management of Open Ocean fixed observatories run by European nations.

FixO³ is expected to start in Autumn 2013, and for 4 years will coordinate the convergence of data management best practice across a constellation of moorings in the Atlantic, in both hemispheres, and in the Mediterranean. To ensure the continued existence of these unique sources of oceanographic data as sustained observatories it is vital to improve access to the data collected, both in terms of methods of presentation, real-time availability, long-term archiving and quality assurance.

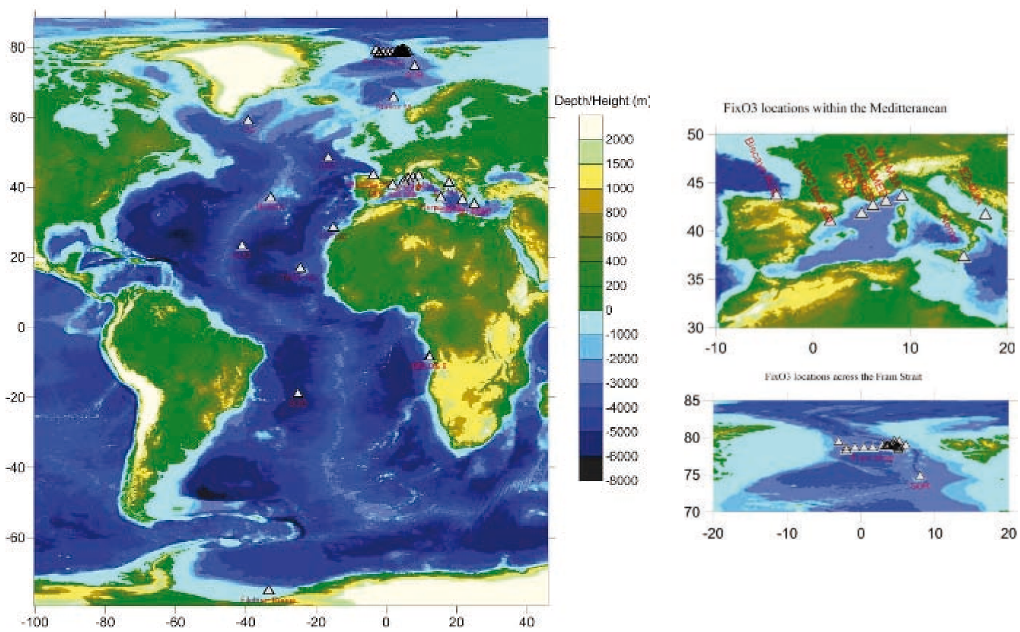


Fig. 1 - FixO³ observatories.

The data management component of FixO³ will improve access to marine observatory data by harmonizing data management standards and workflows covering the complete life cycle of data from real time data acquisition to long-term archiving. Legal and data policy aspects will be examined to identify transnational barriers towards open-access to marine observatory data. To overcome these barriers, it is intended to harmonise data policies and to provide a formal basis for data exchange between $FixO^3$ infrastructures. Presently, the interpretation and implementation of accepted standards has considerable incompatibilities within the observatory community, and these different approaches will will be synthesised into a single $FixO^3$ approach. Further, $FixO^3$ aims to harmonise data management and standardisation efforts with other European and international marine data and observatory infrastructures. The $FixO^3$ synthesis will build on the standards established in other European infrastructures such as EDMONET, SEADATANET, PANGAEA, EuroSITES (European contribution to JCOMMP OceanSITES programme), and MyOcean (the Marine Core Service for GMES) infrastructures as well as relevant international infrastructures and data centres such as the ICOS Ocean Thematic Centre.

The data management efforts are central to $FixO^3$. Combined with the procedural and technological harmonisation, tackled in separate work packages, the $FixO^3$ network of observatories will efficiently and cost effectively provide a consistent resource of quality controlled accessible oceanographic data

Data access services in MyOcean: needs, technical challenges and implementation

Tony Jolibois, CLS, tjolibois@cls.fr (France)
Vincent Claverie, CLS, vclaverie@cls.fr (France)
Thomas Loubrieu, Ifremer, Thomas.Loubrieu@ifremer.fr (France)
Laia Romero, Altamira, laia.romero@altamira-information.com (Spain)
Jon Blower, University of Reading, j.d.blower@reading.ac.uk (United Kingdom)
Rémi De Dianous, CLS, rde-dianous@cls.fr (France)
Guy Griffiths, University of Reading, guy.griffiths@reading.ac.uk (United Kingdom)

Since 2009, MyOcean projects (http://www.myocean.eu) provides an operational service, for forecasts, analysis and expertise on ocean currents, temperature, salinity, sea level, primary ecosystems and ice coverage.

The Central Information System (CIS) of MyOcean manages and gives access to the ocean data products, generated by many Production Centres and distributed by 25 Dissemination Units spread out all over Europe. An objective is to offer the users a single unified service from this highly distributed system. To do so, standardized download and view services have been deployed at Dissemination unit level.

These distributed components are connected together and federated into a seamless reliable service to the user thanks to central components (catalogue, monitoring, user directory) which are presented in the "MyOcean Central Information System".

This presentation focuses on certain needs specified in the MyOcean project regarding these access services (which represent technical challenges) and technical choices made to answer them.

These needs include:

- User needs:
  - Access a standard and harmonized service to reach all types of data (in MyOcean there are gridded data, along-track data and in-situ data)
  - Have accurate and up to date information on the datasets (geospatial coverage, parameters, updated temporal coverage, latest update...)
  - Be able to request the data with selected geospatial and temporal parameters, and download or view only the data needed
  - Be able to download large volumes of data
  - Use robust services, with good performances
  - Connect to the access services of different project with a unique login and password (for secured access services)

- Project needs:
  - Secure download access with authentication and authorization system
  - Have the knowledge of who access what and when, taking into account the fact that users can come from different projects
  - Reduce the time spent to configure and deploy the data distribution systems
  - Comply with Inspire rules*
  - Disseminate netCDF4 data

To answer these needs, the data access services in MyOcean are implemented using international standards, when possible and relevant. It allows answering the INSPIRE directive (notably with ISO and OGC standards). It also ensures the cost-effective and efficient communication involving the geospatial products, and the settlement of the system on well-proven standards and practices. These standards are helping us extend the lives of legacy systems and quickly deploy new capabilities using COTS and mature software.

The tools used to provide the services are:

- Centralized Authentication Service (CAS), which allow to access multiple applications through HTTP protocol while providing credentials only once
- Thredds Data Server, which provides inventory and opendap interface for gridded data
- Oceanotron, which provides inventory and opendap interface for in-situ data
- Motu, which provides a secured and robust interface (connected to CAS) to retrieve data files through a web GUI or through machine to machine interface ; gridded and along-track data are disseminated through this tool
- ncWMS, a tool embedded in the TDS, which provides WMS interface for gridded data
- vsftpd, connected to the MyOcean user directory, so that user can connect the FTP server providing his MyOcean credentials

We'll present in detail these needs and technologies, the high level architecture, and we'll finish by describing the technical challenges that remain.

MEDIN - A UK Partnership for Managing Marine Data and Information

David Cotton, MEDIN, dcott@oceannet.org (UK)
Terry Allen, MEDIN (UK)
Lesley Rickards, BODC (UK)

Introduction

MEDIN (the Marine Environmental Data and Information Network) is a collaborative and open partnership, established in April 2008, working to improve the management of marine data and information, and to provide better access to the UK's marine data resources. MEDIN was established following recognition in the early 2000's that potential data users were finding it difficult, if not impossible, to discover, access and reuse the wide variety of marine data that had been collected by UK organisations.

MEDIN is supported by sponsorship from a range of organisations including government departments, research councils, environmental and conservation agencies, trading funds and commercial organisations. It operates under the auspices of, and reports to, the Marine Science Coordination Committee (MSCC).

Why is MEDIN Needed?

Marine data and information are acquired, maintained and used for a wide variety of different purposes by numerous public and private sector organisations to support their statutory, regulatory, development, commercial and compliance activities.

Common to all these activities is the recognition that good quality comprehensive marine data and information are essential as input to good management and evidence based decision making.

What Does MEDIN Offer?

MEDIN has established a UK-wide operational framework for the management of marine data and information, to provide:

- A single point of access for UK marine data and information – The MEDIN Discovery Portal
- A robust network of definitive integrated Data Archiving Centres (DACs).
- A suite of standards/guidelines for marine data and metadata to ensure discoverability and support data re-use
- Access to authoritative marine reference data - the data sets that are most commonly required to provide context to mapping and analysis

The MEDIN Discovery Portal

The MEDIN Marine Discovery Portal (http://portal.oceannet.org/search/full) aims to provide a single central site to search for all UK marine data, and so avoid the need to search a number of different web locations. It offers geographic and keyword searching, and specific "Regional" views. Coverage, already significant, will increase further as more organisations publish metadata on their holdings.

MEDIN Data Archive Centres

MEDIN has established a network of linked marine Data Archive Centres, each of which has passed an accreditation process, and specialises on a particular theme of data, as follows:

❍ The British Geological Survey – seabed and sub-seabed geology and geophysics
❍ The British Oceanographic Data Centre – Water column oceanography
❍ DASSH – Marine flora, fauna and habitats
❍ The UK Hydrographic Office – Bathymetry.
❍ The Met Office – MetOcean
❍ CEFAS and Marine Scotland – Fisheries Survey Data
❍ Archaeology Data Service – Historic Environment Data

MEDIN Metadata and Data Standards and Guidelines

The adoption of common standards for metadata and data is essential to enable discovery and reuse of data.

MEDIN has established a standard for discovery metadata, which is compliant with all the relevant national (UK Gemini) and international (ISO 19115, INSPIRE) standards, and has developed a number of tools to support the generation of metadata in this format. Support is available through the MEDIN metadata helpdesk.

MEDIN has developed a series of data guidelines, which provide excel format templates and notes, which are built on existing standards and designed to ensure that all necessary auxiliary metadata are recorded. These included a set of guidelines for seabed survey data, which are recommended to the Soundings readers. We welcome feedback on user experience, and are committed to updating the guidelines to ensure they are both practical and effective.

MEDIN for Data Collectors / Managers – What should I do with my Data?

It is important for all members of the marine community to recognise and accept their own responsibilities in terms of putting in place good data management practices.

MEDIN has established a common set of Basic Principles for good data management, which it encourages all in the marine sector to adopt:

• Collect data according to agreed and recognised standards.
❍ Generate and publish metadata in MEDIN Discovery format.
❍ Establish clear conditions for data reuse and access.
❍ Make arrangements for data to be archived within MEDIN accredited data archive centres.

Wider Relevance of MEDIN

Although MEDIN is a national initiative, building a co-operative framework for marine data management and access in the UK, it also has a wider relevance. The problems that MEDIN was set up to address are common and not unique to the UK, including difficulties in identifying, sourcing, and re-using marine data from different sources, where the application of standards is patchy, and the quality of metadata variable. Thus the solution we have developed, in the form of a cooperative framework, with agreed common standards, may also be relevant and applicable elsewhere.

In addition, a key driver for MEDIN at present is to support UK agencies in their implementation of the European Marine Strategy Framework Directive, and the INSPIRE directive. Thus MEDIN is building resources and processes to make metadata and data available according to the agreed specifications and mechanisms. As part of this work, MEDIN's UK partners are involved in relevant key European projects such as SeaDataNet and the EMODNET thematic projects.

Summary

The MEDIN framework acts for the whole UK marine data and information community, public and private sector, delivering a range of practical benefits, including:

- An improved evidence base for decision making and marine planning.
- Efficiency gains in sourcing and ingesting data to meet project aims
- Improved access to data, supporting its re-use and so maximising past investment in data
- Coordination of marine survey and research activities, resulting in efficient use of expensive marine facilities.
- Consistent and clearer terms and conditions for data use, resulting in lower uncertainty, sustainability and more accurate project cost estimates.
- Support to marine public sector information holders in meeting their INSPIRE and Government Transparency agenda obligations.

CTD calibration description using SensorML files

Joan Olive, Marine Technology Unit- CSIC, olive@utm.csic.es (Spain)
Anna Vetrano, ISMAR-CNR, anna.vetrano@sp.ismar.cnr.it (Italy)
Mireno Borghini, ISMAR-CNR, mireno.borghini@sp.ismar.cnr.it (Italy)
Jordi Sorribas, Marine Technology Unit- CSIC, sorribas@utm.csic.es (Spain)
Marc Nokin, IFREMER, marc.nokin@ifremer.fr (France)

Abstract

This paper presents an approach, based on standards and rich semantics, on the use of OGC Sensor Web Enablement SensorML standards to describe instruments and sensors, and in particular to maintain the calibration information related to the CTD (often also CTD-rosette or CTD water sampler). The proposed solution must deal with the inherent difficulty to describe set-up over the time, in a complex instrument that involves different sensors.

Introduction

In general terms, oceanographic instruments are supposed to be calibrated before and after each cruise. Calibration history is a very important information in terms of data quality. The delivery of raw data together with the calibration history of the CTD is a good approach to achieve high data quality levels. All of these calibration events that occur are a relevant part of the sensor description and it's important to have record of them together with the sensors description and the specifications of the sensor calibrations over the time. Depending on the organization, it's a common practice to have all this calibration events explained in different ways, for example, using a textual description (Word file) or a spread-sheet (Excel File).

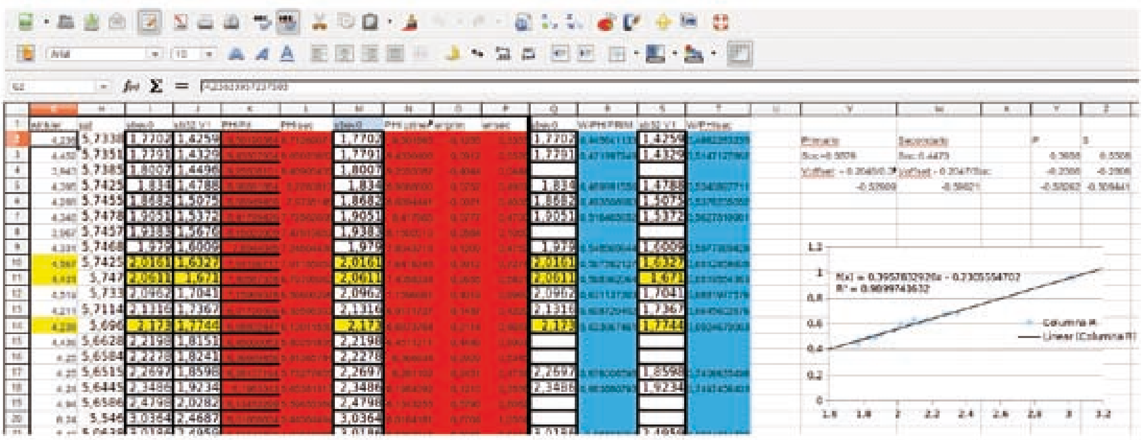


Fig. 1 - Calibration description using a spreadsheet.

Calibration implementation

This work proposes to use the history tag inside of the SensorML files. Using this approach it's possible to use standardized interfaces and mechanisms to retrieve these descriptions and calibrations querying a Sensor Observation Service, SOS.

The history, as can grow quiet large, is not stored directly inside of the SensorML file and is referenced through **xlink:href**.

```
<!--Description of the Conductivuty sensor-->
- <sml:component name="Conductivity Detector">
  - <sml:Component gml:id="CONDUCTIVITY_DETECTOR">
    - <gml:description>
        Conductivity detector connected to the SBE37SMP Recorder
      </gml:description>
      <!--History of the Conductivity Sensor Calibrations-->
    - <sml:history xlink:title="hystory of conductivity sensor calibrations">
      - <sml:EventList>
          <sml:member xlink:href="./historyCTD_0682_Conductivity.xml"/>
        </sml:EventList>
      </sml:history>
```

Fig. 2 - Xlink to calibration external file.

This approach is applied to describe the calibration history of the CTD conductivity, oxygen and temperature sensors.

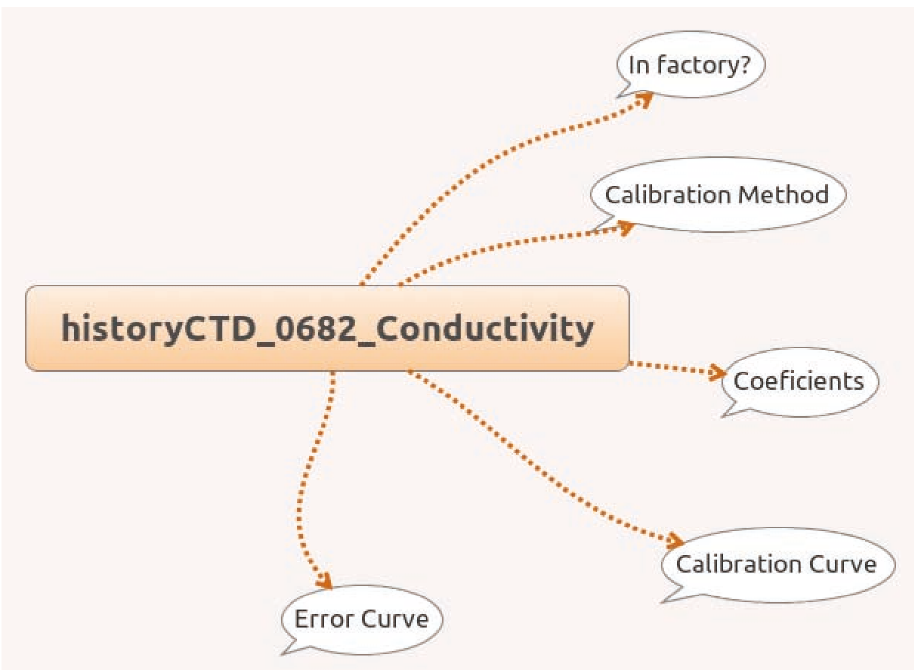


Fig. 3 - Conductivity calibration elements.

To have a complete description of the CTD calibration it's possible to include these five elements (see Fig. 3):

- In factory: to expose if the calibration have been taken place in factory or not. The possible values are true or false.
- Calibration Curve: specification of the calibration curve used to do the calibration:
- Conductivity = $(af + bf^2 + c + dt) / [10(1 + ep)]$
- Calibration Method: description of the calibration method used. It's possible to include an inline description of the calibration or a xlink to an external file with this information: <gml:description xlink:href="http://www.seabird.com/application_notes/AN10.htm"/>
- Coefficients: specification of the calibration coefficients used for the calibration.
- Error curve: description of the calibration method used. It's possible to include an inline description of the calibration or a xlink to an external file with this information: <gml:description>residual=instrument_conductivity-bath_conductivity </gml:description>

When a client query the SOS to retrieve a specific sensor description, the petition must specify the period of time of this description. This petition can query for the description of the sensor and the calibration valid for a specific "Time instant". In this case the SOS retrieves to the client only one history calibration property. If the client asks for the description of the sensor and the calibration valid for a range of time (Time period), the SOS will retrieve the different calibration properties that are valid for this period of time, that can be one or more.

References

[1] OGC 10-037, OGC® Sensor Observation Service Implementation Standard.

[2] Arne Bröring, Christoph Stasch, Johannes Echterhoff. "SOS_2.0_final_draft". (OGC 10-037). Open Geospatial Consortium Inc. 2012-01-06.

[3] Mike Botts, "SensorML for In-situ and Remote Sensors". (OGC 04-019r2)2004-11-01.

ICES Data portal, a single entry point to marine knowledge

Carlos Pinto, ICES, carlos@ices.dk (Denmark)
Neil Holdsworth, ICES, neil.holdsworth@ices.dk

The ICES data portal (ecosystemdata.ices.dk) is designed to become a single entry point to marine data managed by ICES. By providing better access to marine data derived from observations, surveys or samples from all the georeferenced databases that coexist in ICES, all end-users benefit. All data in the portal is free of login restrictions, and the data policy ensures open and free access rights. This is intended to stimulate innovation, encourage the cross-cutting and coordinated activities within marine science, in order to improve our understanding of the marine environment and behaviour of the sea.

Fig. 1 - Current data inventory of the ICES data portal.

Why build a single entry point for the ICES network and marine community?

ICES has invested its efforts in the last years to develop the data portal and it has been committed to facilitate the inclusion of more datasets. This has helped raise the profile and relative importance of data, and data management activities, to data providers and users within the ICES community. The ICES community now have easier access to data by having a single place to see the inventory of data, view and check the data on a map, download the data or the layers in standard formats for their work, studies or layers.

The data portal makes it easier to know what kind of data ICES manages, from when and where, and to get a regional or sub-regional overview across a data theme. Data sets can become more coherent when they can be compared with others. ICES portal facilitates this, and allows scientific assessments to use one or more datasets. This can make them more consistent allowing for better international coordination to identify data gaps, and promote harmonization of methods and terms.

It provides the knowledge and facilitates the growth of a sustainable process that increases the interest and use of marine data..

Disparate datasets harmonised in the same portal, map, format and download

One of the biggest tasks was to design a single home for all the marine data without losing the granularity that primary data offers. Data are assembled from disparate data sets and collated in the ICES data portal form distinct source databases. The datasets currently in the ICES Data portal are:

- Biological Community (34 years of data)
- Contaminants and biological effects (36 years of data)
- Eggs and Larvae (87 years of data)
- Fish predation or stomach contents (historical dataset with 12 years of data)
- Fish trawl survey (49 years of data)
- ICES Historical plankton (11 years of data)
- Oceanographic data (124 years of data)

All these datasets are overlayed onto the blue marble layer that represents bathymetry (Fig. 2). Users can also enable the following layers (ICES EcoRegions, ICES Areas, OSPAR Regions and HELCOM – Baltic Sea sub-basins). All the data in the database are compatible and comparable and can be filtered by geographical areas such as : ICES Areas, OSPAR regions, HELCOM – Baltic Sea Sub-basins. The user has the option of dowloading the data in a standard format i,e, text file or in a shapefile.

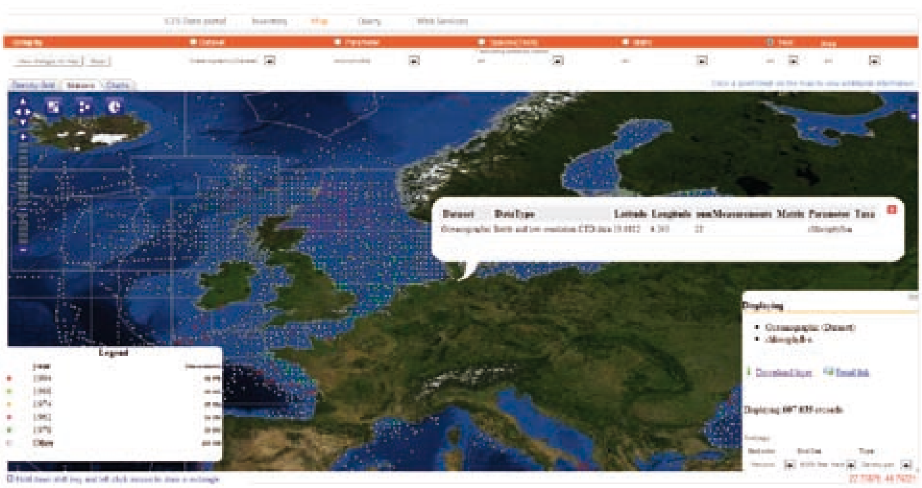


Fig. 2 - Spatial distribution of measurements of chorophyll-a per year shown in the Data Portal Map.

A Service for Statistical Analysis of Marine Data in a Distributed e-Infrastructure

Gianpaolo Coro, ISTI-CNR, gianpaolo.coro@isti.cnr.it (Italy)
Antonio Gioia, ISTI-CNR, antonio.gioia@isti.cnr.it
Pasquale Pagano, ISTI-CNR, pasquale.pagano@isti.cnr.it
Leonardo Candela, ISTI-CNR, leonardo.candela@isti.cnr.it

In computational statistics, the continuous design of algorithms for implementing statistical analysis procedures is producing rather a lot of specialized statistical software applications. At the same time, domain experts need environments in which they can easily manipulate data and run experiments by themselves. This is the case of computational biology, in which data mining procedures are usually applied to big datasets. Such procedures can address several problems, ranging from ecological modelling to vessels activity monitoring. In ecological modelling, the algorithms try to model complex phenomena, in order, for example, to predict the impact of climate changes on biodiversity, prevent the spread of invasive species, help in conservation planning, guide field survey and estimate the geographical distribution of species[1]. Vessel monitoring systems, instead, use data mining to monitor and control fishing activity in the oceans[2]. In such domains, distributed e-infrastructures have been indicated as possible enabling layers for scientists who want to share data, results and experiments with other remote collaborators[3]. Among the others, D4Science is a distributed e-infrastructure that aims at supporting large-scale resource sharing (both hardware and software) and allows data to be processed with distributed computing. D4Science allows also for the creation of Virtual Research Environments (VREs), web-based cooperation environments equipped with all the resources needed to accomplish a scientific investigation.

In this paper we introduce a specific workbench software, named Statistical Manager (SM), that aids in the application of statistical computing and data mining to a variety of biological and marine related problems. The integration with D4Science allows for distributing it "as a Service". The SM is able to exploit the heterogeneous resources offered by the e-Infrastructure to both retrieve and store data. Furthermore, it relies on the D4Science computational resources to run processes on large datasets. The design follows very closely the standard OGC WPS specifications[4]. The differences with

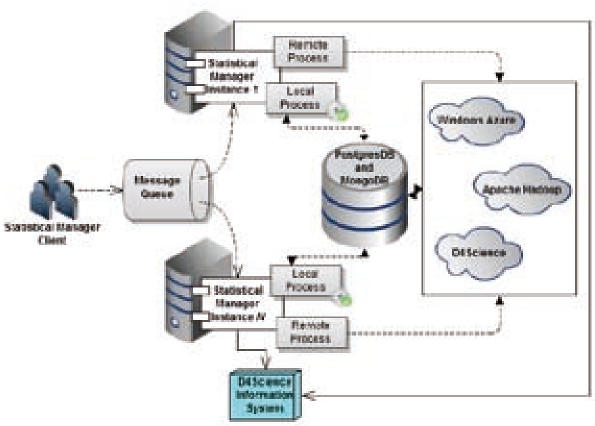


Fig. 1 - Architecture of the D4Science Statistical Manager.

respect to other WPS services[5] is that SM is much flexible in terms of algorithms *plugability* and in the *dynamism* of the web interface, which allows users to execute the experiments. The service interface defines five base operations that can be requested by a client via SOAP protocol:

1. **GetCapabilities** - Allows a client to request and receive service metadata (Capabilities), which describe the algorithms supplied by the service. Such operation provides the name and the general description of each process;
2. **GetProcessDescription** - Allows a client to request and receive detailed information about the processes that can run on the service instances. Information includes the inputs and the outputs. The formats are formally defined on the basis of a proprietary xsd schema;

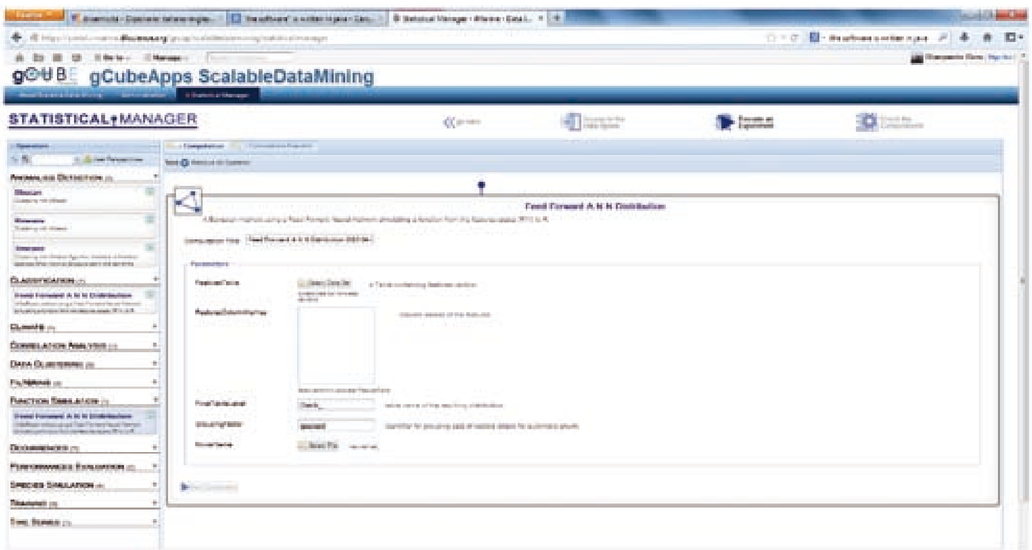


Fig. 2 - Web interface of the Statistical Manager. The left side shows the grouped list of available algorithms. The right panel reports the interface of a Feed Forward Artificial Neural Network.

3. **Execute** - Allows a client to run a process included among the Capabilities;
4. **GetProcessStatus** - Allows a client to check the status of a computation as completion percentage;
5. **GetProcessOutput** - Allows a client to retrieve the process output.

SM uses a software plugin manager to dynamically add new capabilities and algorithms. A web GUI helps users in the setup of the experiments. The GUI is automatically generated on the basis of the definition of the inputs and outputs of the procedures. The software is written in Java and distributed under the open source EUPL license. A software developer who wanted to contribute to extend the service capabilities would have to develop an algorithm implementing some basic interface, which defines its inputs and outputs. The SM will automatically import the new component, publish it among the SM Capabilities and produce the GUI. By extending some specific SM Java interfaces, the developers can use the D4Science cloud computing facilities which run algorithms with a Map-Reduce approach. By using the general D4Science Java libraries it is possible to retrieve, in simple way, the biological datasets which are hosted by the infrastructure in several formats. It is possible to retrieve information also from external biodiversity data providers like GBIF[6] or OBIS[7]. Geographical data are accessible by means of a

GIS connector which communicates with a GeoNetwork[8] instance that indexes and registers the GIS layers hosted by the infrastructure.

Fig. 1 shows the architecture of the Statistical Manager. The requests coming from the clients are distributed among the available services, which have the same Capabilities. The first service that is available in terms of computational resources accepts the request and notifies its status to the D4Science Information System. This is responsible for registering all the resources of the infrastructure, including the computational ones. The SM executes the algorithm, which can exploit the local computational resources as well as the D4Science cloud capabilities or some other external computing facility. The outputs of the algorithms are stored as tables or files. In the former case, the SM relies on a PostgreSQL database shared among the services. In the latter case, the files are stored on a distributed storage system based on MongoDB[9], which ensures high availability and scalability. In both the cases the results are registered in the infrastructure as shareable resources.

Fig. 2 shows the web GUI of the interface. On the right side the list of capabilities is presented, which are grouped according to a user's perspective. For each capability, the **GetProcessDescription** is called in order to get the types of the inputs and outputs. In the right panel, the GUI generates the required fields and displays them with the proper format. In the figure, a Feed Forward Artificial Neural Network (FFANN) is asked to process some tabular data. The *StartComputation* button sends the request to the available Statistical Manager instances in D4Science.

Currently, SM supplies 30 algorithms which include: real valued features clusterers, simulations of generic functions and of climate scenarios, niche models, evaluators of models performances, analysis of time series of catch statistics, analysis of occurrence data about marine species and of vessels transmitted information. The Statistical Manager is currently used by the i-Marine European project[10] and the interface is accessible through the related web portal. We are currently working to include compliancy with WPS specifications and to integrate R scripts.

References

[1] Richard G. Pearson. Species distribution modeling for conservation educators and practitioners. 2012. Synthesis. American Museum of Natural History. Available at http://ncep.amnh.org.

[2] G. Coro, L. Fortunati, P.Pagano. Deriving Fishing Monthly Effort and Caught Species from Vessel Trajectories, 2012. TBP in Oceans 2013 conference, Proceedings of MTS/IEEE.

[3] M. Kia, E. Shayan, and F. Ghotb. The importance of information technology in port terminal operations. vol. 30 (3), pp. 331 – 344, 2000. J. M. Davis. Monitoring control and surveillance and vessel monitoring system requirements to combat iuu fishing, pp. 15–19, 2000. [Online]. Available: http://www.fao.org/docrep/005/Y3274E/y3274e0g.htm

[4] WPS specifications 1.0.0. http://portal.opengeospatial.org

[5] 52North WPS Service. http://52north.org/communities/geoprocessing/wps/

[6] GBIF - Global Biodiversity Information Facility. www.gbif.org

[7] OBIS – Ocean Biogeographic Information System. www.iobis.org

[8] GeoNetwork Opensource. A catalog application to manage spatially referenced resources. http://geonetwork-opensource.org/

[9] Mongo DB. An open source document database, and the leading NoSQL database. www.mongodb.org

[10] The i-Marine European Project, 2011. www.i-marine.eu.

The HNODC Data & Information Management Services: Description & Recent Upgrades

Sissy Iona, Hellenic Centre for Marine Research-Hellenic National Oceanographic Data Centre(HCMR-HNODC) , sissy@hnodc.hcmr.gr (Greece)
Stavroula Balopoulou, HCMR-HNODC, smpalop@hcmr.gr
Pelopidas Karagevrekis, HCMR-HNODC, pkar@hcmr.gr
Angelo Lykiardopoulos, HCMR-HNODC, angelo@hcmr.gr

The Hellenic National Oceanographic Data Centre is managing a great diversity of different data types that come from its participation in a numerous of National, European and International. For the efficient management and exploitation of these data, a multidisciplinary oceanographic database has been developed using a relational database management system. All the information related to the data such as oceanographic cruises, descriptions of environmental datasets, description of research projects and monitoring systems are also organized in relational databases compliant with international established standards. In order to improve its services for dynamic on-line access of the environmental data and their information and by exploiting the increasing capabilities of the Geospatial Web Services, HNODC has built a new integrated management system that follows the basic concepts of a Service Oriented Architecture (SOA). The main features of the system's three layers architecture can be denoted as follows:

- At the Back end layer Open Source PostgreSQL DBMS (and not only..) stands as the data storage mechanism with more than one Data Base Schemas. UMN Map Server and Geoserver are the mechanisms for:
  - representing geographically geospatial data and metadata information via the Web Map Service (WMS),
  - querying and navigating Geospatial and Meta data via the Web Feature Service (WFS),
  - transacting and processing new or existing Geospatial Data via the Web Processing Service (WPS), in the near future.
- *WhereGroup's MapBender*, a geospatial portal site management software for OGC and OWS architectures, acts as the integration module between the Geospatial Mechanisms.
- Apache and Tomcat stand as the Web Service middle Layers. Apache Axis2 with its embedded implementation of the SOAP protocol acts as the Non-spatial data Mechanism of Web Services *(These modules of the platform are still under development but their implementation will be completed in the near future)*.
- In the presentation layer a Web user Interface has been

Fig. 1 - Web GIS Interface.

developed based on enhanced and customized version of a Mapbender GUI. This application is an online Search, Map and download service. allows searching, visualizing and downloading data from the Hellenic oceanographic data base. (A data shopping basket is now under development)

One latest progress is that HNODC database is hosting data from the Poseidon Monitoring, Forecasting, and Information System for the Greek seas. (currently two Poseidon stations are in operation and a third one is expected to be put in service very soon). In addition, within the SeaDataNet Project, the Poseidon CDIs and EDIOS descriptions have been updated and linked with EMODnet -Physical Parameters web portal).

The next step planned for the near future is to upgrade the HNODC database in order to be INSPIRE compliant.

A new updated web site is also under development. Visit our portal at http://hnodc.hcmr.gr

STOQS: The Spatial Temporal Oceanographic Query System

CMike McCann, Monterey Bay Aquarium Research Institute, mccann@mbari.org (USA)
Julio Harvey, Monterey Bay Aquarium Research Institute, jharvey@mbari.org (USA)

Motivation

With increased ability to acquire measurements from various platforms such as ships, moorings, drifters, gliders and autonomous underwater vehicles, the need to efficiently access and visualize multi-parameter spatial temporal data is growing. The Monterey Bay Aquarium Research Institute has designed and built the Spatial Temporal Oceanographic Query System (STOQS) specifically to address this issue.

The fundamental issue of providing efficient management and access to multidisciplinary data is addressed by embracing existing standards and employing geospatial relational database technology along with modern web frameworks to build a tool that enables deep exploration of complex data sets.

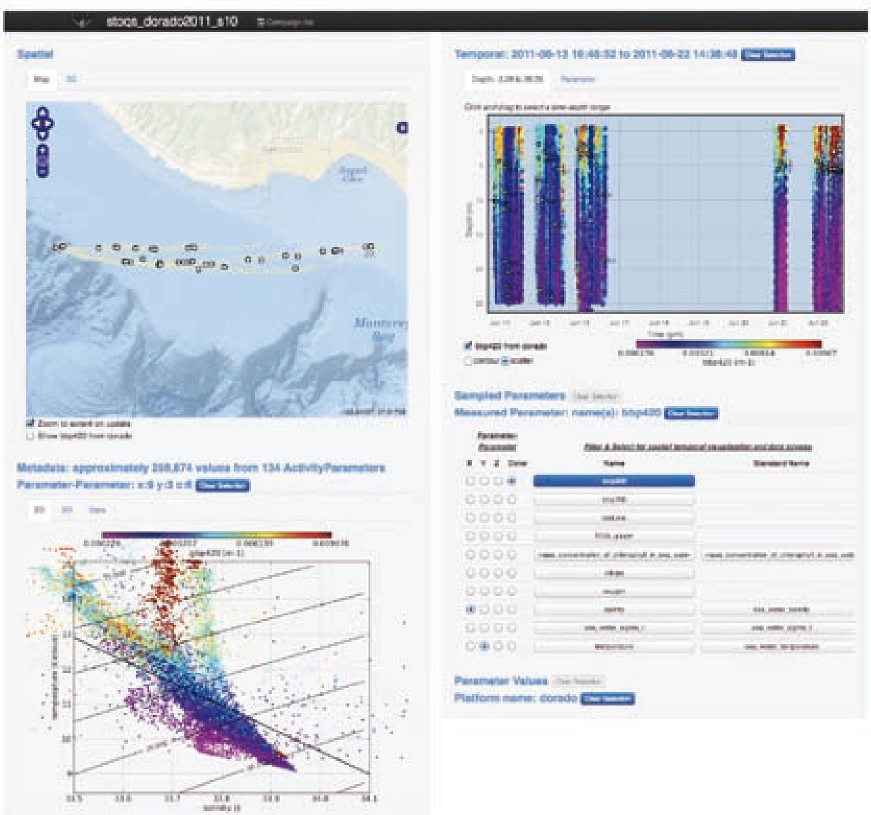


Fig. 1 - STOQS web user interface.

Technology

STOQS consists of a PostgreSQL/PostGIS database, Mapserver, and Python-Django running on a server and client-side technology (jQuery, OpenLayers, Twitter Bootstrap) running in a modern web browser.

The web application provides faceted search capabilities allowing a user to quickly drill into data of interest. Data selection can be constrained by spatial, temporal, and depth selections as well as by parameter values and platform names. The web application layer also provides a REST (Representational State Transfer) Application Programming Interface allowing tools such as the Matlab stoqstoolbox to retrieve data directly from the database. New capabilities provided by X3DOM are being explored for providing interactive 3D views of the data in browsers that support WebGL.

Operation

The STOQS software and its prerequisites are free to use under standard open source licenses. A capable Unix system administrator can have a system up and running within a day. The specific steps of operation include:

1. Install the STOQS software on a Unix server (tested on CentOS, Red Hat and Ubuntu)
2. Conduct oceanographic missions that produce in situ measurement data
3. Create files of the data using the CF-NetCDF conventions for point observation data
4. Make those files available through an OPeNDAP server and create a PostgreSQL database
5. Construct and execute a load script that loads data from OPeNDAP into the database
6. Explore and access the data through the STOQS web user interface

Data Access

Within the user interface measurement data are retrieved directly from the database and transformed to products and formats for further visualization and analysis. A typical campaign produces dozens of NetCDF files containing millions of measurement values. Performing comparable data access without STOQS would require reading all of the NetCDF files, sub-selecting the data and creating desired products within the client tool, an intractable proposition for certain queries.

Data Exploration

With data access and subselection performed on a high performance data server (rather than within the client software), new ways of exploring relationships amongst the data are enabled. A typical oceanographic campaign produces a few million *in situ* electronic measurements and about 100 water samples.

The water samples are subsampled and subjected to further examination ranging from simple chemical analysis and microscopy to more complex genetic assays. Numeric values resulting from these analyses can be loaded and linked to electronic sensor measuement data, permitting easy comparisons and exploration.

The web application allows immediate association of any parameter with any other parameter in a visual way such that an investigator can quickly form hypotheses and test them by making appropriate selections and associations.

Acknowledgements. Development of STOQS has been supported by the David and Lucile Packard Foundation at the Monterey Bay Aquarium Research Institute. STOQS is an open source software project built upon a framework of free and open source software and is available for anyone to use. For more information please see: http://code.google.com/p/stoqs/

ALBATROSS: Retrieval, processing and visualization of INIDEP's oceanographic data and products

Joaquín Gatti, Universidad Fraternidad de Agrupaciones Santo Tomás de Aquino, UFASTA, gattijoaquin@gmail.com, (Argentina)

Juan Arroqui, UFASTA, juan.arroqui@gmail.com

Santiago Trigo, UFASTA, santiago.trigo@gmail.com

Ana Baldoni, INIDEP, baldoni@inidep.edu.ar

Graciela Molinari, INIDEP, gnmo@inidep.edu.ar

Raúl Guerrero, INIDEP, raul.guerrero@inidep.edu.ar

Since its foundation in 1977, INIDEP (Institute for Research and Fishery Development – Mar del Plata, Argentina) through the Physical Oceanography Department (POD) has made a significant effort measuring physical variables and developing procedures to automate data entry, post-processing and quality control. POD implements a standard protocol in the acquisition and processing of data throughout SIAVO, an application that compiles and safewards the oceanographic data, synchronizes land and on board terminals and performs header quality control tests according the IOC – IODE – GETADE recommendations. Header QC procedures includes profile identification (ship, instrument,etc), date and time, localization (ship velocity, track visual inspection), max profile depth vs sonic depth, global and regional ranges for all header variables (lat, long, depth, meteorological data).

In this routine group we added increasing depth control and the identification of missing values with the interpolation option. Delayed QC procedures are applied over individual profiles using the Quality Control Editor (QCed) developed by the Global Temperature Salinity Profile Project (GTSPP- NOAA- NODC). Nowadays 30.680 quality controlled oceanographic stations are stored and managed in a MS SQL-Server database, the INIDEP Regional Oceanographic Data Base (BaRDO). At the present time INIDEP's database is accessed via SQL scripts or using different queries and visualization tools that require a subset of BaRDO data as input.

In this paper we present Albatross, a new tool that provides agile interaction with the stored data and fluent connection with the database.

The developed system aims to systematize queries, providing visual tools to facilitate the data analysis and enable a web access to data and products. Albatross access directly to BaRDO extracting the stations geospatially selected by the user. Selection is made using several graphic tools associated to the stations map. Each tool allows the user to select geographical areas through a polygon, a polygonal line (that includes the stations within a user defined width) or define a regular grid in which every cell contains statistics calculations of the selected stations, all in the same map.

The system also permits the addition of logic filters to any database attribute or parameter, allowing the user to constrain the dataset visualized as needed. In the same way, the system admits categorization rules providing a coloured distinction of the returned data sub-set.

The data is graphically shown as the result of the selected area in the map, allowing a simple interaction between users and the station's specific information that includes max-min, averages, and deviations of each station measurements.

Different plots can be generated depending on the user's selection on the map: T/S plot with isopycnals viewing; Volumetric T/S, histograms, scatter and line plots. Data and products can be exported as a CSV and in case of charts these can also be downloaded as a PNG, JPG or SVG file. The system has a web client - server architecture.

The backend was developed entirely in PHP, which is necessary for the communication of the application with the database engine, as well as to process the retrieved data and its subsequent display on the client. MS-SQL 2005, and Spatial Index library are used to collect oceanographic stations, allowing geographical information filtering. The library implements HTM algorithms (John Hopkins University), that using scalar functions allow the stations indexation stored in BaRDO for later reference.

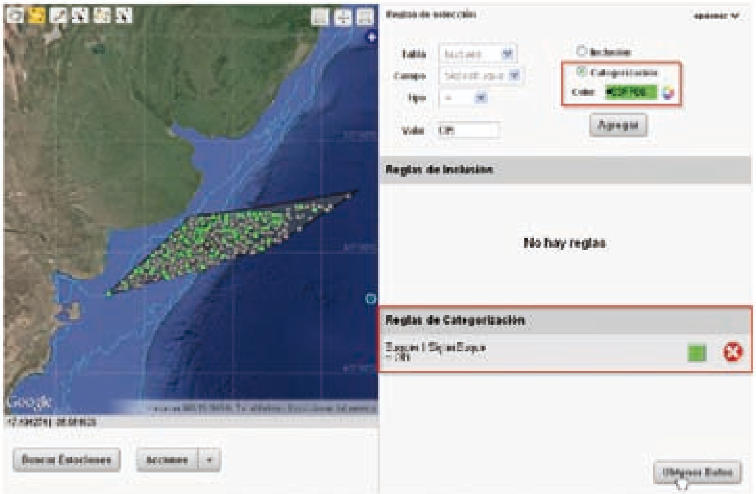
OpenLayers and Highcharts (Javascript based libraries) are the most important tools used since both facilitate interaction with maps and plots. OpenLayers is used for display maps, regions selection, and show the results. Highcharts is the basis for a graphical representation module of the selected data, which allows the definition of plotting scales, limits on both axes and select plot. Maps used can come from various sources (like Google, Bing, etc) as well as maps provided by the library. It is important to remark that the system was built mostly with free languages, except for the database engine which granted access to the existing georeferenced stations.

This project contributes substantially to processing and publication of INIDEP oceanographic data and products over the Argentine Continental Shelf and South Atlantic Ocean.

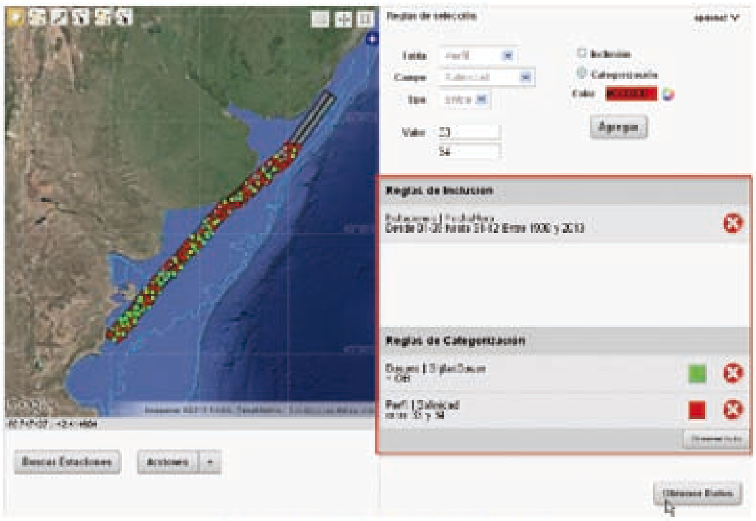
Appendix

The following screenshots show some of the features described in this document.

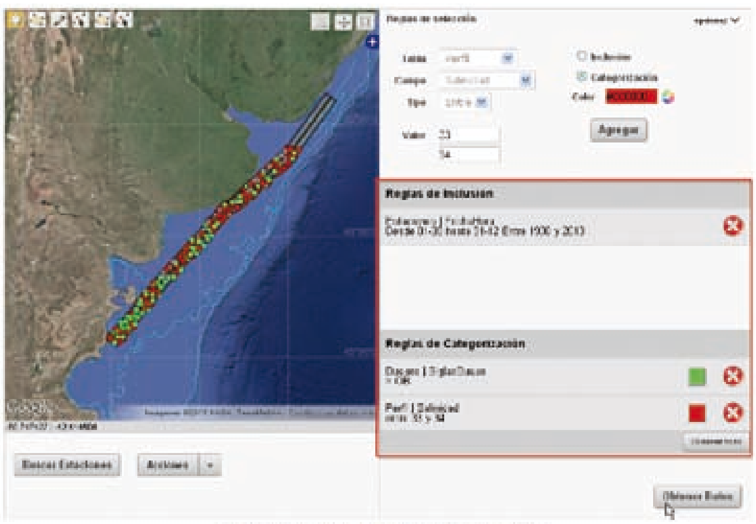
Polygon selection displaying retrieved stations.



Polygon selection with stations data and categorization rules applied.



Polygonal line with multiple rules applied.



Polygonal line with multiple rules applied.

A participatory framework to collect data for understanding the linkages of coastal and marine resources and human use

Marcia Moreno-Báez, Scripps Institution of Oceanography, University of California San Diego, mmorenobaez@ucsd.edu
Catalina López-Sagastegui, University of California Riverside, catalina.lopez-sagastegui@ucr.edu
Jimenez Esquivel, V.M. Centro para la Biodiversidad Marina y la Conservación, A.C., victoria.jimenez@gocmarineprogram.org
Aburto-Oropeza, O., Scripps Institution of Oceanography, University of California San Diego, maburto@ucsd.edu

Cooperative fisheries research provides a mechanism for integrating the unique knowledge, experience, and skills of fishermen and scientists. It is a cooperative effort that begins with the agreement of cooperating in a project and continues until its final stages, with each group having mutual investment in and ownership of the project. Decentralized approaches such as cooperative research mechanisms promotes science communication, transparency, trust and empowerment among fishing communities and can provide much-needed scientifically valid data for fisheries monitoring and management. These programs can enhance the ability to understand the interaction between humans and marine resources, often with impressive success, but have yet to become integrated into mainstream Mexican fisheries management. This paper presents a specific citizen science approach called collaborative research (Miller-Rushing et al., 2012) that incorporates fishers' participation to understand the spatial and temporal dimension of small-scale fishing activities and their interactions with coastal and marine resources using two case studies in Mexico. We incorporate tools that bring a multidisciplinary group of people together in participatory fisheries research. The use of technology and building trust within the work group are key components of this process. The first step is to identify the social and conservation context of the community (Fig. 1). Second, design and implement the participatory data collection process in which local community members become highly involved. Participants are trained through public meetings and workshops to prepare them with necessary skills for the collection and transfer of

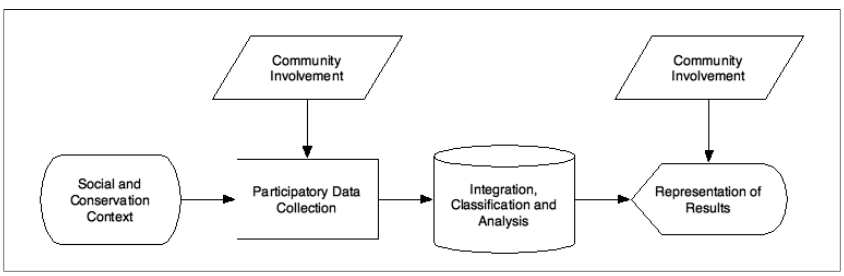


Fig. 1 - Schematic of the participatory fisheries research process.

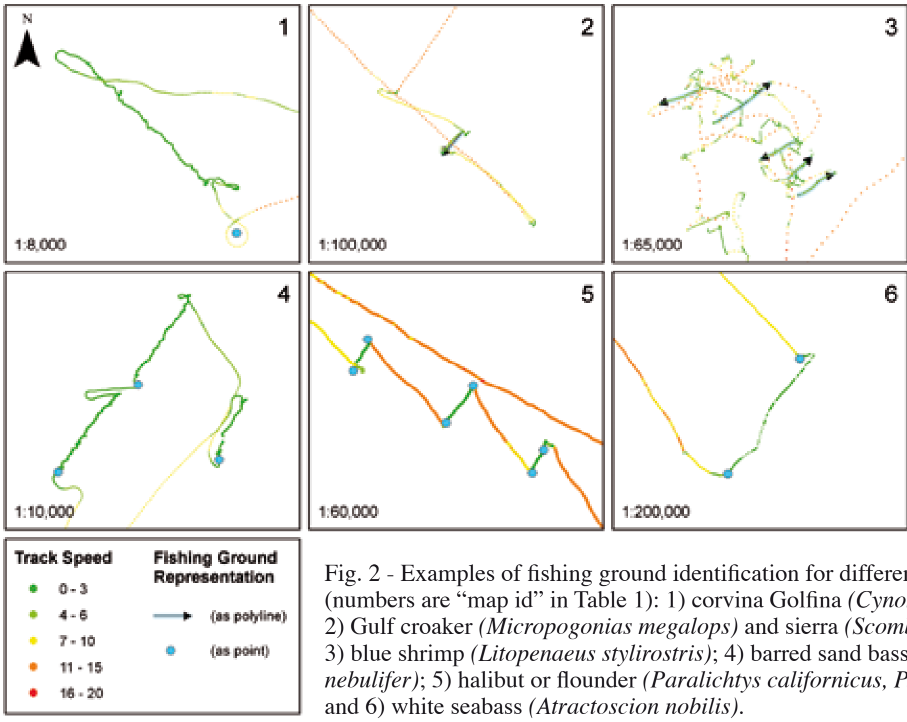


Fig. 2 - Examples of fishing ground identification for different fisheries (numbers are "map id" in Table 1): 1) corvina Golfina *(Cynoscion othonopterus)*; 2) Gulf croaker *(Micropogonias megalops)* and sierra *(Scomberomorus spp.)*; 3) blue shrimp *(Litopenaeus stylirostris)*; 4) barred sand bass *(Paralabrax nebulifer)*; 5) halibut or flounder *(Paralichtys californicus, P. Woolmani)*; and 6) white seabass *(Atractoscion nobilis)*.

data. Third, the integration and analysis of the data into a GIS, which includes four main steps: a) integration of the trackers data extracted from a GPS Data-logger device into a GIS database; b) incorporation of data from the harvest logbooks into the GIS database, c) identification of fishing grounds (Fig. 2); and d) application of spatial analysis techniques. One of the challenges was the development of a technical guide to identify fishing grounds for different fishing methods when data points for fishing grounds where not provided. Finally, the fourth step of the process is choosing the most appropriate format (i.e. maps, graphs) (Fig. 3) for the presentation of results in public meetings in order to involve participants and generate a constructive discussion around various topics including: content, formats, additional audiences, and meaning and interpretation of results. The two case studies highlights the effectiveness of and the essential steps in developing our cooperative fisheries research and monitoring projects.

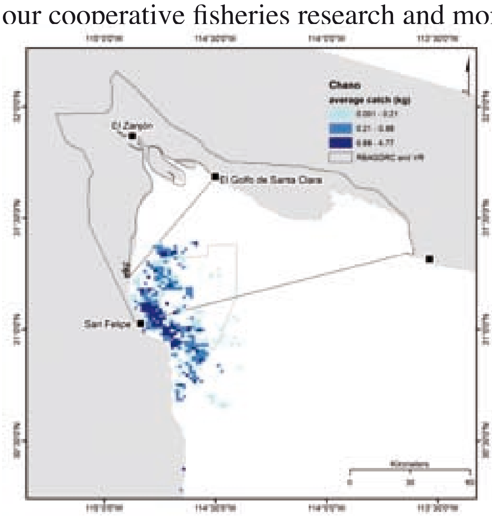


Fig. 3 - An example of a map generated with the data and presented to the community.

A Quality Management Approach to the Delivery of Marine Data and Services

Greg Reed, Australian Ocean Data Centre Joint Facility, greg@metoc.gov.au, (Australia)

Marine data are fundamental to our understanding of the processes that control the environment. These data are a key requirement for effective strategic decision making and play an important role in promoting the development of economic activities and underpin many of our activities. Reliable marine data are key inputs to the efficient management of marine resources and it is essential that accurate data be accessible in a timely manner to facilitate informed decision making. The value of such data increases when they are provided by organizations that have established and adhere to quality management principles.

Quality management is the process for ensuring that all the activities necessary to design, develop, and deliver a product or service meet the requirements of the end-user. Quality management focuses not only on product and service quality, but also on the means to achieve it. Users of marine data, products and services are increasingly calling for quality management systems to be put in place to provide a level of confidence in quality.

The International Oceanographic Data and Information Exchange (IODE) of has recognized the need for a quality management framework to ensure its National Oceanographic Data Centres (NODC) are able to provide quality data and services to meet the needs of a broad community of users and has endorsed a Quality Management Framework (IODE-QMF) to provide the overall strategy, advice and guidance for NODCs to design and implement quality management systems for the successful delivery of marine and related data, products and services. The IODE-QMF addresses the implementation of a quality management system by NODCs that will assure the quality of final data, products and services.

An accreditation process will be implemented whereby NODCs will need to satisfy a minimum set of requirements to ensure compliance with IODE standards and to establish a mechanism to regularly monitor and assess the quality of their data and services. IODE has defined the accreditation criteria to ensure NODCs meet these requirements.

This paper discusses the IODE-QMF which addresses the implementation of quality management systems to ensure NODCs can demonstrate their capabilities to provide data, products and services in compliance with established standards and procedures that will lead to accreditation.

Using GeoServer and spatial database for receiving and analyzing of georeferenced data

Damir Ivankovic, Institute of Oceanography and Fisheries, ivankovic@izor.hr (Croatia)
Ivan Vucic, Institute of Oceanography and Fisheries, vucic@izor.hr
Vlado Dadic, Institute of Oceanography and Fisheries, dadic@izor.hr

In the frame of operational monitoring according to Water Framework Directive benthic macroalgae were used as a biomonitoring tool in the Adriatic Sea using CARLIT method (Cartography of littoral rocky-shore communities). The result of CARLIT method application is assignation of water quality status expresed as EQR value (Ecological Quality Ratio) to coastline sectors and an average EQR to the whole water body. Each segment was predefined, and problem was how to enter acquired data and associate data to particular segment. As EQR have 5 classes (High, Good, Moderate, Poor and Bad) another problem was how to show segment EQR on the map and calculate and show EQR for water body (calculated as average from all segments that are inside in particular water body).

System components

As server platform we choose CentOS 5.8, Linux operating system. For GIS publishing software we choose GeoServer, because of good references and easy implementation. GeoServer is running under Tomcat 5 application server. GeoServer allows users to share and edit geospatial data. Designed for interoperability, it publishes data from any major spatial data source using open standards. GeoServer functions as the reference implementation of the Open Geospatial Consortium Web Feature Service standard, and also implements the Web Map Service and Web Coverage Service specifications. GeoServer is open source free software. As we already use Oracle database on Institute, for storage and spatial analysis we use Oracle database (instead of files). Geoserver also can be used with open source spatial relational databases as is Postrgres database.

Java script libraries

For displaying GIS content, described system use various Java script libraries. Java script libraries are predefined JavaScript functions packed together with visual CSS components for advanced web content presentation. Advantage of this conception is that most of processing power required for user - system interaction comes from client side, JavaScript functions are running inside client web browser. For ours web application two JavaScript libraries were used:

- OpenLayers
- GeoEXT

While OpenLayers is used as base tool for web gis application developement, using of GeoExt allow creation of familiar and easy to use user interface.

Databse storage and data analysis

In the database three tables were used: segments, water_bodies and crossover table segment_water_bodies. In segments table each segment was inserted whit spatial column (line). From data column EQR was calculated for each segment. In the water_body table definitions of water bodies (polygons with holes) with some attribute data were stored. Also column for average EQR was added. At the end using spatial function SDO_ANYINTERACT (segments_geom_column, water_bodies_geom_column) crossover table was populates in the way that for each water_body *id* all segments *id*'s that have any interaction (spatial interaction) with that water body were found. After that founding average EQR for each water body was trivial (group by with crossover table).

Results

On the Fig. 1 resulting map with coloured segments and water bodies according EQR class are shown (http://jadran.izor.hr/geo/eqr.html). All data were in WGS 84 (SRID 8307 in Oracle notation, and EPSG 4326 in GeoServer) datum which made overlapping with GoogleMaps (which is used as base layer) pretty straightforward, since GeoServer does it automatically. Ecological classes have standard representations with colours. For colour representation of attributes values on the map (in this case EQR) we used SLD. The OGC Styled Layer Descriptor (SLD) standard defines a language for expressing styling of geospatial data. GeoServer uses SLD as its primary styling language. In the SLD ranges can be easily defined so each chart element is coloured depending to value of the EQR attribute.

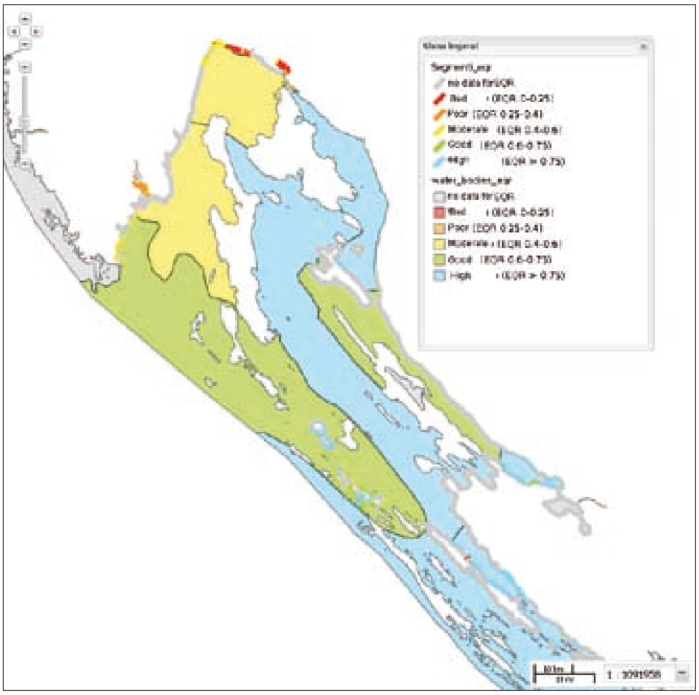


Fig. 1 - Map from GeoServer (web application) showing resulting layers (segments and water bodies) in the north Adriatic Sea http://jadran.izor.hr/geo/eqr.html.

Conclusion

For data insertion and validation is very convenient to have web application because it supports field work and it is basically independent to client side hardware and OS. For georeferenced data maps are almost mandatory and web oriented GIS is native solution. Resulting layers can be shown and overlapped with many other layers and this adds possibilities for additional use of acquired data. Although Geoserver by itself has possibilities for spatial data analyses they become more easialy and quicker in case of using spatial database as data source.

Bridging the gap between marine environmental and molecular data: The Micro B3 Information System

Renzo Kottmann, Max Planck Institute for Marine Microbiology, rkottman@mpi-bremen,de (Germany)
Aleksandar Pop Ristov, InterWorks , aleksandar.popristov@interworks.com.mk
Julia Schnetzer, Max Planck Institute for Marine Microbiology, jschnetz@mpi-bremen,de
Antonio Fernandez-Guerra, Max Planck Institute for Marine Microbiology, afernand@mpi-bremen,de
Frank Oliver Glöckner, Max Planck Institute for Marine Microbiology and Jacobs University Bremen, fog@mpi-bremen.de

Background

Sequencing technologies for molecular analysis of marine organisms are becoming increasingly cheap and powerful. Consequently not only large sequencing centers such as Genoscope or the Wellcome Trust Sanger Institute in Europe expand their capacities but also an increasing number of marine biology laboratories set up their own small- or medium-scale sequencing facilities. This trend establishes molecular sequence data as an important and routine part of the scientific analysis of marine organisms.

Several large-scale marine sampling and sequencing campaigns like the Tara Oceans (oceans.taraexpeditions.org), Malaspina (www.expedicionmalaspina.es) or the upcoming Ocean Sampling Day (www.oceansamplingday.org) lead to a wealth of data which engages the whole marine molecular research community.

Sophisticated data management, including the efficient processing of sequence information and the collation and organization of multidimensional contextual information, will be the key to ensure successful exploitation of these data to their full potential. The open source Micro B3 Information System (MB3-IS) provides numerous components and modules for integrating and managing molecular- and environmental data.

From Sampling and Sequence Analysis to Public Access

The aim of the Micro B3 Information System (MB3-IS) is to provide the bioinformatics software capacity for marine biodiversity data processing, analysis and biotechnological exploitation. It is developed within EU 7FP project Micro B3 (Biodiversity, Bioinformatics, Biotechnology: www.microb3.eu).

It supports users in the complex process of handling the sequence data arising from studies of a range of organisms (protists, viruses and prokaryotes), both in isolation and in mixed organism samples, and a variety of sequencing platforms whilst integrating interpretations with related environmental data. To perform these tasks, MB3-IS consists of three main components (see Fig. 1): 1) The data processing component comprises bioinformatics pipelines which perform tailored, automatic sequence annotation for metagenomes and genomes from protists, viruses and prokaryotes. 2) In the data integration component the quality controlled and processed data will be continually integrated with contextual (meta)data from the environment obtained by on site and remote sensing measurements. The integration of marine environmental data is developed in close collaboration with marine data infrastructures like SeaDataNet, Pangaea and EurOBIS among others. At this point, it may be that the sequence data are submitted to other databases and archives as appropriate, in order to maximize data interoperability.

Additional information from sequence data archives helps finding recurring patterns of genes, metabolic capabilities, as well as to determine the biogeography of organisms with respect to habitat conditions. Overall, the integration component lays the basis for large-scale statistical analysis and modeling in ecosystems biology and it provides candidates for biotechnological applications. 3) The Micro B3 service component which provides end-user tools for community sequence annotation, visualization of large-scale, multidimensional datasets and statistical tools for ecological analysis. Additionally, Micro B3 will utilize social web technologies to share and collaborate on data. All data and services will be published on the web for browsing and for programmatic access via web services. Each component is composed of standards-based, integrated, and modular open source software to reach a maximum of interoperability with existing and future data systems.

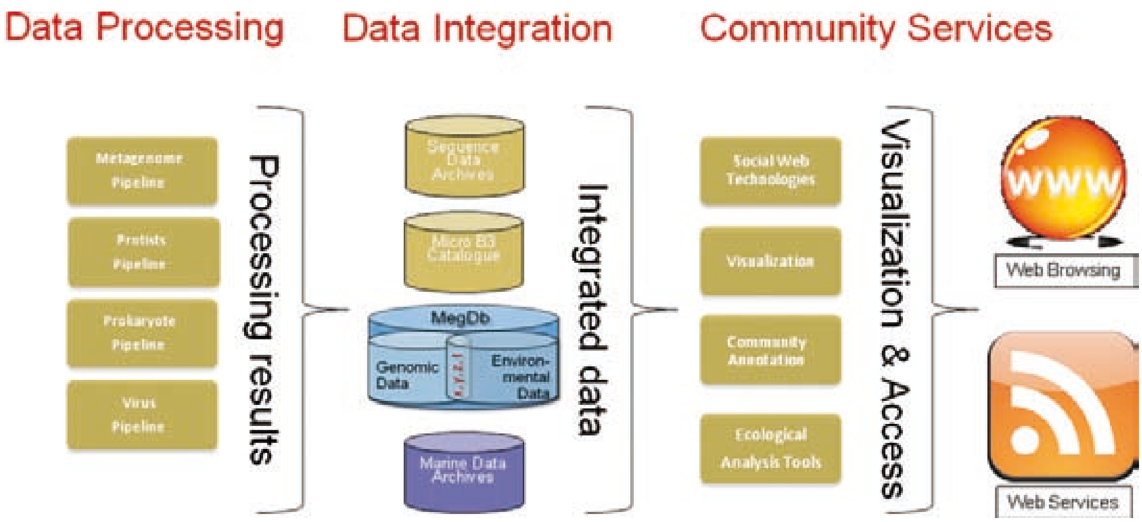


Fig. 1 - High level view on the data flow of the Micro B3 Information System.
New sequence data from marine samples will be processed and analyzed by bioinformatics pipelines.
In the next step the new sequence data will be integrated with existing sequence - and marine environmental data for higher level analysis in order to then provide web browsing and programmatic access to integrated information.

Supporting the Ocean Sampling Day

The Ocean Sampling Day (OSD) is a simultaneous sampling campaign of the world's oceans. It is organized by Micro B3 and will take place on the summer solstice (June 21st) in the year 2014. These cumulative samples, related in time, space and environmental parameters, will provide insights into fundamental rules describing microbial diversity and function and will contribute to the blue economy through the identification of novel, ocean-derived biotechnologies.

It is expected that these data will provide a reference data set for generations of experiments to follow in the coming decade. That renders the OSD as the primary use case which the Micro B3-IS is supporting. Hence, the Micro B3-IS supports the OSD with the Ocean Sampling Day App, a smart phone application for early consistent digital metadata acquisition. It allows users to submit data right from the field to the Micro B3-IS for world-wide near real-time sharing of sampling information. The App is available for Android and iPhone.

Summary

The Micro B3-IS is a modular open source software for the analysis, integration and efficient management of marine environmental as well as marine molecular data from all domains of life. It is based on current state of the art technologies and standards in order to achieve maximum interoperability with other systems. It can be freely downloaded from https://colab.mpi-bremen.de/wiki/x/EwAL.

Acknowledgements. Micro B3 is financially supported by the 7FP Ocean of Tomorrow Grant #287589

Sub-surface Underway Data Validation at IEO Data Center

Andrea Pannocchi, Laboratory of Experimental Oceanology and Marie Ecology
Department of Biological and Ecological Science of Tuscia University, andreapannocchi@gmail.com (Italy)
Lucía Arranz, Spanish Institute of Oceanography (IEO), lucia.arranz@md.ieo.es
Alberto Hernández, Marine Technology Unit (UTM/CSIC),ahernandez@utm.csic.es
Elena Tel, Spanish Institute of Oceanography (IEO), elena.tel@md.ieo.es
Amaia Viloria, Spanish Institute of Oceanography (IEO), amaia.viloria@st.ieo.es

Almost the entire Spanish Research Vessels fleet, together with VOSs (Voluntary Observing Ship) and SOOPs (Ship of Opportunity Program), are equipped with automatic sampling devices able to collect physical, chemical and biological data along the survey routes. On board thermosalinographs (TSG) measure the sub-surface seawater temperature and salinity. Frequently these devices have coupled fluorometers to measure the chl-a fluorescence, for the estimation of phytoplankton biomass.

Data quality control is essential to ensure an international acceptance and to detect deviations and/or errors in the sensor operations, missing information and duplicates. At this work there are presented the procedures to check and validate TSG data coming from Spanish research vessels from two oceanographic institutions: the Spanish Oceanographic Institute (IEO) and the Spanish Marine Technological Unit (UTM/CSIC).

IEO datasets start in 2002 and cover the Spanish coasts and adjacent seas. At the present time, five research vessels (Cornide de Saavedra, J.M. Navaz, J. Rioja, R. Margalef and A. Alvariño) are equipped with TSG. Quality control is realised using the MATLAB tool **procesaTSG**[1] developed at the IEO following GOSUD's[2] (Global Ocean Surface Underway Data) protocols that is continuously improved. These routines include an automated quality control of dates, global and regional positions, temperature and salinity ranges, duplicate control, detection of constant values, and a velocity test[3]. Daily data are published online for operational purposes.

The high facilities of visualization are used to validate the data by checking the vessels tracks, eliminate stabilization time periods and spikes, check gradient variations and compare with average values. The scientific validation is supported by information provided by the water profiles obtained in the IEO monitoring programs and oceanographic cruises.

UTM datasets coming from surveys activities of the Sarmiento de Gamboa and Hespérides research vessels spanning from year 2008 to 2011, undertaken around the Spanish coasts, and along ocean routes (North and South Atlantic, Pacific and Caribbean Sea). Due to the information provided by UTM for each dataset (metadata and ancillary data which includes cruise events, navigation, track map, etc), datasets were checked for validation following this procedure:

1. NEMO[4] software: to obtain a dataset coupled with metadata and for parameter code matching with SDN common vocabularies.
2. ODV[5] software: for pre- and post-control procedures, data representation and quality flag management. ODVs visualization gives a good understanding of the overall trends of the sampled parameter along each track, revealing their latitudinal and circadian gradients.
3. Navigation file: for vessel speed checking (*ad-hoc* Matlab routine), in order to delete all that data corresponding to a vessel velocity less than 1 kn.
4. NOAA World Ocean Database: to verify doubtful data against climatological reference values.

Discussion

As the international scientific community welcome such instrumentation on board[6], it is very important to consider the problems associated with the management of these big amount of data. Datasets are received at the data centre and, following ICES WGMDM Guidelines[7] and GOSUD recommendations, are archived in metadata and data format standards in order to facilitate data interoperability, sharing, and re-using.

In some cases, at the beginning of the cruise or during the oceanographic stations in the research survey, anomalies are found probably due to the sensor stabilization or the overheating of the seawater at the pipe intake. The data visualization permits the detection of those situations (Fig.1, Fig.2). In the same way, plotting the track lines shows errors associated to GPS position data.

Conclusions

Continuous measurements along navigation tracks give valuable information that complete the one obtained in the oceanographic surveys.

The control procedure highlights the convenience of a careful planning for the sampling procedures. Furthermore, in some cases, it should be recommendable to turn off the instrument during a fixed sampling station, especially in estuarine areas. Enclosed maps with the sample stations points are useful to detect these anomalies.

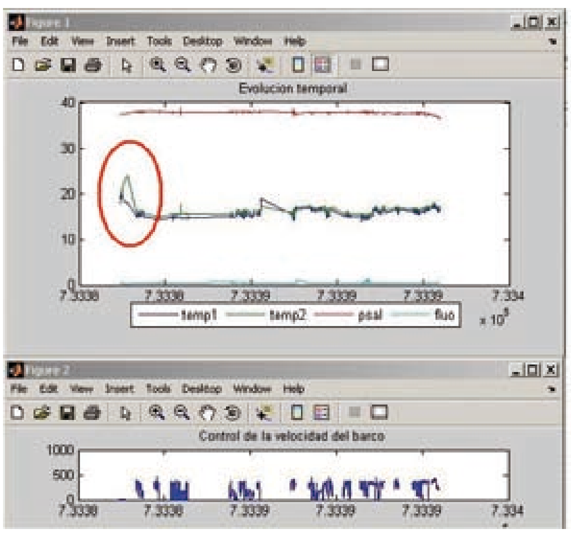


Fig. 1 - Sensor stabilization and vessel velocity, output from ProcesaTSG.

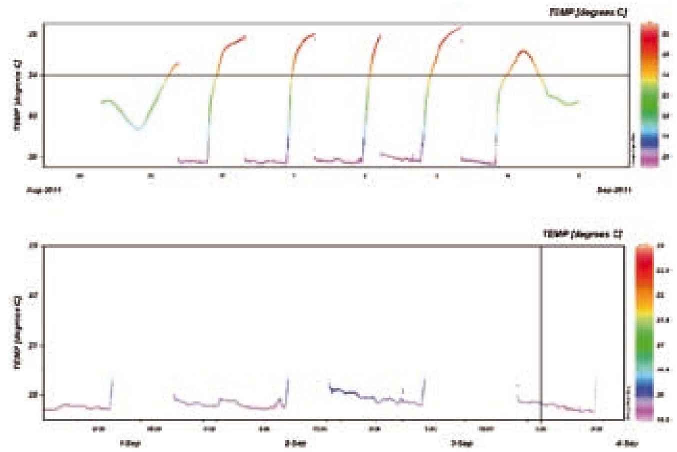


Fig. 2 - Original data with incorrect variation of temperature at stations (above); validated data of temperature only along the track (bottom).

The two procedures showed along this work achieve similar results in systematic errors detection and validation of the final datasets, although the output formats are different: ODV-SDN and MEDATLAS. The assembly of the two dataset after this validation improves the spatial coverage and allows to elaborate spatial distribution maps for the regions with enough data.

References

[1] Tel E. and Arranz L., 2008. ProcesaTSG: Analsiis de Datos de Termosalinógrafo. Aplicación Informática en Matlab. Manual de Usuario. IEO.

UNESCO-IOC 2010. GTSPP Real-Time Quality Control Manual, First Revised Edition.

[2] IOC Manuals and Guides No. 22, Revised Edition. IOC/2010/MG/22Rev.

[3] Viloria A., Tel E., González-Pola C., Merino A., Reguera I., Rodriguez-Puente C. and Lavín A. 2012. 10 years of TSG records on the Cantabrian Coast. Proceedings 7th Spanish-Portuguese Assembly for Geodesy and Geophysics.

[4] Fichaut M. and Bregent S., 2012. NEMO 1.4.4. User's Manual, IFREMER.

[5] Schlitzer R., ODV OceanDataView:. Schlitzer R., Ocean Data View, http://www.awi-bremerhaven.de/GEO/ODV, 2002. Alfred-Wegener-Institute for Polar and Marine Research (AWI).

[6] Marcelli M., Pannocchi A., Piermattei V. and Mainardi U., 2012. New technological developments for oceanographic observations Oceanography, InTech Pubb. ISBN 979-953-307-541-8.

[7] Guidelines for Surface Underway Data. ICES WGMDM (2001).

Using OGC standards (SOS, O&M) to standardize data flow from cabled marine observatories to the PANGAEA data archive

Andree Behnken, MARUM - University of Bremen, Germany, abehnken@marum.de
Robert Huber, MARUM - University of Bremen, Germany, rhuber@uni-bremen.de
Christoph Waldmann, MARUM - University of Bremen, Germany, waldmann@marum.de
Mikhael Kononets, Department of Chemistry, University of Gothenburg, Sweden, m.kononets@chem.gu.se
Anders Tengberg, Department of Chemistry, University of Gothenburg, Sweden, anderste@chem.gu.se
Per Hall, Department of Chemistry, University of Gothenburg, Sweden, perhall@chem.gu.se
Henrik Stahl, Scottish Marine Institute, Scotland, Henrik.Stahl@sams.ac.uk

Overview

Two cabled marine observatories were deployed during the EU FP7 HYPOX project – one in Loch Etive, Scotland, and one in Koljoefjord, Sweden. We will present a data access workflow which was implemented by us, based on a set of OGC standards, namely Sensor Observation Service (SOS) and Observations & Measurements (O&M).

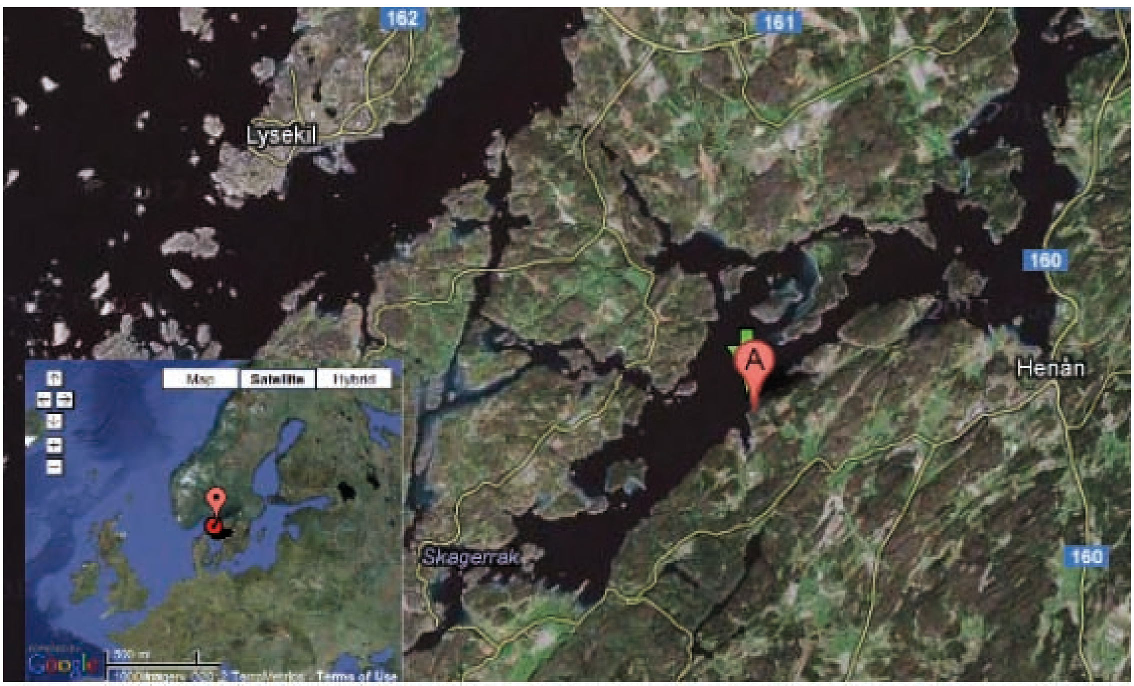


Fig. 1 - Location of the Koljoefjord observatory, Sweden.

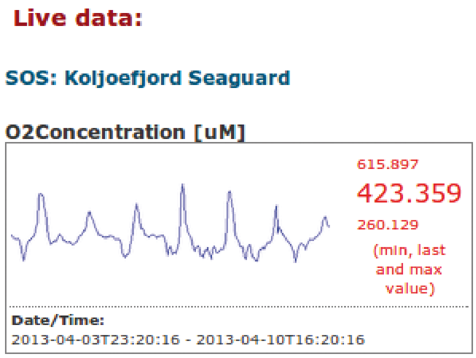


Fig. 2 - Display of near real-time data on the HYPOX portal page.

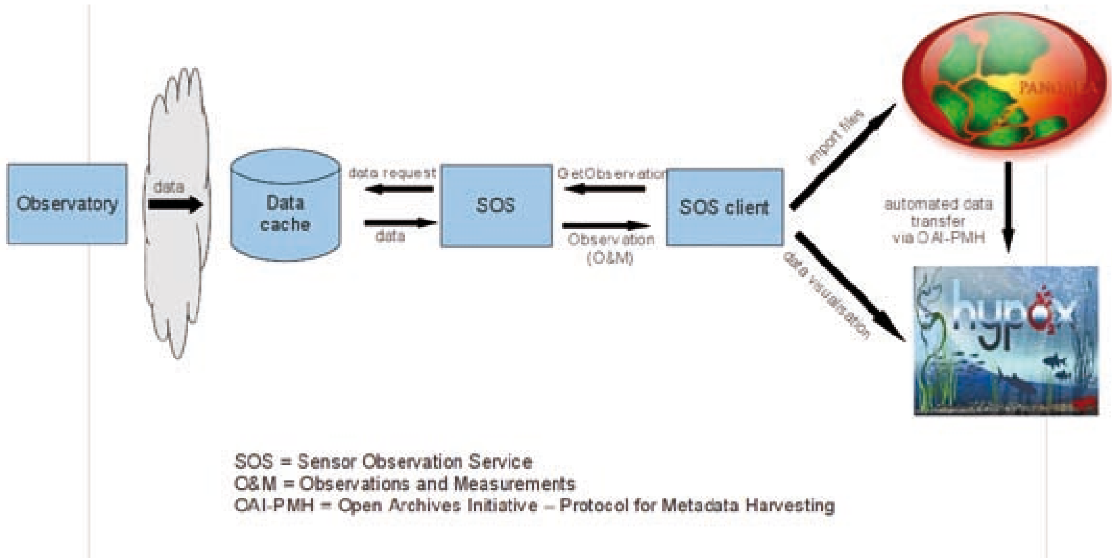


Fig. 3 - Standardized data flow from observatory to PANGAEA and HYPOX portal.

Standardized data transport

Real-time or near real-time data access as well as data retrieval for archiving purposes at PANGAEA for these two observatories has been achieved by implementing a SOS and two SOS clients. The SOS defines a web service interface which allows for querying of metadata, information about observed features and observations (data) in a standardized way. A SOS instance deployed at MARUM is constantly retrieving data from the Koljoefjord observatory. One of the implemented SOS clients is used to request real-time data from the SOS and pass it on to an AJAX web client for display on the HYPOX portal page (http://dataportals.pangaea.de/hypox). The second client can be regarded as a harvesting service, requesting data from the SOS once monthly in order to prepare it for semi-automatized archiving in the PANGAEA long-term data archive. In both scenarios, requested data is streamed conforming to OGC's Observations and Measurements XML implementation standard. Using the described OGC standards allows for managing sensor-/observatory data in an interoperable way. Fig. 3 gives a schematic overview on the described setup.

Monitoring services in the framework of SeaDataNet II Project

Angelos Lykiardopoulos, Hellenic Centre for Marine Research, angelo@hcmr.gr (Greece)
Konstantinos Kalkavouras, Hellenic Centre for Marine Research, ckalkav@hcmr.gr (Greece)
Flavian Gheorghe, Marine Information Service (MARIS), flavian@maris.nl (Netherlands)
Athanasia Iona, Hellenic Centre for Marine Research, sissy@hnodc.hcmr.gr (Greece)

Network monitoring refers to the practice of overseeing the operation of a computer network using specialized management software tools. Network monitoring systems are used to ensure availability and overall performance of computers (hosts) and network services.

Fig. 1 - Screenshot from nagvis visualization tool.

A monitoring system is capable of detecting and reporting failures of devices or connections. It will often send messages over the network to each host to verify it is responsive to requests. When failures, unacceptably slow response, or other unexpected behavior is detected, these systems send additional messages called alerts to designated locations (such as a management server, an email address, or a phone number) to notify system administrators.

The main objective of Network Monitoring development during SeaDataNet II is to deliver a monitoring system, more stable, accurate and capable to monitor the operation of a computer network. The system will ensure the availability of computers (hosts) and network services in the network. It will also be capable to detect and report failures of devices or connections. Even more in case of failures occurrences or other unexpected behaviors, the system will send messages (alerts) to designated locations to notify system administrators.

The appropriate monitoring software was selected after a comparison of available tools on the market. The selected software decided to be Nagios. This is an Open Source software package widely used in Academic and Commercial community.

In order to view the state of the services in real time in a more friendly way a visualization tool developed based on NagVis add-on (Fig. 1).

The following modules are been or have to be developed as main parts of such a system constitution:

- Additional Nagios system plug-ins in order to monitor web forms, web services, Download Manager and Central Authentication Service (CAS).

- Definition of SeaDataNet global availability indicator (formula and weights)
- A Web portal for delivering and sharing monitoring results. The portal will host global statistics and information of all SeaDataNet sources as well as detailed statistics for each separate service which will be accessible only from the administrators of each service.
- A messaging system for notifying the administrators
- Duplication of the whole system to a second physical premise and development of software tools for comparison and merging of the measurements results.

As far a total of 90 services are being monitored:

- 39 SeaDataNet Download Managers
- 7 SeaDataNet core Services
- 25 GeoSeas Download Managers
- 23 UBSS Download Managers

The production of this monitoring system provides the following monitoring benefits:

- Monitoring in real time and alerting when incidents are detected so the administrators are able to correct them as soon as possible;
- In a longer term, identification of critical components within widely distributed systems and to update them to improve their robustness;
- Information for the users and stakeholders of the system on the overall availability of provided services.
- Specially by the production of the messaging system for the administrators the monitoring system will became a “live” component of SeaDataNet acting as valuable tool to improve the overall availability of whole platform

The network monitoring project is now an operational monitoring system which can produce statistics for 90 services. So far the results of the system indicates that SeaDataNet core services availability is up to 99.90% which is enough to conclude that core services are accessible constantly. At the same time on other outcome of monitoring system is the availability of download managers which be considered satisfactory as the most services’ availability is up to 99,5 %.

Progress in developing the ICSU World Data System

Lesley Rickards, PSMSL, Liverpool,UK (ljr@bodc.ac.uk)
Jean-Bernard Minster, Chair WDS Scientific Committee, Scripps Institution of Oceanography, USA
Mustapha Mokrane, WDS International Programme Office, NICT, Tokyo, Japan

On behalf of the WDS Scientific Committee and the WDS International Programme Office

The World Data System (WDS) was created by the 29th General Assembly of the International Council for Science (ICSU) and builds on the 50 year legacy of the former ICSU World Data Centres (WDCs) and Federation of Astronomical and Geophysical data analysis Services (FAGS). It aims to transition from existing standalone services to a common globally interoperable distributed data system that incorporates emerging technologies and new scientific data activities.

The primary goals of the ICSU WDS are to:

- enable universal and equitable access to quality-assured scientific data, data services, products, and information,
- ensure long-term data stewardship,
- foster compliance of agreed-upon data standards
- provide mechanisms to facilitate and improve access to data and data products.

The WDS is striving to become a worldwide ‘community of excellence’ for scientific data, with searchable common data directories and catalogues that ensure long-term stewardship and provision of quality-assessed data and data services to the international scientific community.

The governing body of ICSU-WDS is the WDS Scientific Committee (SC). It is made up of leading scientists that cover a broad range of disciplines and geographical areas. They are actively involved with data/computer sciences and deal with large dataset issues. Financially supported and hosted by the National Institute of Information and Communications Technology until April 2016, the WDS International Programme Office (IPO) was officially opened in Tokyo on 9 May 2012. The IPO acts, with guidance from the Scientific Committee, to coordinate WDS operations under the leadership of its Executive Director. It takes responsibility for organizing biannual SC meetings, biennial WDS conferences, as well as outreach and promotional activities.

Membership has been growing steadily, starting with a first group of certified Members from the former WDCs and FAGS, as well as several new organizations. New applications are also evaluated regularly by the WDS Scientific Committee on criteria such as access to high quality data, data stewardship, and participation in broad harmonization and interoperability efforts. Furthermore, WDS members adopt open and equitable access to quality-assured scientific data, data services, products, and information. Membership has a broad disciplinary and geographic base and establishment of nodes in emerging countries will be encouraged. Different membership-types are available for organizations, dependent on their activities and interests. Thus far, WDS has certified and accredited 54 Members, and co-opted another 14. The Flanders Marine Institute is a regular member and the IOC International Oceanographic Data and Information Exchange (IODE) programme is one of the first network members.

WDS main focus is on the long-term stewardship of scientific data, as well as on data and metadata services such as analysis and publication services. Strong relationships are thus being forged with science publishers and libraries. Emphasis is also on linking with data production facilities, in particular science projects and programmes. WDS will furthermore provide common data and information infrastructure and interoperability conforming to global standards. Hence, efficient usage of supplied data and services, in addition to exchanges with other networks, is ensured.

New activities include:

(i) Developing an online open WDS Metadata Catalogue that can be used to discover and access the quality-assessed scientific data and information holdings of WDS Regular and Network members.

(ii) Promoting and establishing the data publication concept among data centers, science publishers and bibliometric service providers.

COCONET PROJECT (Towards COast to COast NETworks of Marine Protected Areas) - the architecture of a common Geodatabase for Marine Data Management and synthesis

Valentina Grande, CNR-ISMAR, valentina.grande@bo.ismar.cnr.it (Italy)
Federica Foglini, CNR-ISMAR, federica.foglini@bo.ismar.cnr.it
Francesca Lisi, CNR-ISMAR, Bologna, Italy
Francesco De Leo, CONISMA, Lecce, Italy
Aurore Chassanite, CNRS, Université de Perpignan, France
Oleksandr Neprokin, UkrSCES, Odessa, Ukraine
Richard Lisovsky, UkrSCES, ricgardl@te.net.ua

Abstract

This paper illustrate the architecture of an integrated marine Geodatabase developed within The COCONET project (Towards COast to COast NETworks of Marine Protected Areas) to manage marine protected areas and Ocean Wind Farms (OWF) in Mediterranean and Black Sea. We are creating a common framework for management, storage and analysis of marine data, that come from various and heterogeneous sources. We are following the INSPIRE Directive to homogenize metadata, data and database structure, considering also the SeaDataNet standards and tools for data and metadata storage. The final goal is a digital and interactive map of marine protected areas with information about designations, habitats, species, biological connectivity, geology, threats, clean energy, etc, representing a useful tool for decision makers to assess the effectiveness of marine protected sites network and to propose new potential sites.

Key words

Integrated management, marine data, Geodatabase, INSPIRE Directive, marine protected area network.

Introduction

COCONET is a European project that will produce guidelines to design, manage and monitor network of MPAs and Ocean Wind Farms. The Project covers a high number of Countries and involves researchers covering a vast array of subjects, developing a timely holistic approach and integrating the Mediterranean and Black Seas scientific communities through intense collective activities and a strong communication line with stakeholders and the public at large. Within this project we aim at providing a common framework for marine data management and final synthesis of the outcomes of different scientific topics from heterogeneous sources. An integrated

Geodatabase and a WebGIS system is the linking tool for all partners, regions and thematic research. It involves the entire consortium at different levels in topics such as data provision and integration, GIS products, GIS interpretation, data archiving and data exchange.

Objectives

The work is organised around the following main objectives:

- assess the rules for data and metadata sharing between partners reviewing the existing common European protocols and standards (INSPIRE and SeaDataNet);
- design and implement data repositories (Marine Geodatabase) following the INSPIRE Directive, to store and retrieve the spatial data collected during the lifespan of the project for the Mediterranean and Black Sea areas and for the pilot study areas;
- develop the COCONET WebGIS to integrate the multi scale GIS layers derived from all regions going towards an integrated management of the marine resources;
- develop an analytical and evaluative framework for designing, managing and monitoring regional networks of MPAs, including wind farms, centred on science-based guidelines, criteria, concepts and models.

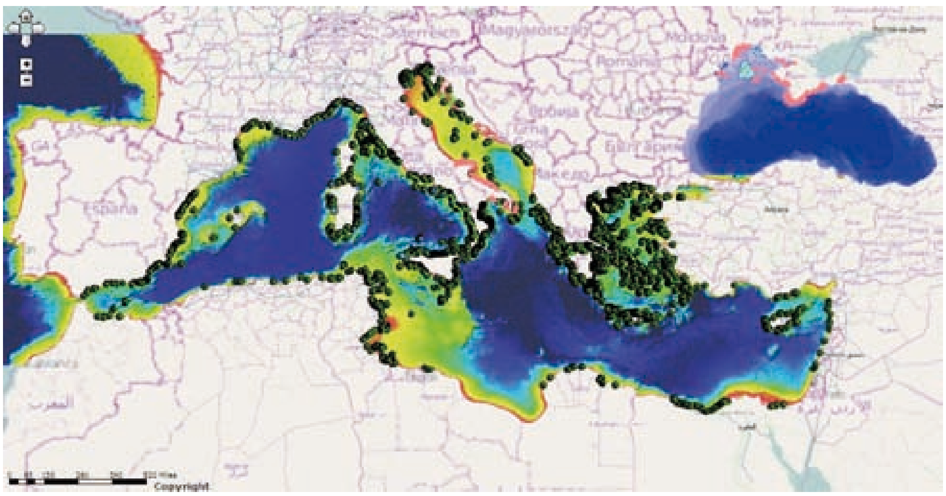


Fig. 1 - The image shows the preliminary COCONET WebGIS view: in this example we can see the bathymetry (base map from EMODnet Hydrography Portal, http://www.emodnet-hydrography.eu/), the seagrasses distribution (green points) and some example of threats (pink points) in Mediterranean and Black Sea.

Geodatabase structure – Data Layers

We are designing the architecture of six Geodatabases:

- Protected sites
- Habitat and Biotopes
- Threats
- Geology
- Species distribution
- Offshore Wind Farms

Each theme is developed as a physical UML schema (by using Miscrosoft Visio) and implemented in a dedicated Geodatabase (RDBMS) with layers and related tables (by using

ArcGIS) that can be visualized and queried by all partners interactively trough the WebGIS. A first version of the WebGIS will be available from August 2013, but the structure of Geodatabases will be implemented and improved step by step according to available data during the lifespan of the project.

Conclusions

The final goal is to deliver digital maps of networks of marine protected areas and offshore wind farms as final synthesis of the outcome from all scientific topics. The integrated Geodatabase will be a fundament tool to produce the guidelines to design, manage and monitor network of MPAs, and an enriched wind atlas for both the Mediterranean and the Black Seas. COCONET metadata will be integrate in the SeaDataNEt infrastructure using the pan European metadata standards. The Project will identify groups of putatively interconnected MPAs in the Mediterranean and the Black Seas, shifting from local (single MPA) to regional (Networks of MPAs) and basin (network of networks) scales. The identification of physical and biological connections with clear the processes that govern patterns of biodiversity distribution. This will enhance policies of effective environmental management, also to ascertain if the existing MPAs are sufficient for ecological networking and to suggest how to design further protection schemes based on effective exchanges between protected areas.

SESSION 2

Marine environmental data bases: infrastructures and data access systems

- Coastal and deep-sea operational oceanography metadata/data systems
- Physical and bio-chemical databases for climate studies
- Geophysical and geological metadata/data systems

ORAL PRESENTATIONS

ICES Report on Ocean Climate (IROC) - Online

Hjalte Parner, ICES, hjalte.parner@ices.dk (Denmark)

The ICES Working Group on Oceanic Hydrography (WGOH) closely monitors the oceanographic conditions in the ICES region (North Atlantic) by updating and reviewing the results from standard hydrographic sections and stations.

The material presented at the WGOH meetings each year is consolidated and published in the annual ICES Report on Ocean Climate (IROC).

With the IROC, the Working Group analyses multiple time-series in a consistent way to give an overview of the state-of-the-environment in the North Atlantic that includes:

- North Atlantic climate headlines
- Summary of upper ocean conditions
- The North Atlantic Atmosphere
- Detailed area descriptions, part I: The upper ocean
- Detailed area descriptions, part II: The deep ocean

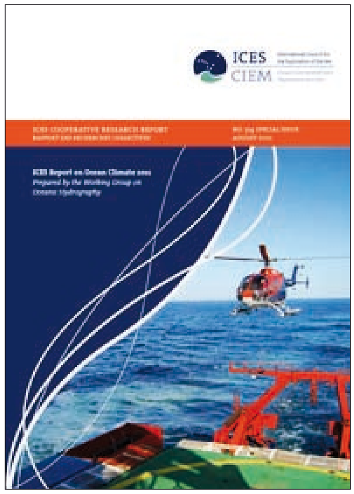


Fig. 1 - ICES Report on Ocean Climate (IROC).

IROC data series on the web

A new web application http://ocean.ices.dk/iroc has been developed that enables the users to explore the data series behind the standard hydrographic sections and stations found in the IROC report.

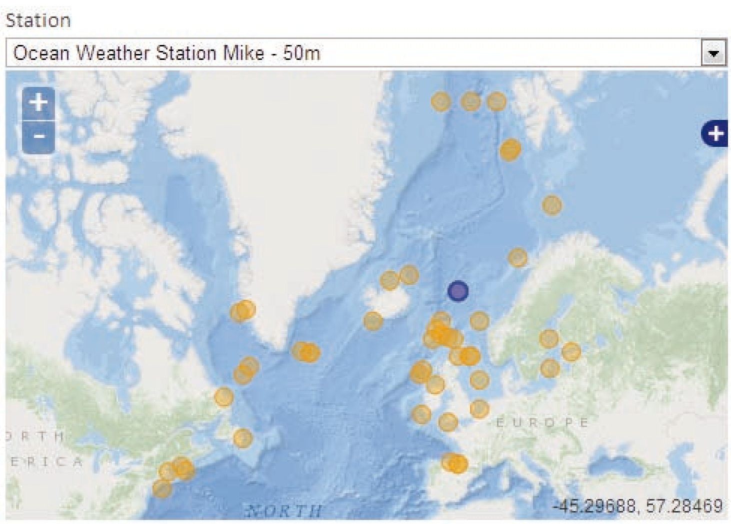


Fig. 2 - Map of the IROC sections and stations.

By selecting a given section or station in the area of interest the users are able to view and download the figures or the actual data series behind. Likewise a link is provided to the latest IROC report itself.

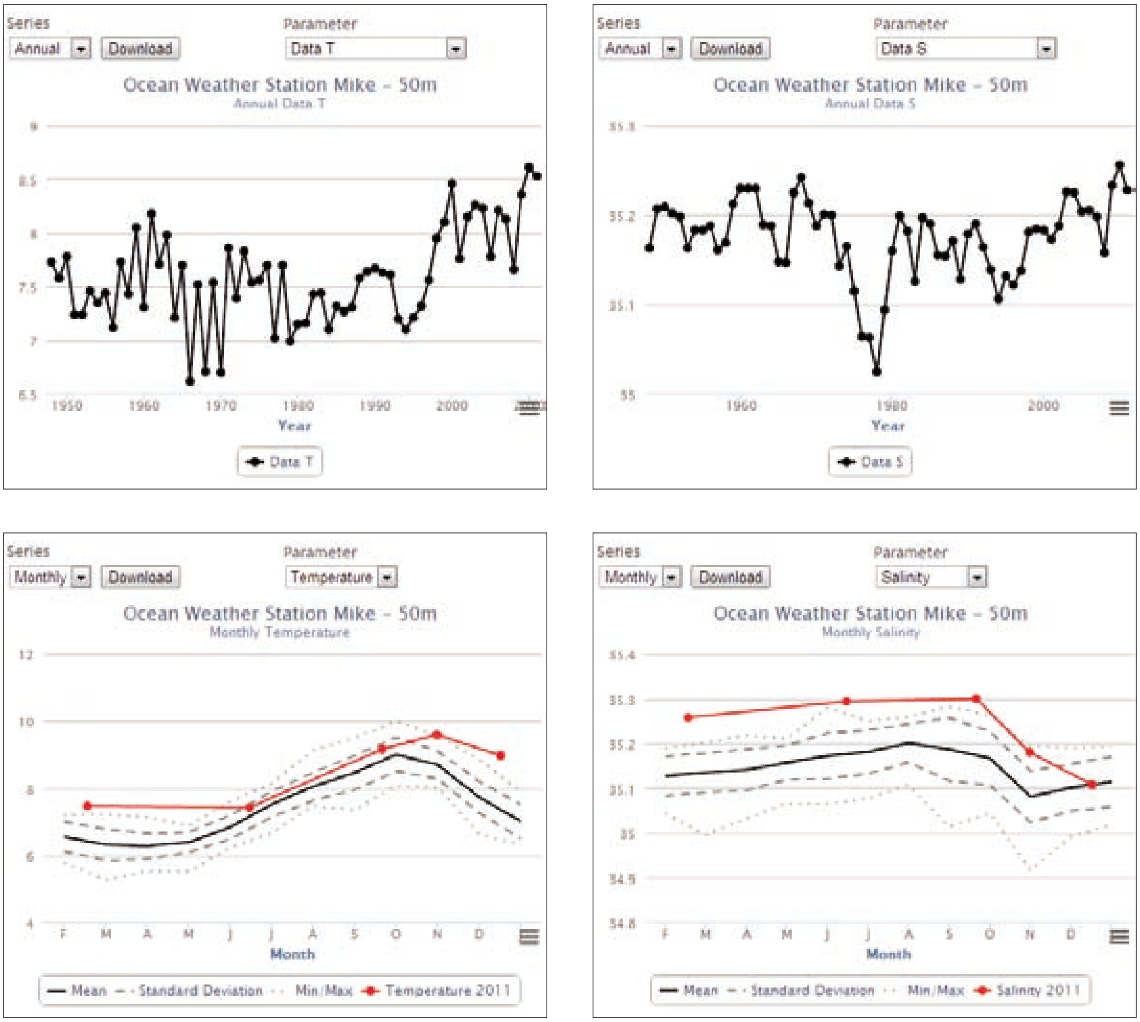


Fig. 3 - Annual (upper panel) and Monthly (lower panel) Temperature (left panel) and Salinity (right panel) series of Ocean Weather Station Mike in the Norwegian Sea at 50 meter.
Data Provider: Geophysical Institute – University of Bergen - Norway.

Ocean Data Portal: from data access to integration platform

Sergey Belov, All-Russian Research Institute of Hydrometeorological Information World Data Center (RIHMI-WDC), belov@meteo.ru (Russia)

Nikolai Mikhailov, All-Russian Research Institute of Hydrometeorological Information World Data Center (RIHMI-WDC), nodc@meteo.ru (Russia)

Kristina Belova, All-Russian Research Institute of Hydrometeorological Information World Data Center (RIHMI-WDC), christy@meteo.ru (Russia)

Development of a distributed information infrastructure is highly important for the marine science and maritime activities because of the trend towards integration of both local and geographically distributed applications that provide access to heterogeneous data and information resources of the marine environment.

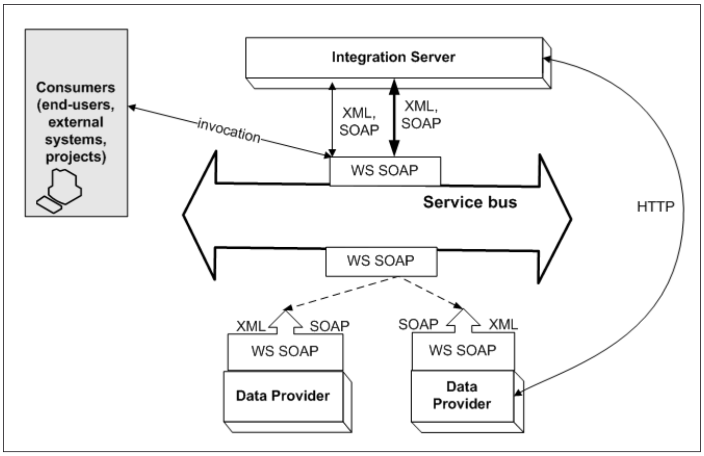


Fig. 1 - Simplified architectural view of the ODP integration platform.

The use of service - oriented approach to build the ocean data distributed network allows as to give different users access to data on the marine environment and marine activities, and to organize a remote call processing methods that operates on the data owners side.

Data integration platform, developed within the framework of the IODE Ocean Data Portal (ODP) project is called to organize and manage distributed access and processing of information about the world oceans. Platform provides interoperability components, their integration into the system, management and monitoring, provides a dispersion of components in the interacting information technology system nodes.

Interaction of components is based on the service bus technology, web-services and wide implication of widely used IT standards (ISO, XML, SOAP, etc.) and documents control interfaces of interacting components. ODP service bus is a specialized software that allows to encapsule web-services of all components, and provides access to these services through a single point. All together they are combined into the integration platform for data and services provided via Ocean Data Portal technology.

Implementation of the integration platform technology is based on client-server model with respect to multithreading, load balancing, low resource consumption and reliability. Integration platform includes core componets dealing with data – ODP Data Provider and ODP Integration Server (see Fig.1).

The Ocean Data Portal is developed as a component of the IODE system to provide modern data exchange and dissemination infrastructure to achieve the IODE objectives:

- to facilitate and promote the exchange of all marine data and information including metadata, products and information in real-time, near real time and delayed mode;
- to ensure the long term archival, management and services of all marine data and information;
- to promote the use of international standards, and develop or help in the development of standards and methods for the global exchange of marine data and information, using the most appropriate information management and information technology;
- to assist Member States to acquire the necessary capacity to manage marine data and information and become partners in the IODE network; and
- to support international scientific and operational marine programmes of IOC and WMO and their sponsor organizations with advice and data management services.

The iMarine Data Bonanza: Improving Data Discovery and Management through an Hybrid Data Infrastructure

Donatella Castelli, ISTI-CNR, donatella.castelli@isti.cnr.it (Italy)
Pasquale Pagano, ISTI-CNR, pasquale.pagano@isti.cnr.it
Leonardo Candela, ISTI-CNR, leonardo.candela@isti.cnr.it
Gianpaolo Coro, ISTI-CNR, gianpaolo.coro@isti.cnr.it

Nowadays both data owners and data consumers are very conscious of how the availability of "data" in contexts different from where it was produced benefits science and society at large. Hence there is a larger availability of data, also across domains, that potentially opens the way to new types of scientific practices, *e.g.* experiments, analysis, modelling, that were not possible few years ago. Currently datasets are accessible through diverse Information Systems, via different protocols and through different user interfaces, therefore a considerable amount of scientists' time is spent in understanding how to access the datasets, in selecting the most appropriate ones, homogenizing them and, more in general, preparing the datasets that fit the purpose of the planned scientific investigation. As a result researchers and technologists in computer science are challenged to investigate new approaches for data sharing and management practices that make the underlying complexity transparent to scientists. Given the heterogeneity and dynamicity of the data contexts these approaches must be flexible and powerful enough to adapt to the multitude of different and evolving situations.

This scenario is also true for Marine data as there is no one single Information System that offers the community of practice access to all the data needed (e.g. environmental data to statistical data, fact sheets, and distribution maps) in a seamless way.

iMarine (www.i-marine.eu) is a data infrastructure conceived to support the application of the principles of the Ecosystem Approach (a strategy for the integrated management of land, water and living resources that promotes conservation and sustainable use in an equitable way) to fishery management and conservation of marine living resources. Such an infrastructure offers its community of practice facilities both to access a rich array of marine data in a seamless way as well as a comprehensive set of data processing offered *"as-a-Service"*. This paper briefly introduces iMarine and the services it offers for data discovery and management.

The iMarine Hybrid Data Infrastructure in a Nutshell

The iMarine Infrastructure is the implementation of a Hybrid Data Infrastructure, i.e. an IT infrastructure conceived to enable a data-management-capability delivery model in which computing, storage, data and software are made available *as-a-Service*. In essence, it builds on the cloud paradigm to offer resources (data, services, computing) as a utility, *elasticity* of resources and the perception of infinite capacity as key features to make data and data management services available *on demand*.

This infrastructure is not built from scratch. Rather, it is a "system of systems" built by exploiting other existing infrastructures, services and Information Systems such as the European Grid Infrastructure (EGI), Global Biodiversity Information Facility (GBIF), the Ocean Biogeographic Information System (OBIS), myOcean, World Ocean Atlas, FAO Data, Catalogue of Life, World Register of Marine Species (WoRMS), and many more. The iMarine Infrastructure integrates these systems with the aim of exploiting the synergy amongst them, and thus provide its community of practice with a set of novel and enhanced services.

The infrastructure also supports the creation and operation of *Virtual Research Environments* (VREs), *i.e*. web based working environments where groups of scientists, usually geographically distant from each other, have transparent and seamless access to a shared set of remote resources (data, tools and computing capabilities) needed to perform their work.

The iMarine Data Access Services

The iMarine Data Infrastructure caters for scientists and practitioners seeking access to marine data and in particular it offers services for seamless access to species data, geospatial data, statistical data and semi-structured data from diverse data providers and Information Systems. These services can be exploited both via web based graphical user interfaces and web based protocols for programmatic access, e.g. OAI-PMH, CSW, SDMX.

For species data, the infrastructure is equipped with a Species Data Discovery (SDD) Service which mediates over a number of data sources including taxonomic information, checklists and occurrence data. The service is equipped with plug-ins interfacing with major information systems such as Catalogue of Life, Global Biodiversity Information Facility, Integrated Taxonomic Information System, Interim Register of Marine and Nonmarine Genera, Ocean Biogeographic Information System, World Register of Marine Species. To enlarge the number of information systems and data sources integrated into SDD, the VRE data manager should simply implement (or reuse) a plug-in. Each plug-in can interact with an information system or database by relying on a standard protocol, e.g. TAPIR, or by interfacing with its proprietary protocol. Plug-ins mediate queries and results from the language and model envisaged by SDD to the requirements of a particular database. SDD promotes a data discovery mechanism based on queries containing either the scientific name or the common name of the target species. Furthermore, to tackle potential issues related to taxonomy heterogeneities across diverse data sources, the service supports an automatic query expansion mechanism, i.e. the query could be augmented with "similar" species names. Discovered data is presented in a homogenised form, e.g. in a typical Darwin Core format.

For geospatial data, the infrastructure is equipped with services generating a Spatial Data Infrastructure compliant with OGC standards. In particular, it offers a catalogue service enabling the seamless discovery of and access to every geospatial resource registered or produced via the iMarine infrastructure or its services. These resources include physical and biochemical environmental parameters like temperature and chlorophyll, species distribution and occurrence maps, and other interactive maps. Some of these resources are obtained by interfacing with existing Information Systems including FAO GeoNetwork, myOcean and World Ocean Atlas. New resources can be added by linking data sources to the SDI via standards or ad-hoc mediators. On top of the resulting information space, iMarine offers an environment for identifying resources and overlays them through an innovative map container that caters for sorting, filtering, and data inspection further to standard facilities such as zoom in.

For statistical data, the infrastructure is equipped with a dedicated statistical environment supporting the whole lifecycle of statistical data management including data ingestion, curation, analysis and publication. This environment provides its users with facilities for creating new datasets and code lists by using sources like CSV or an SDMX repository, curating the datasets (by using controlled vocabularies and code lists, defining data types and correcting errors), manipulating the datasets with standard operation like filtering, grouping, and aggregations, analysing the datasets with advanced mining techniques like trend and outliers detection, producing graphs out of the datasets, and finally publishing datasets in an SDMX registry for future use.

Besides accessing this plethora of data, iMarine offers a number of facilities for processing and inspecting all the available data, e.g. data mining techniques can be applied to any kind of dataset. Moreover, it is always possible to simply “save” the discovered data in various formats and share them with co-workers through a dedicated workspace. Thus, it is possible, for example, to access a time series via SDMX and publish it in the Spatial Data Infrastructure or to analyse the distribution trend of species in geographical space.

Bridging the gap between data and metadata (Part 2): looking backward at what happened during a survey with the Eurofleets Automatic Reporting System (EARS)

Paolo Diviacco, OGS, pdiviacco@ogs.trieste.it (Italy)
Alessandro Busato, OGS, abusato@ogs.trieste.it (Italy)
Marie-Paule Corre, IFREMER, marie.paule.corre@ifremer.fr (France)
Karien De Cawer, MUMM, Karien.DeCauwer@mumm.ac.be (Belgium)
Marc Nokin, IFREMER, Marc.Nokin@ifremer.fr (France)
Serge Scory, MUMM, serge.scory@mumm.ac.be (Belgium)
Jean-Marc Sinquin, IFREMER, Jean.Marc.Sinquin@ifremer.fr (France)
Jordi Sorribas, UTM-CSIC, sorribas@utm.csic.es (Spain)
Yvan Stojanov, MUMM, yvan.stojanov@mumm.ac.be (Belgium)

The problem

The Eurofleets EU FP7 project aims at bringing together the European research fleets to enhance their coordination and promote the cost effectiveness of their facilities. Among the many activities of the Project, a specific task was devoted to the development of a software tool intended to record and report what happens during a survey, aside the planned and actual data acquisition.

Automatic underway data acquisition as well as manual event logging had to be foreseen. The former takes care of the registration of data such as geographic positioning, echo sounding, or thermosalinograph values, while the latter takes care of events that can range, for example, from a sample taken over an observation made to any malfunction that might be identified during the cruise. In both cases it can be difficult to evaluate whether the recorded information pertain to metadata or to the actual data. For example recordings of sea-waves can be used to correct vertical positioning in VHR seismics but can be an observation in itself if used in a meteorological context. Sound speed in the water column can be used to calibrate multibeam surveys but, being related to temperature, salinity etc, can be used to detect water masses. In fact, the traditional rather strict separation between data and metadata reveals here its limits so that event logging becomes a bridge between the two words. Storing all events in a database, it is possible to produce survey events reports that can be used in the process of data discovery, for example, exploiting the identification and explanation of possible anomalies while defining the quality of data.

The implemented solution

To put all this in practice it was decided to leverage the possibilities of the Ifremer software CASINO+ (I) broadening its adoptability by any vessel in the project and (II) extending its functionalities, to obtain a new system that has been named Eurofleets Automatic Reporting System (EARS).

Automatic event logging

The problems of adoptability were related to the fact that CASINO+ worked in tandem with another Ifremer system called TECHSAS for automatic data acquisition, which has not entered Eurofleets. It was necessary therefore to introduce a new set up where devices with serial outputs have been connected to serial-to-Ethernet boxes linked to a private network sensed by EARS. We proved that this configuration does not introduce delays so that data are correctly referenced in time and therefore in space. CASINO+ setting has been very flexible but at the expenses of simplicity. Therefore a Graphical User Interface has been developed to ease the creation of a catalogue of instruments where setups can be recalled easily from survey to survey but also exchanged across the fleets whenever the same devices are used.

Manual event logging

To handle manual events, a completely new paradigm was designed. It started from the very basic design of the concept of event, which has not been identified with a flat term, rather with the combination of seven components:

- **Subject**: refers to the domain in which the event takes place, for example a seismic system or a sediment sampler
- **Tool**: the device that produces the event, for example an anemometer or a Niskin bottle
- **Category**: describes the kind of event taking place, for example a malfunction, sampling, operation, or phase change
- **Action**: which is the actual task performed, for example start sampling, close bottle or start recording.
- **Comment**: which is a free text field that is left to the operator to enter more information like for example the visual description of a sediment sample
- **Actor**: is the person performing the action
- **Action_property**: any additional characteristic or parameter accompanying a given action for which the user needs to enter a value onboard e.g. volume of water centrifuged, installation height of a sensor or distance trawled.

This event structure is very flexible and allows a high variety of events to be stored together with a timestamp. For some of the concepts, the semantic framework of the Natural Environment Research Council (NERC) Vocabulary Server can be used. The controlled vocabulary for 'Action' and the relations between the terms of the different concepts are currently under development within the Eurofleets ontology governance.

Due to difficult sampling conditions and time constraints on board, the entry of manual events should be really easy to do. With one double-click or drag-and-drop, users can register an event using a personal, discipline-specific configuration tree. This tree shows a selection of tools to be used during the cruise with the possible actions. The configuration can be made in advance and re-used later on. Also the need for use outside on the deck for example, has been considered as well as multilingualism. Developments during Eurofleets II aim to use the event ontology as input for EARS improving further user-friendliness.

Reporting

Once manual and automatic events are entered in the database, reporting is possible. EARS offers a two "resolution" automatically produced reporting output that has been named Ship Summary Report (SSR) and that is coded following SensorML OGC standard. The first level of SSR is very generic and reports only very basic information such as, for example, the position of a ship or its identification. The core SSR is queried from the Eurofleets EVIOR system which is intended to represent the current status of all the vessels participating in the Project.

At the same time the core SSR links to another document where all information from the database can be listed. This is the extended SSR allowing very detailed information retrieval to be done using for example XSLT.

EARS has been tested and proved to be capable of addressing the above mentioned issues, so that within the following Eurofleets 2 project it is planned to install it in most of the partner's vessels.

Modernised Cruise Summary Reports Management, in particular with a Link to Data

Anne Che-Bohnenstengel, BSH-DOD, anne.che-bohnenstengel@bsh.de (Germany)
Friedrich Nast, BSH-DOD, friedrich.nast@bsh.de

Cruise Summary Reports (CSR) are one component of the end-to-end data management strategy and have a long tradition since the 1960s. BSH-DOD, running the SeaDataNet CSR inventory, holds more than 45000 cruises of ~2000 different ships. The CSR perspective is twofold:

- **From the funding agency's point of view**
  It is current practice that upon cruise application the funding agencies demand a data management plan. After the cruise the data must be accessible for other scientists as well as for the public so that the high cost for the expedition will be justified. A clear definition for the final data archiving has to be stated. Thereby various types of data measured during a single cruise might be finally archived in different repositories. Therefore the CSR serves as a central reference for what has been measured at sea and as a link to the final storage of the data – and at best the CSR is directly linked to the data.

- **From the scientist's point of view**
  Scientists looking for data as well as scientists working for global data compilations use the CSR inventory not only for the metadata but also for data discovery. A link between the SeaDataNet Common Data Index(CDI) and the respective CSR is indispensable.

At present the DOD is extending its Data Management system to make it more efficient and to automate as much of this process as possible. Through DOD's collaboration in SeaDataNet and MaNIDA (the German Marine Network for Integrated Data Access) some synergy effects flow into our sequence of operation. In particular the SDN CSR directory will be equipped with CS-W harvesting capabilities for harvesting records from partner data centres in addition to the already existing XML ingestion, thus turning the CSR data flow into a de-central automation.

A major step towards a comprehensive data archiving is to include the cruise programme metadata in the DOD database as a CSR checklist and consequently also a checklist for potential measurement data sources. A template for the Cruise Summary Report can easily be generated from the cruise programme metadata which again can be completed via the online content management system for CSRs by the chief scientist once the cruise is launched or completed.

Another essential renovation is to make use of the automatic logging system of the "underway" data for additional information such as station, device and parameter lists, which will gradually be archived in the database. In the MaNIDA project it is also planned to generate the CSRs automatically in SDN XML-format, which will be ingested as new submission or updates for existing records in the CSR inventory. This is in line with the FP7 EUROFLEETS project in which systems will be defined and implemented for registering data acquisition events during cruises resulting in automatic CSR reporting after and even during cruises (en-route CSR).

Quality control and ingestion of the data into the DOD Oracle database are finally processed using a flexible interface which is capable of handling data in various formats. As services for the users the DOD provides the following:

- CSR web service in INSPIRE compliant format
- Web map service for track charts
- Data download for each cruise in ODV format
- Link between the SDN CDI inventory and the CSR inventory

But for all automations and innovations detailed description of cruise measurements and goals in the CSRs by the chief scientists remain the basic prerequisites for an efficient system.

The CNR-ISAC Informatics Infrastructure for the Satellite Climatological and Oceanographic data: production, harmonization and dissemination in Interoperability Frameworks

Vega Forneris, ISAC-CNR GOS, vega.forneris@artov.isac.cnr.it (Italy)
Cristina Tronconi, ISAC-CNR GOS, cristina.tronconi@isac.cnr.it (Italy)
Gianluca Volpe, ISAC-CNR GOS, Gianluca.Volpe@isac.cnr.it (Italy)
Simone Colella, ISAC-CNR GOS, Simone.Colella@isac.cnr.it (Italy)
Bruno B. Nardelli, ISAC-CNR GOS, bruno.buongiornonardelli@cnr.it (Italy)
Andrea Pisano, ISAC-CNR GOS, andrea.pisano@artov.isac.cnr.it (Italy)
Rosalia Santoleri, ISAC-CNR GOS, Rosalia.Santoleri@isac.cnr.it (Italy)

The Satellite Oceanography Group (GOS) CNR-ISAC is one of the leading Units of the Earth Observation Division of ISAC-CNR since 1987 and focuses on operational Satellite Data processing and analysis. GOS is responsible for the Mediterranean and Black Seas satellite operational systems in the framework of several Partnership, such as MyOcean (http://www.myocean.eu/), SeaDataNet (http://www.seadatanet.org/), MOON (http://moon.santateresa.enea.it/), and others (some still under development). As Production Unit, CNR-ISAC GOS systems acquire satellite data from different data upstream (NASA, ESA, etc), process them in a distributed environment (grid, clusters/blades) and produce Near Real Time (NRT), Delayed Time (DT) and Re-analysis (RAN/REP) of Ocean Color (chlorophyll-a and several optics parameters such as KD490, RRS at several wave length bands, etc) and Sea Surface Temperature (at different spatial resolutions) covering the Mediterranean and the Black Seas.

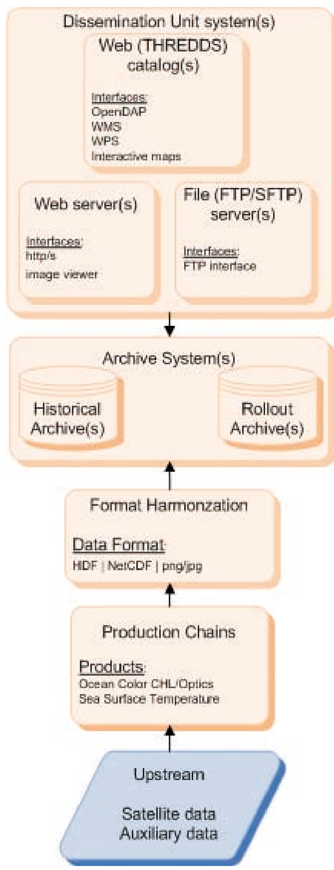


Fig. 1 - Satellite processing workflow.

Working in such frameworks/projects, which involve several partners and institutes spread through Europe, made clear the importance of the interoperability of different realities (*"Being able to accomplish end-user applications using different types of computer systems, operating systems, and application software, interconnected by different types of local and wide area networks"* James O'Brien, George Marakas: Introduction to Information Systems, McGraw-Hill/Irwin). One of the main objectives of such shared projects was in fact to define standards and common approaches in order to standardize formats and interfaces, to provide end users with homogenous results in terms of data formats, conventions and interfaces for accessing those products.

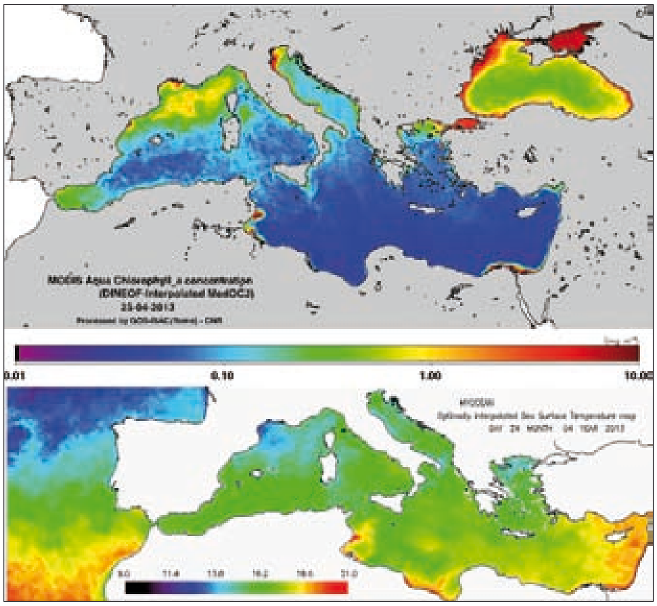


Fig. 2 - OC (upper panel) and SST (lower panel) L4 products (static images).

An overview of the processing procedures involved in such projects will be presented:

- **Satellite data processing**: based on GRID technology, clusters and blades, GOS's processing chains download data input files (L0/L1/L2P_GHRSST) and process them up to final products (L3/L4) using regional algorithms, optimized through the implementation of C++ and/or MPI libraries and taking advantage of distributed computing

- **Products harmonization**: will be presented a quick overview of formats, conventions and specifications commonly used, in particular: NetCDF Climate and Forecast (CF) Metadata Convention (http://cf-pcmdi.llnl.gov/), INSPIRE DIRECTIVE (http://inspire.jrc.ec.europa.eu/), GDS (GHRSST Data Set) documents: (https://www.ghrsst.org/documents/q/category/gds-documents/)

- **Products dissemination**: MyOcean Project will be used as use case, showing the THREDDS catalog and its interfaces (OpenDAP and WMS) and MOTU (custom software deployed in the MyOcean Project framework) used for creating higher levels services (MOTU).

Data management at the CATDS-PDC, the operational processing center for high level data of the SMOS satellite

Stéphane Tarot, IFREMER, stephane.tarot@ifremer.fr (France)
Sarah Guibert, CNES, sarah.guibert@cnes.fr (France)
Yann Kerr, CESBIO, yann.kerr@cesbio.cnes.fr (France)
Carole Larigauderie, CNES, carole.larigauderie@cnes.fr (France)

The SMOS (Soil Moisture & Ocean Salinity) satellite was launched in November 2009 and is fully operational since May 2010. The data processing facilities are split into 2 main parts:

- The ESA DPGS (Data Processing Ground Segment) in charge of level 1 (L1) and level 2 (L2) products,
- The CATDS (Centre Aval de Traitement des Données SMOS) in charge of level 3 (L3) and level 4 (L4) products.

The CATDS itself is split into three components:

- A C-PDC (CATDS Production and Dissemination Centre) which routinely produces and disseminates L3 and L4 data from L1B and auxiliary data, which performs the reprocessing campaigns of L3 and L4 data and which provides assistance to users. It is located at SISMER (IFREMER) in Brest.
- Two C-EC (CATDS Expertise Centre) which host the algorithms definition, which assess the quality of the products and which provide specific information to users. The first C-EC, dedicated to Soil Moisture, is located at CESBIO in Toulouse. The second C-EC, dedicated to Ocean Salinity, is co-located at LOS (IFREMER) in Toulouse and at LOCEAN in Paris.

The C-PDC processing chain to generate Soil Moisture and Ocean Salinity L3 data from L1B products is composed of several processors which are chained together. Some of the processors also require auxiliary data, either static or dynamic.

The L3 products generated by C-PDC are the following:

- 1 day global map of soil moisture values (daily aggregation of SMOS orbits)
- 3 day/10 day/monthly global map of soil moisture values
- 1 day ocean salinity values
- Average salinity fields (monthly 200km average for the moment, soon extended to 10 day and monthly 25, 50, 100 and 200 km averages).

C-PDC also produces and distributes global map of polarized brightness temperature at ground level (arranged by incidence angle values). At the moment, there isn't yet any L4 product.

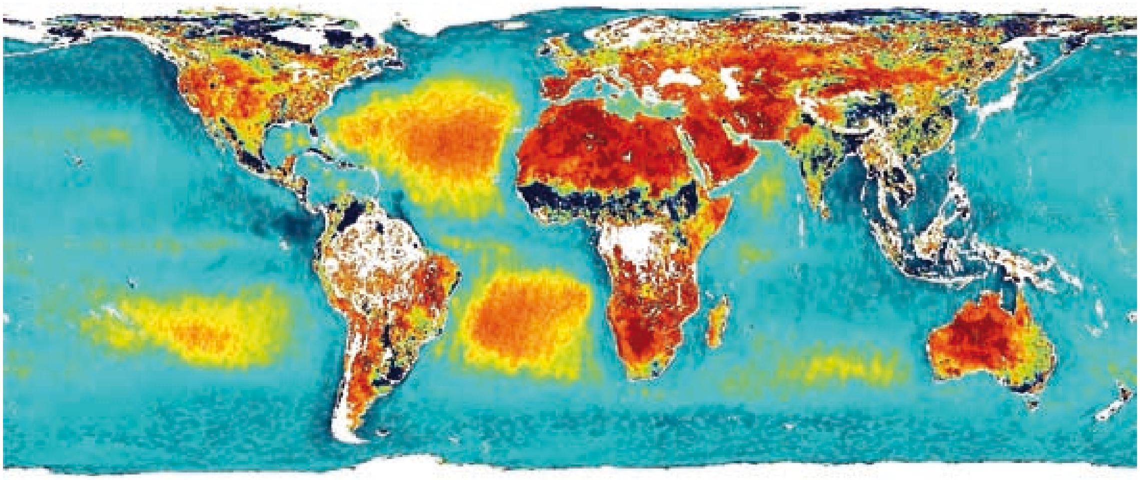


Fig. 1 - Combined map of Ocean Salinity and Soil Moisture.

The C-PDC is operational since the middle of year 2011. Each day, it process 110 Go of operational data (including input data) and uses 110 hours of CPU time. The first reprocessing campaign was performed mid-2012, taking into account more than 2 years of measurements (January 2010 to April 2012). The second reprocessing campagn is planned for September 2013 and will concerned almost 4 years of data.

A web site (www.catds.fr) describes the products and gives information to users. The products are distributed :

- on FTP,
- through SIPAD, a web-based tool which allows aggregation and sub-settings.

The presentation shows the CATDS-PDC with a focus on the data management and operational challenges, such as the processing constraints, volume of data, ... The L3 data dissemination process and the way to retrieve CATDS L3 products are also described.

Geo-Seas: delivering harmonised marine geoscience data on a European scale

Helen Glaves, British Geological Survey, hmg@bgs.ac.uk (UK)
Dick Schaap, Mariene Informatie Service (MARIS), dick@maris.nl (Netherlands)

Significant volumes of both raw and interpreted marine geoscience data are held by a wide range of institutions throughout Europe. This data, which is largely unique and therefore irreplaceable, has been acquired for a diverse range of purposes and often at significant cost. Facilitating re-use of this data is therefore a priority for data managers. However, there are challenges associated with the re-purposing of this data due to the use of different formats, standards, nomenclatures, scales and co-ordinate systems at the organisational, national and international level.

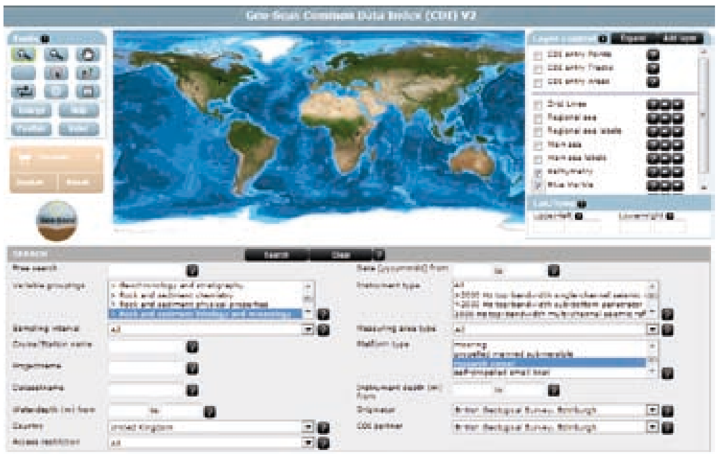


Fig. 1 - Geo-Seas data discovery and access service.

To address these barriers the Geo-Seas project has developed and implemented an e-infrastructure for the discovery and access of marine geological and geophysical data.

The dedicated portal (http://www.geo-seas.eu), which provides access to more than 120,000 standardised datasets from 26 federated data centres in 17 European countries, has been developed by adopting and adapting the technologies, standards and methodologies developed by the SeaDataNet project for the management and delivery of marine and oceanographic data.

Through the re-use of this pre-existing architecture, including the associated common standards and vocabularies, a joint infrastructure for the management of both marine geoscientific and oceanographic data has been created. The Geo-Seas project has also developed a suite of visualisation tools which assist the user in assessing the suitability of individual data sets for their own particular requirements. These tools and services allow users to visualise seismic, borehole and 3D DTM data as part of the data selection process.

The Geo-Seas initiative had lead to a major improvement in the availability of standardised marine geoscientific data throughout Europe allowing end users better access to marine geological and geophysical data for use in a range of applications. This approach has also lead to the development of collaborative links with other European projects such as EMODNET, Eurofleets and EGDI-Scope, as well as with similar initiatives in other global regions including the Rolling Deck Repository (R2R) initiative in the USA and IMOS in Australia.

EMODNet
Physical Parameters

Antonio Novellino, ETT s.r.l., antonio.novellino@ettsolutions.com (Italy)
Giuseppe Manzella, ENEA, giuseppe.manzella@enea.it
Dick Schaap, MARIS, dick@maris.nl
Patrick Gorringe, EuroGOOS, Patrick.Gorringe@smhi.se
Lesley Rickards, NERC BODC, ljr@bodc.ac.uk
Sylvie Pouliquen, Ifremer, sylvie.pouliquen@ifremer.fr

European Marine Observation and Data Network (EMODnet) has been created to improve the usefulness to European users for scientific, regulatory and commercial purposes of observations and the resulting marine data collected and held by European public and private bodies, wherever that data has been collected from. European Commission, represented by the Directorate-General for Maritime Affairs and Fisheries (DG MARE), is running several service contracts for creating pilot components of the ur-EMODNET and it is assisted by a Marine Observation and Data Expert Group (MODEG) .

The Emodnet Physical Parameters Portal (http://www.emodnet-physics.eu/) is one of the EMODnet portals and it is aimed at providing access to archieved and real time data catalog on the physical condition in Europe's seas and oceans. The overall objectives of the EMODNet Physics preparatory action is to provide access to archived and near real-time data on physical conditions in Europe's seas and oceans by means of a dedicated portal and to determine how well the data meet the needs of users from industry, public authorities and scientists. The objectives are achieved through:

- A portal that allows
  - Access to marine data from measurement stations and ferryboxes. Both near real-time and archived data of time series are to be made available.
  - Metadata for these data sets using EMODNet/INSPIRE standards
  - Metadata maps and overviews for whole seabasins showing the availability of data and monitoring intensity of that basin.
- Monitoring and reporting of the effectiveness of the portal in meeting the needs of users in terms of ease of use, quality of information and fitness for purpose of the product delivered.

EMODnet Physics aims to contribute to the broader initiative 'Marine Knowledge 2020', and in particular to the implementation of the European marine monitoring programme and marine services (GMES). It is based on a strong collaboration between EuroGOOS associates and its regional operational systems (ROOSs), MyOcean and SeaDataNet consortia. The portal also respects INSPIRE standards for discovery and access and it is operational 24 hours a day, 7 days a week, and provides information and tools to potential users (managers, policy makers, researchers, specialized users) in fact it:

- ❍ Gives access to thematic monitoring data that can be queried/selected
- ❍ Gives access to monitoring observations
- ❍ Provides data to GMES, researchers and specialised users

The EMODNet PP is providing access to the following types of measurements:

- wave height and period;
- temperature of the water column;
- wind speed and direction;
- salinity of the water column;
- horizontal velocity of the water column ;
- light attenuation;
- sea level.

The geographical coverage includes all the maritime regions with the following expected input providers:

- the Western Mediterranean Sea (ROOS Med and SeaDataNet);
- the Adriatic Sea (ROOS Med and SeaDataNet);
- the Ionian Sea and the Central Mediterranean Sea (ROOS Med and SeaDataNet);
- the Aegean-Levantine Sea (ROOS Med and SeaDataNet);
- the Greater North Sea, including the Kattegat, and the English Channel (ROOS NWS (North West Shelf) and SeaDataNet);
- the Celtic Seas (ROOS IBI and SeaDataNet);
- the Bay of Biscay and the Iberian Coast (ROOS IBI and SeaDataNet);
- the Atlantic Ocean, the Macronesian bio-geographic region, being the waters surrounding the Azores, Madeira and the Canary Islands (Coriolis and SeaDataNet), and Cape Verde (E subtropical Atlantic)

Parameters Group	Number of Stations
Currents	22
Light attenuation	2
Sea levels	267
Sea water salinity	45
Sea water temperature	132
Waves and winds	165

Tab. 1 - Number of stations and available parameters

	Currents	Light attenuation	Sea levels	Sea water salinity	Sea water temperature	Waves and winds	Number of Stations
Artic/Barrents/Greenland/Norvegian Sea			3				3
Atlantic/Bay of Biscay/Celtic Sea	11	1	103	30	60	53	258
Baltic Sea			74	1	28	8	111
Black Sea					1		1
Mediterranean Sea	6	1	26	5	29	38	105
North Sea	5		61	9	14	66	155
Number of stations	22	2	267	45	132	165	

Tab. 2 - Number of stations and available near real time data per sea basin

- North Atlantic (including Porcupine Abyssal Plain, Central Irminger Basin, Norvegian Sea) (ROOS Artic and SeaDataNet)
- the Baltic Sea (ROOS BOOS and SeaDataNet)
- the Black Sea (ROOS Black Sea and SeaDataNet)

In two years of activity, by means of joint acivities with its pillars EuroGOOS, SeaDataNet and MyOcean, EMODnet Physical Parameters was able to connect about 400 stations that provides:

New EU efforts to assess the state of the marine environment: the EMODnet Chemistry pilot project

Matteo Vinci, OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale)
mvinci@ogs.trieste.it (Italy)

Alessandra Giorgetti, OGS, agiorgetti@ogs.trieste.it

Alberto Brosich, OGS, abrosich@ogs.trieste.it

The EU Green Paper for Marine Knowledge 2020 highlighted that Seas and Oceans that surround Europe provide an essential part of our wealth and well-being but they are also under huge pressure from human activities and climate change, they offer opportunities that we have to develop in a sustainable way. Central to this strategy is the concept to develop a European Marine Observation and Data Network (EMODnet), a network of marine organizations that would provide a single entry point for accessing and retrieving marine data derived from observations, surveys or samples from the hundreds of databases maintained on behalf of agencies, public authorities, research institutions and universities throughout Europe. EMODnet started as a pilot project with four thematic service contracts for a final operational European Marine Observation and Data Network, launched by the Directorate-General for Maritime Affairs and Fisheries (DG MARE).

EMODnet Chemistry pilot project aims to assemble fragmented and inaccessible marine data into interoperable, continuous and publicly available data streams for complete maritime basins, focusing on the marine data groups of chemicals required for the monitoring of the Marine Strategy Directive: pesticides, antifoulants, pharmaceuticals, heavy metals, radionuclides, fertilizers, organic matter, hydrocarbons including oil pollution.

It concerns the following geographical regions:

- North Sea;
- Black Sea;
- 5 spots of the Mediterranenan Sea.

EMODnet Chemistry is focused on multidisciplinary interoperability by adopting and adapting existing and well established data sharing initiatives and being at the same time compliant with the EU INSPIRE directive. Great attention was devoted to the definition of data and product viewing and downloading services to fit the purpose of the Marine Strategy Framework Directive. The Chemistry Lot was developed according to the principle of adopting and adapting the SeaDataNet V1 infrastructure. This because SeaDataNet is a “de facto” European open infrastructure that can give access to a continuously increasing number of data centres across sectors and countries, increasingly meeting the standards needed for INSPIRE compliance. It is an efficient distributed Marine Data Management Infrastructure conceived for the management of large and diverse sets of Physical Oceanography data deriving from in situ and remote observation of the seas and oceans which is developing its standards and tools and now adopted by a growing number of other communities.

As an European platform building upon SeaDataNet, the European Marine Observation Data Network - EMODnet could provide a solid framework for the structured development of a network of distributed data centres using a common lexicon and ensuring broad accessibility for diverse users, from scientists to policy makers, as well as provide user-friendly assembling tools.

The Chemistry Lot adopted SDN Standards for metadata , data and products as:

- metadata CDI (xml ISO 19115);
- Standard Vocabularies (P021, P011, P061...) for common terms;
- ODV data format for background data exchange.

One of the main challenges of the activities of the three year Chemical lot Pilot Project was represented by the Data complexity and heterogeneity management. The Lot had to manage data coming from 8 groups of compounds measured in three matrices (sediment, water column and biota). Seventeen parameters were selected for the data-products generation. The data collected presented a high heterogeneity in all the 3 matrices considered, in relationship with the sampling methods, the data distribution (coastal time series stations vs homogenous sampling at basins level), the analytical methods (instrument, method, target species, target basis, grain sizes). The heterogeneity was also depending on the geographical distribution of the target species considered in the data collection. This is the case of the Mytilus sp that in the Mediterranean Sea was mainly represented by Mytilus galloprovincialis, while in the North Sea area was represented by Mytilus edulis.

The complexity on the management of this information was approached by adapting:

- the Metadata and data description:

the SeaDataNet infrastructure, that was created to handle mainly physical oceanographic data, was adapted to fit it to the management of chemical data which need a more detailed metadata description. In fact, in order to be able to store and handle chemical data also after several years from their measurement and to have the necessary information to be able to compare data from different areas, more details on analytical methods, target species, target basis are required.

- Products generation and handling:

suitable products were defined to represent the different features of the datasets collected and to satisfy the needs of a standard for reporting environmental data.

Data with a homogeneous distribution in time and space, collected at basin level were used to generate seasonal and annual interpolated maps for nutrients, metals and radionuclides in the North Sea, in the Black Sea and in selected areas of the Mediterranean Sea. Data collected at coastal stations repeated in time, were used to generate time series plots for hydrocarbons, pesticides, metals, fertilizers over more than 160 stations.

Through the EMODnet portal it is possible to visualize the maps of selected parameters, the pre-computed time series plots and the original data. The objective of the obtained results and of future efforts is the development of an efficient infrastructure for the data collection, management and analysis and then for the integrated assessment of the marine environment.

DataMARES: Marine Access Repository for the Gulf of California, Mexico

Marcia Moreno-Báez, Scripps Institution of Oceanography, University of California San Diego, mmorenobaez@ucsd.edu

Obregón-Noriega, V., Centro para la Biodiversidad Marina y la Conservación, A.C., vobregon@datamares.org

Sánchez-Rodríguez, A., Centro para la Biodiversidad Marina y la Conservación, A.C., alexandra@gocmarineprogram.org

Aburto-Oropeza, O., Scripps Institution of Oceanography, University of California San Diego, maburto@ucsd.edu

In Mexico, research institutions are generating a large body of well-inventoried spatial information such as geo-referenced and tabular databases. However, this data is often hard to find with only a few of the essential datasets having been compiled, analyzed, and distributed at an appropriate spatial or temporal scale. This lack of availability and accessibility of information hampers effective planning and decision-making. Using technology to enhance the speed of data process and accessibility, is possible to reshape the way we compile, distribute and access relevant information, facilitating the communication of scientific information to any interested to study or preserve the coastal and marine ecosystems. The idea of “open science” (Reichmann et al., 2011), which makes scientific research, data and dissemination accessible to all levels of the society, could facilitate and improve access to information with better transparency and reproducibility of results. Additionally, this accessibility to science promotes collaboration among scientists, decision-makers and stakeholders and brings opportunities to resolve trans-disciplinary problems related to the coastal and marine ecosystem.

We created a free, open source interactive platform called DataMARES (Ocean Data in Spanish) (Fig. 1) to address the need for accessible, secure, and robust data. A spatial display interfase and tools with the most updated functionalities (Google engine) were used to geographically visualize relevant information in a user-friendly format (e.g., graphs, images, videos). Different technologies were used such Ubuntu Linux server, MySQL DBMS as its back-end, PHP as its API provider, MapServer as a mapping application and Google feature technologies in the front-end. We developed a beta version of the portal, and thus are well on our way to having a publicly available resource within the life of this project. Additionally, we build a tool called EcMonitor (Fig. 2) to capture and view all information collected by users. The software was designed based on the requirements of our monitoring program but the idea is to adapt the tool for any other data collection processes held by scientists. EcMonitor facilitates the capture and access to information, which can be exported into other software for analysis.

DataMARES and EcMonitor are an open forum through which everyone interested in the preservation of coastal and marine resources of the Gulf of California, have equal opportunity to disseminate scientific information. Ultimately, the DataMARES initiative aims to empower anyone interested in science and costal and marine natural resources and increase their participation in the production, broadcasting and use of scientific research. We believe that access to scientific information will increase stakeholder participation in conservation actions and produce a broader understanding regarding the importance of incorporating scientific knowledge in policy making. Also, the participation of different scientific institutions opens the opportunity to build a collaborative center of information and promotes open science in the region.

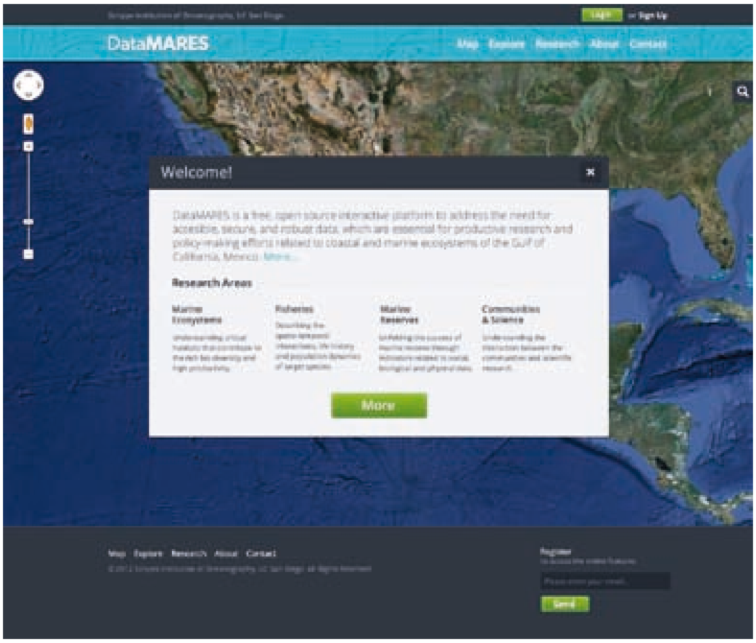


Fig. 1 - Screen shot of DataMARES websiteavailable near real time data per sea basin.

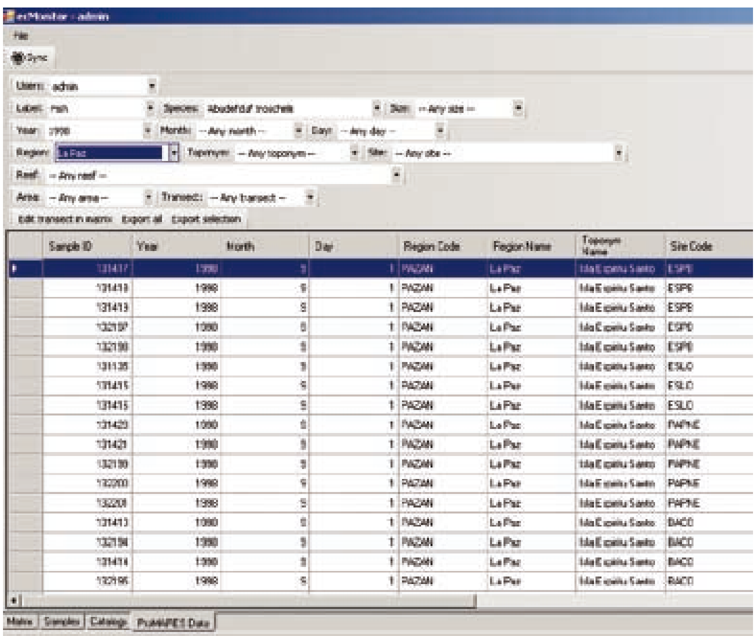


Fig. 2 - ecMonitor software that transfers the monitored data into DataMARES.

Data Management and applications at SOCIB

David March, IMEDEA (CSIC-UIB), david@imedea.uib-csic.es, (Spain)
Sebastian Lora, SOCIB, slora@socib.es
Joan Pau Beltrán, SOCIB, jbeltran@socib.es
Kristian Sebastian, SOCIB, ksebastian@socib.es
Sonia Gómara, SOCIB, sgomara@socib.es
Joaquín Tintoré, SOCIB and IMEDEA (CSIC-UIB), jtintore@socib.es

SOCIB Description

SOCIB, the Balearic Islands Coastal Ocean Observing and Forecasting System (http://www.socib.es), is a Marine Research Infrastructure, a multiplatform distributed and integrated system, a facility of facilities that extends from the nearshore to the open sea and provides free, open and quality control data. SOCIB is a facility o facilities and has three major infrastructure components: (1) a distributed multiplatform observing system, (2) a numerical forecasting system, and (3) a data management and visualization system. We present the principles, major components and actions implemented in the 2010-2013 period by SOCIB Data Centre, also providing some examples of on-going developments.

Data Centre Facility Goals

The Data Centre is the core of SOCIB. The general goal of SOCIB Data Centre is to provide users with a system to locate and download the data of interest (near real-time and delayed mode) and to visualize and manage the information. Following SOCIB principles, data need to be (1) discoverable and accessible, (2) freely available, and (3) interoperable and standardized. In consequence, SOCIB Data Centre Facility is developing and implementing a general data management system to guarantee international standards, quality assurance and interoperability. The combination of different sources and types of information requires appropriate methods to ingest, catalogue, display, and distribute this information.

The data managed by SOCIB mostly come from its own observation platforms, numerical models or information generated from the activities in the SIAS Division. The Data Centre also performs the management of data coming from external data providers through various collaborations.

Data Centre Facility Components

SOCIB Data Centre is responsible for directing the different stages of data management, ranging from data acquisition to its distribution and visualization through web applications. The system implemented relies on open source solutions, following other architectures adopted within the context of marine spatial data infrastructures. In other words, the core Data Centre system main components are:

An instrumentation application developed at SOCIB to manage all platforms centralized in a database through a web interface.

- A processing application developed at SOCIB to deal with all collected platform data performing data calibration, derivation, quality control and standardization.
- A THREDDS catalogue implemented at SOCIB to archive data and distribute them through services such as OPeNDAP, OGC services, HTTP and ncISO.
- A layer of RESTful web services developed at SOCIB to ease the development of both internal and external applications, such as web or mobile applications.
- A set of tools for data visualization and real time monitoring developed at SOCIB

The general structure and some of the components are illustrated in Fig. 1 and some specific examples of visualization developments are illustrated in Fig. 2.

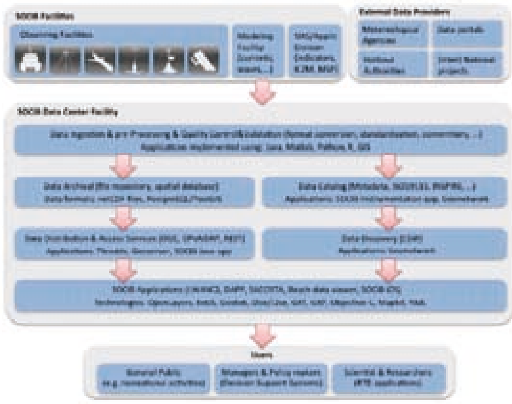


Fig. 1 - Data Centre Conceptual Structure and SOCIB developed applications.

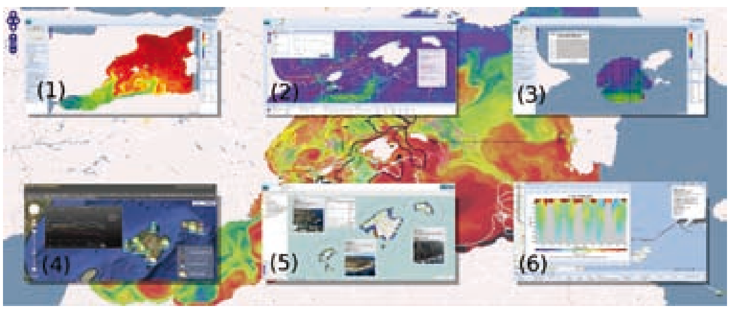


Fig. 2 - Data and applications developed at SOCIB: (1) web-based map viewer for numerical models output; (2) web application for mobile platforms (e.g.,gliders, Argo profilers, drifters, etc.); (3) HF radar output; (4) fixed stations web section; (5) web-based map viewer for cartographic data, ESI, etc.; (6) glider real-time monitoring tool.

Conclusions

The organizational and conceptual structure of SOCIB's Data Centre and the components developed provide an example of marine information systems within the framework of new ocean observatories and/or marine research infrastructures. The main principles of standardization, interoperability and open access are in line with the INSPIRE Directive and other important European initiatives such as Emodnet or Copernicus, which are also the information pillar of the Integrated Maritime Policy and the Marine Strategy Framework Directive.

XBT Data Management and Quality Control in Japan

Toru Suzuki, Marine Information Research Center, suzuki@mirc.jha.jp (Japan)
Masayoshi Ishii, Meteorological Research Institute, maish@mri-jma.go.jp
Toshiya Nakano, Japan Meteorological Agency, nakano_t@met.kishou.go.jp
Hiroyuki Yoritaka, Hydrographic and Oceanographic Department, Japan Coast Guard
Akihiro Seta, Japan Oceanographic Data Center
Shoichi Kizu, Tohoku University, kizu@pol.gp.tohoku.ac.jp
Kanako Sato, Japan Agency for Marine-Earth Science and Technology, k_sato@jamstec.go.jp
Yugo Shimizu, Fisheries Research Agency, yugo@affrc.go.jp
Ken-ichi Amaike, Tsurumi-Seiki Co., Ltd., amaike@tsk-jp.com
Toshio Suga, Tohoku University, suga@pol.gp.tohoku.ac.jp
Yutaka Michida, Atmosphere and Ocean Research Institute, the University of Tokyo, ymichida@aori.u-tokyo.ac.jp

The expendable bathythermograph (XBT) is a device for obtaining a vertical profile of temperature from a moving ship and has been used by oceanographer for long time. Temperature is measured with a thermistor within an expendable probe but depth is determined by fall-rate at probe sinks with constant time intervals. Japanese marine science organizations and agencies have submitted to Japan Oceanographic Data Center (JODC) about 270 thousand XBT casts and about 14 million temperature data since 1970. Recent studies pointed out that the time change of estimated ocean heat content involves a bias against natural variability which is caused by various errors involved in XBT data.

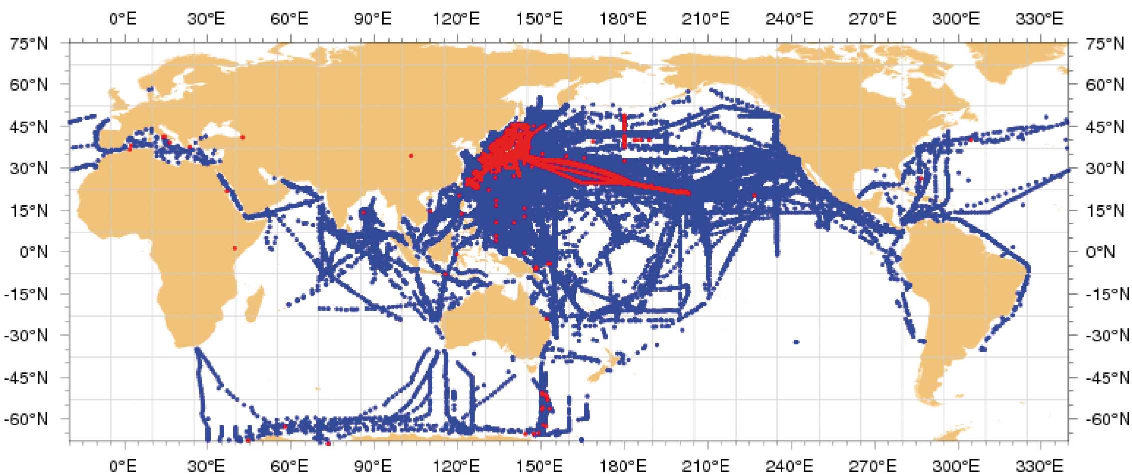


Fig. 1 - Station map of all XBT casts archived in Japan Oceanographic Data Center. Red circle indicates error in date/time/position record or error of sea/land check.

For this reason we established a working group on reconstruction of XBT database under the national committee for promotion and cooperation with IOC and have started to search and rescue of XBT data. Primary quality control procedure as same as World Ocean Database, i.e., checks of date/time/position, temperature range check by basin as a function of depth (e.g., see thick lines in Fig. 2) and excessive inversion/gradient, have already been processed. In addition, estimated ship speed check has also been done in order to seek out errors by not only date/time/position but also land/sea effectively. Some depth error were found by comparison with gridded bathymetry data or rated depth range for each probe type.

Some depth exceed errors in historical data were caused by mis-transcription from field note or fail to format conversion. We will search historical and high resolution XBT data and also metadata as much as possible following concept of Global Data Archaeology and Rescue (GODAR) project recommended by IODE.

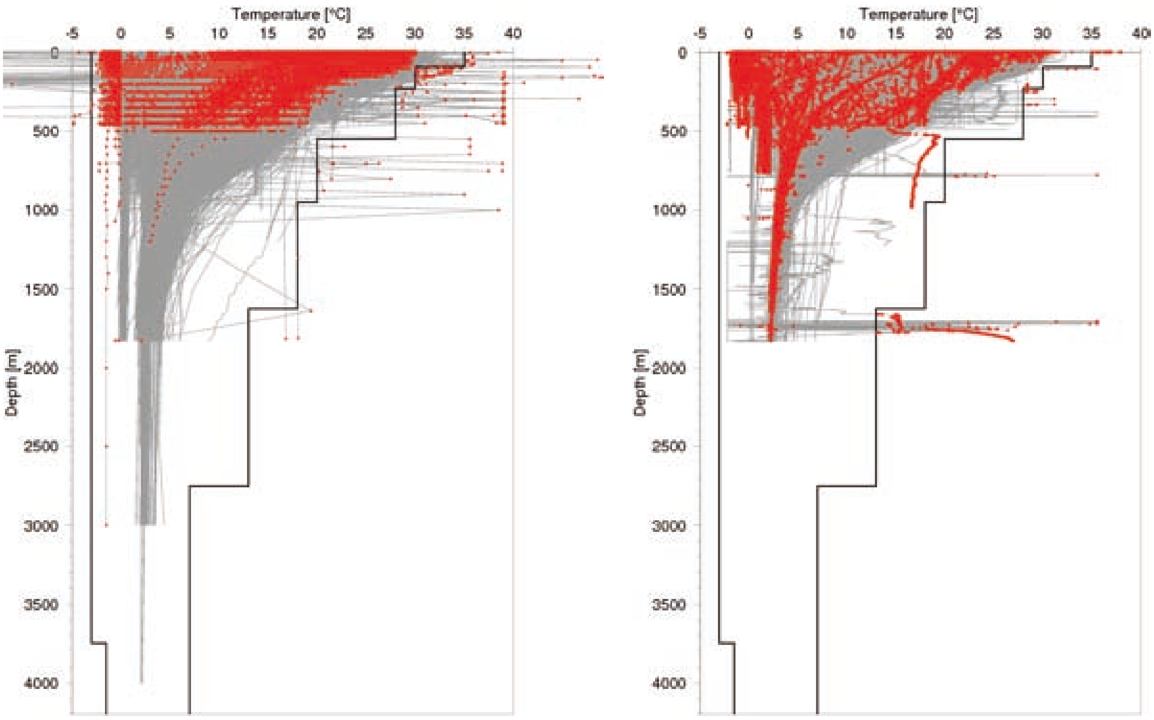


Fig. 2 - Profiles of XBT casts: (left panel) standard depth data, (right panel) high resolusion (1m intervals) data. Red circle shows questionable or bad value by range and/or excessive gradient/inversion check and thick line shows temperature acceptable range as a function of depth in the North Pacific Ocean same as WOD09. Note that profiles are included in all basin.

The German Marine Data Infrastructure (MDI-DE)

Johannes Melles, Federal Maritime and Hydrographic Agency (BSH), johannes.melles@bsh.de (Germany)

Rainer Lehfeldt, Federal Waterways Engineering and Research Institute (BAW), rainer.lehfeldt@baw.de

Current questions related to changes to marine ecosystems, global warming and the resulting requirements for costal protection as well as reporting obligations of the European Union with respect to effective framework directives (INSPIRE, MSFD etc.) require interdisciplinary access to the related spatial data. Therefore a novel "Marine Data Infrastructure" has been established in Germany. Coastal and marine data collected by 11 Federal and State agencies are made available by OGC compliant Web services and documented with metadata according to the ISO standard. A new Web portal serves as central entry point for data and information from the German coastal zone and the adjacent marine waters. This facilitates intersectoral views of resources by providing technological solutions of networking and distributed data management.

The integrated national marine and coastal information system was set up within the co-operative project "Marine Data Infrastructure (MDI-DE)". Each participaiting agency or institute operates a node to provide metadata and services such as WMS, WFS etc. Such a node could easily be built based on existing Open Source or Commercial of the Shelf Software. The benefit of hosting the data locally is that the data from different sources could be merged in almost any way, custom-made compositions of thematic data layers can be compiled without touching the data itself.

To achieve such a seamless integration, each node within this network needs to be equipped with a few basic components: services to provide the data and metadata and a database to feed the two. For the provision of spatial data, the Open Geospatial Consortium (OGC) has developed a number of open and international standards. The most appropriate in this context are the Web Map Service (WMS) to generate and visualize digital maps in the Web and the Web Feature Service (WFS) to download the data in an interoperable form such as GML. For both several software packages in the Open Source an commercial sector are available which could easily be used. The underlying database needs to support these services. This could be achieved by either installing a spatial database dedicated for this purpose or adding a data view to an existing database to adapt existing structures to the needs of the services. After thus ensuring the data distribution, the last piece of the puzzle is the corresponding metadata, that again is available through a specialized database or a data view and provided in the Web by any metadata management tool that offers a standardized Catalogue Service for the Web (CS-W) interface.

Within the MDI-DE Portal such catalogue will act as a harvester, collecting the metadata stored in all the sub-catalogues situated at the various nodes which are connected to the main catalogue through the network. A search on the central catalogue will find all relevant datasets from all associated nodes.

The central entry point for data and information from the German coastal zone and the adjacent marine waters is the MDI-DE Portal (www.mdi-de.org). Connecting the local nodes it will provide a wide range of data coming from different disciplines working in the marine environment. The MDI-DE Portal will be maintained by the Federal Maritime Agency (BSH) in Hamburg.

In short, the MDI-DE will provide the following functionalities:

- Simple and expert search using a thesaurus for controlled vocabulary and a coastal gazetteer for geographic names in the marine environment,
- Intersectorally consistent data structures for interdisciplinary views on marine data,
- OGC Web Services for visualization and data access,
- Prototype services for data analysis and automatic report generation to meet the requirements from the EU framework directives.

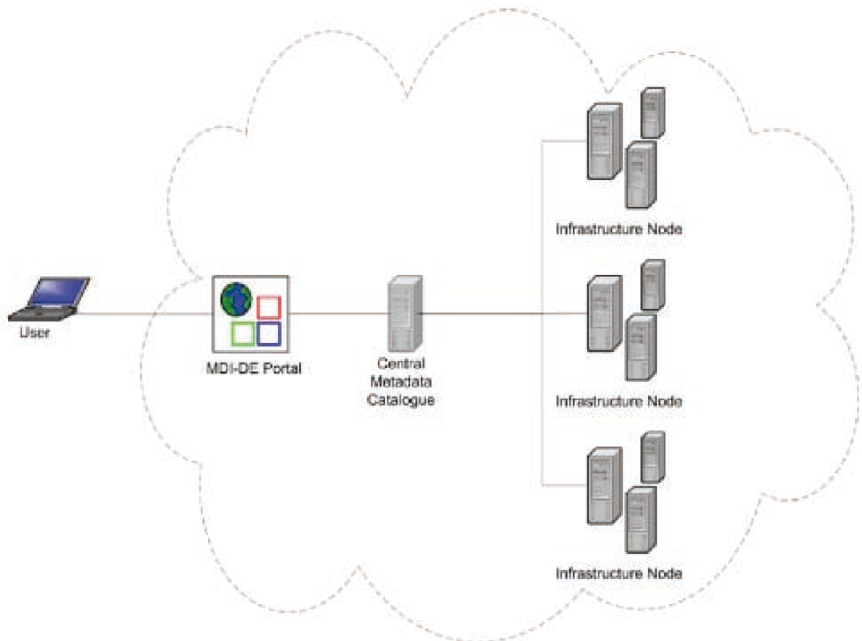


Fig. 1 - Scheme of the MDI architecture.

Beyond the basic services like WMS and WFS other services like Web Processing Services (WPS) for data analysis or the Sensor Observation Services (SOS) for the provision of sensor data from the German marine environment will be incorporated into the the MDI-DE network.

MDI-DE represents not only a state-of-the-art spatial data infrastructure, it also aspires to lose the ‘spatial-only’ attribute and includes other related data into this distributed data source. The future Web-Services provided by MDI-DE will support system analysis applications related to coastal engineering, spatial planning, nature conservation, science and ecology.

We would like to thank the German Federal Ministry of Education and Research for funding the co-operative project “Marine Data Infrastructure (MDI-DE)” from July 2010 until December 2013. Altogether, 11 Federal and State Agencies, which are in charge of coastal engineering and coastal water protection, marine environmental protection, marine nature conservation and accompanying scientific / technical research are involved.

Integrated Ocean Data and Information Management System Oceanographic Data Center of IO PAN

Marcin Wichorowski, Institute of Oceanology Polish Academy of Sciences, wichor@iopan.pl (Poland)
Sławomir Sagan, Institute of Oceanology Polish Academy of Sciences,sagan@iopan.pl

Oceanographic data at IO PAN

Oceanographic data originated from worldwide experiments have growing importance and influence decisions processes and policymakers. Accessibility to information for decision makers is crucial for leading such processes.

All organizations involved in oceanographic research and data providers are responsible for maintaining infrastructure for data storing, processing and running data providing services.

Institute of Oceanology (IO PAN) maintains huge archives of information gathered during research activities performed for more than 50 years. These data, originating from different sources are being processed according to common practice elaborated through generations of oceanographers, however there appears to be a very strong demand for standardization of procedures, especially while facing problems regarding cooperation in data exchange.

Fig. 1 - IO PAN Integrated Ocean Data and Information Management System (ZSPDO).

Development of ODC

To enable cooperation in data exchange on higher level IO PAN has decided to develop data center and deploy system for management of data and information, using well defined and widely used standards of data processing. In consequence IO PAN developed project Integrated Ocean Data and Information Management System (Zintegrowany System Przetwarzania Danych Oceanograficznych - ZSPDO). The project goals were defined as follows:

- to build ocean data repository with data management and processing system ensuring data availability, increasing data accessibilty providing catalogue services and reducing risk of data degradation and data loss;

- to deploy management system for research and development projects and works performed by Institute according to standards and identified processes;
- to deploy management system for resources used and people engaged in research and administrative processes optimizing human work, accessibility and utilization of the resources;
- to develop eLearning platform and data visualisation services providing tools for teachers and trainers at different levels of education; to deliver platform for data and information delivery services;
- to build organizational and technological infrastructure (ensuring development, management and coordination of national and international projects from the marine research domain; data, information and metainformation provisioning for marine research institutes and coastal, environment and marine management organizations; industrial R&D organizations active in shelf, coastal land and river estuaries areas involved in exploitation of marine environment; national and international scientific institutions; administration and authorities at local, regional and national level; others interested in) and
- to increase possibility of hosting data and services at high quality level.

The forefront idea of ZSPDO is facilitating oceanographic data management processes considering their complexity, their heterogenouos nature and demand of interoperability preserved for external systems. ZSPDO enables users to define their own formats of raw data and mapping from native formats to vocabulary parameter and metadata definitions and store them in the data mart for further processing, eg. Performing multidimensional analysis. In spite of data mart functionality, users can also merge raw data and big binary files with metadata and archive files untouched. All data resources have defined access rights, and fields of explotation according to the policy and information delivered with metadata.

Functionality of the system is enhanced with samples management, cruise management, project management, laboratory and equipement management, reporting tools and others. Videoconferencing module allow for interaction between remote users and provide helpdesk functionality. New approach to training and teaching activities is possible using advanced technologies and eLearning platform. Multimedia materials published for educational purposes are easily accessible through the electronic communication channels for dissemination of information as well. Development of tools enabling both of processes is one of the objectives of ZSPDO.

Project has reached key expected results:

- ensurance of security of agregated information resources,
- improvement of data exchange with international organisations,
- automatization of metainforamation discovery and publication,
- normalization of data and information exchange procedures deployed in widely used standards (eg. these implemented within SeaDataNet and other leading initiatives),
- improvement of research project management with developed tools for budget and resources administration
- monitoring of the administration processes and quality management within projects,
- improvement of efficiency of research works providing fast and easy access to data repository.

Data Center Infrastructure

ZSPDO is combination of advanced data processing system and top level technical infrastructure. Software part of the system is build with application server Glassfish using cluster of MySQL database server, and Lotus Notes groupware. System procedures are controled and managed through graphics interface using BPMN notation for processes modelling. Technical infrastructure is build with blade servers technology assuring High Availability access and enhanced with Tesla GPU accelerators. ZSPDO provide for users computing power (more than 6Tflops for single precision computations, 2048 CUDA cores) and storage infrastructure with more than 600TB allocated at tape library and hdd matrix. This infrastructure is replicated at supporting data center located at Tricity Academic Supercomputer and Networking Center (TASK) – partner of the PLATON project providing data storage in the cloud. Data centers are connected with dedicated 10Gbps connection integrating facility provided by TASK with services dedicated for oceanographic community available at ZSPDO.

Conclusion

Growing computation power and storage eqiupment capabilities give oportunity of increasing data volume archived and processed in data centers. This trend is well known as big data paradigm in financial and bussines applications. Exploding data volume also lead to demand of new services eg. mutlidimensional analysis and data visualisation, integrated in unique platform. ZSPDO have been developed to face these identified chalenges and raise operational performance of data processing.

GOSUD: Global Ocean Surface Underway Data From gathering underway data to providing high quality delayed mode data

Loïc Petit de la Villéon, Ifremer (France), petit@ifremer.fr
Thierry Delcroix, Legos OMP, (France), Thierry.Delcroix@legos.obs-mip.fr
Gael Alory, Legos, OMP (France), Gael.Alory@legos.obs-mip.fr
Fabienne Gaillard, Ifremer (France), Fabienne.Gaillard@ifremer.fr
Magali Krieger, Ifremer (France), Magali.Krieger@ifremer.fr

The goal of the GOSUD Project is to develop and implement a system for collecting and managing surface ocean data that are acquired by vessels when they are underway. At the moment, the parameters considered are salinity and temperature.

The Project was set up in 2001, since then the main efforts have been put on gathering data, identifying contributors and setting up a data management infrastructure. To reach this objective, a GDAC –Global Data Assembling Centre- hosted by Coriolis –France- was set up with a mirror site back up facility provided by US-NODC. NetCDF formats have been developped (last version V3 is being implemented at the GDAC level.

In 2011, data from 91 vessels were collected and assembled by the GDAC and in 2012 data from only 81 vessels were gathered. The data are collected either on board research vessels or on board merchant ships. When collected on board merchant vessels the quality of data is better when the management is done within an identified network rather than individual contributions

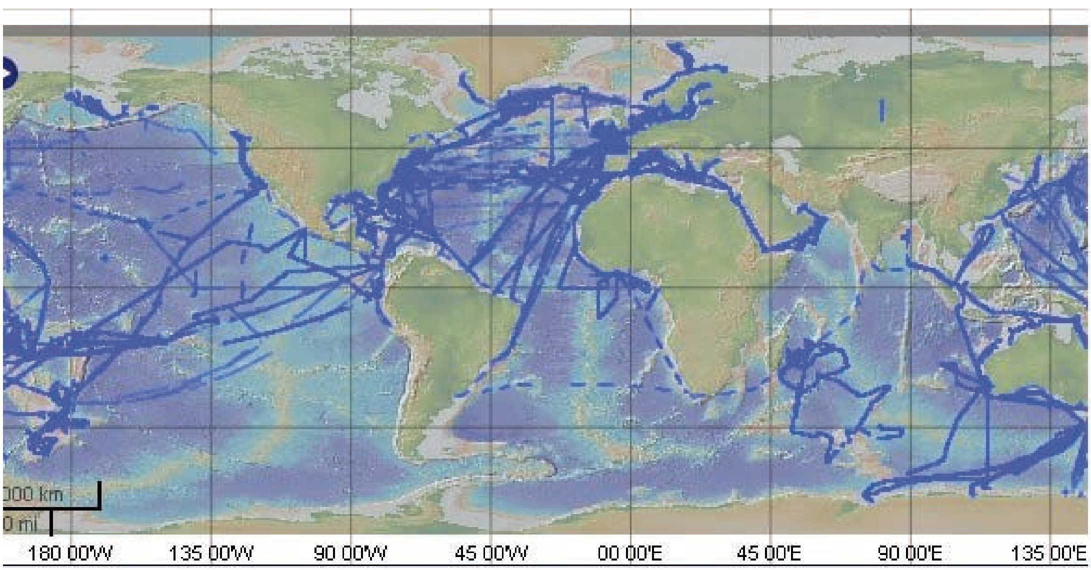


Fig. 1 - The GOSUD network 81 vessels in 2012.

The GOSUD data circulate either in near real-time or in delayed-mode. The two users communities are very different depending on the way they access the data. Access to NRT data is useful for operational ocean modelers,, satellite validation centers and users in charge of monitoring the data acquisition. The Science community (process studies and global climate research) needs carefully checked delayed mode data..

While the data collection and management systems were built up within GOSUD, a consortium of scientists and engineers (IRD, CNRS, Ifremer) have been working on the validation and correction methods that were later integrated in the TSG-QC (A tool for interactive quality control of sea surface temperature and salinity) software developped by IRD which enables to quality control the GOSUD data and to produce a science quality reference dataset.

Using the data received in near real-time or when the ship arrives in the port and with the help of the salinity samples taken with bottles or / and data collected by neighbouring instruments such as Argo, it is possible to evaluate the drift of the conductivity sensor and to propose salinity values adjustments that could fit the water samples analysis results and that take into account the calibration coefficients of the instruments. This software has been developed using Matlab and is freely distributed by IRD (http://www.ird.fr/us191/spip.php?article63).

The first delayed mode dataset is available on http://www.gosud.org

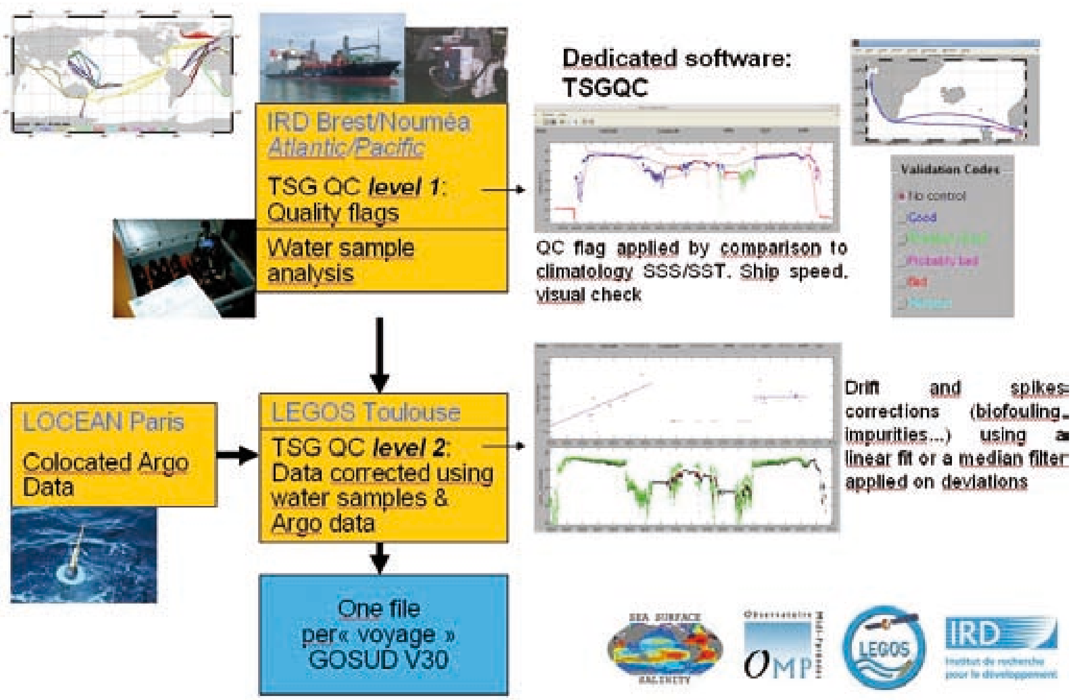


Fig. 2 - GOSUD elaboration of the delayed mode dataset.

This GOSUD presentation will report on the work performed in the frame of GOSUD and will focus on the delayed mode dataset that is now available

Regional project (PERSEUS) oriented system for storage and operative exploration of Mediterranean and Black Sea cast data

Isaac Gertman, Israel Oceanographic & Limnological Research, isaac@ocean.org.il (Israel)
Yevgeniya Krivenko, Israel Oceanographic & Limnological Research, yevgeniya@ocean.org.il
Tal Ozer, Israel Oceanographic & Limnological Research, tal@ocean.org.il
Vladimir Belokopytov, Marine Hydrophysical Institute, belo@vip.sevsky.net
Andrey Ingerov, Marine Hydrophysical Institute,ingerov@rambler.ru
Alexey Khaliulin, Marine Hydrophysical Institute, khaliulin@yahoo.com

PERSEUS is large scale European research project (54 institutions) intended to identify the interacting patterns of natural and anthropogenic pressures in the Mediterranean and Black Sea, assess their impact on marine ecosystem to design research governance framework.

The 4 years project stared in 2011. In order to expedite data availability for project partners, we have created a dedicated multiparametric oceanographic database (DB) to collect all available historical data of Mediterranean and Black Sea observed as vertical profiles (excluding XBT data).

The DB allows online selection and download of any part of the data to a wide community of users with different levels of data accessibility (defined according to PERSEUS data policy). In addition, the DB includes tools for importing new data which will be gathered in the framework of the PERSEUS project.

Fig. 1 - GIS like on line interface for query design.

As the primary source for the historical data the SESAME cast data base (Gertman et al., 2010) was used.

In contrast to the old SESAME system, where the primary DB was MS ACCESS DB, the new system has a primary MS SQL DB which includes direct online interface for data import and QC of new data in ODV format. The identification of parameters in the PERSEUS system is based completely on the Common Vocabularies which were introduced as standard terms by SeaDataNet: P011 (parameter usage vocabulary), P061 (data storage units) and P021 (parameters discovery vocabulary). GIS like on line user interface provides flexible data selection and analysis of metadata (Fig. 1).

The system supports two ways to export and download selected data:

- Cruise by cruise in ODV format. Parameters and units are the same as in the original submission.
- Aggregate of cruises in MS ACCESS DB format. Parameters and units are selected by the user.

In the middle of 2013 the PERSEUS cast DB contained about 192,000 casts with 196 different parameters. All historical data assembled during SESAME project have become public available on 01/APR/2013 and can be downloaded by any registered user.

Two autonomous relocatable applications developed by MHI co-authors are available for tuneable quality control and scientific analysis of the data exported to MS ACESS DB (of aggregated data). These applications are similar to those which were developed and used successfully during the SESAME project (Gertman et al., 2010).

References

Gertman I., Katsenelson B., Lunin I., Belokobytov V., Khaliulin A. (2010). *Multidisciplinary marine cast data management system for SESAME EU project. Inter. Conf.on Marine Data and Information Systems (IMDIS 2010)*, 29-31 March 2010, Paris, France, 73-74.

Monitoring Long Term Variability in the Eastern Mediterranean

Devrim Tezcan, Institute of Marine Sciences-METU, devrim@ims.metu.edu.tr, Turkey
Volodymyr Myroshnychenko, Institute of Marine Sciences-METU, volodymyr@ims.metu.edu.tr, Turkey
Elif Yılmaz, Institute of Marine Sciences-METU, elify@ims.metu.edu.tr, Turkey
Barıs Salihoglu, Institute of Marine Sciences-METU, baris@ims.metu.edu.tr, Turkey
Zahit Uysal, Institute of Marine Sciences-METU, uysal@ims.metu.edu.tr, Turkey
Ahmet E. Kıdeys, Institute of Marine Sciences-METU, kideys@ims.metu.edu.tr, Turkey
And DEKOSIM team

To contribute to the regional ecosystem and climate research, the Institute of Marine Sciences (IMS-METU) set up a long term monitoring program for the Eastern Mediterranean. This program is being carried out under DEKOSIM (Centre for Marine Ecosystem and Climate Research). DEKOSIM is an interdisciplinary centre of excellence funded by the Turkish Ministry of Development, established at the Institute of Marine Sciences of the Middle East Technical University located on the Eastern Mediterranean coast of Turkey (Fig. 1). The main objective is to create an infrastructure that focuses on interdisciplinary research linking physical, chemical, biological and geological marine research. The majority of the IMS-METU academic staff as well as several researchers from other national institutions have contributed to the program. With 35 years experience in marine research in the Mediterranean and Black Sea the IMS-METU through DEKOSIM aims to become a marine research centre serving all national and international researchers in the marine ecosystem and climate fields.

DEKOSIM consists of two main frameworks: Long term monitoring systems and a marine science laboratory. The long term monitoring system involves the deployment of 4 Argo floats in the Black Sea and 2 Argo floats in the Eastern Mediterranean, one coastal mooring system with both surface and underwater sensors to obtain continuous oceanographic, meteorological and atmospheric measurements in the Eastern Mediterranean situated offshore IMS-METU Campus and another in the Black Sea situated offshore Sinop (Fig. 1). The marine science laboratory will comprise several laboratories including, genetics, fisheries, chemistry, geophysics and computer cluster branches.

The Mediterranean mooring system will be set up about 5 nm off the IMS-METU Campus at about 100 meters water depth. The underwater sensors of the system are; T, S, DO, Chl, Tur at near surface, T,S,P at 10 meters, T,S at 20, 30 and 75 meters, T,S,DO at 50 and 100 meters. A 300kHz ADCP will be deployed under the buoy. All the data from underwater sensors will be transferred to the buoy via inductive modems. The meteorological equipment and daily data telemetry via GPRS will be provided by Turkish Met Office (Fig. 2).

Based on the time series measurements since 1997, a new time series project, Erdemli Time Series (ETS) has been started in the Eastern Mediterranean (Fig. 1). Along a section of 10 nm, at 8 stations of 20, 50, 75, 100, 125, 150, 175 and 200 meter water depths, the physical parameters (T,

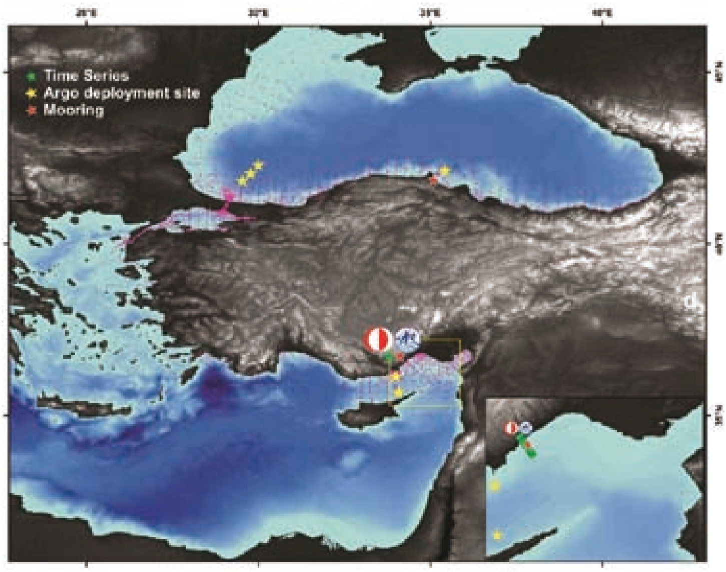


Fig. 1 - Location map of the time series stations, argo deployment sites and mooring stations. Pink dots are the oceanographic stations made by IMS-METU for 35 years.

S, DO, Chl, Tur, Par, Sechi Disk) will be measured weekly and biochemical parameters monthly. As the profile will cross the main Eastern Mediterranean water current, the weekly ADCP measurements will provide important information about the currents and the suspended materials. These data will be analysed to understand the long-term variability (including the climate effect) in the region.

Data Management is an important part of the DEKOSIM activities. The DEKOSIM information system provides data storage and quality control facilities and takes full advantage of the latest web technologies for data sharing and the visualization of data products to support decision making. Particular attention is given to integration of the monitoring installations into the information system thus enabling the transfer of acquired monitoring data from the sensors to the end users in (near) real time. Management and sharing of the spatial data and products as well as supplementary GIS resources is to be organized on base of the ArcGIS Server. The metadata on performed observations will be timely submitted to the SeaDataNet – the pan-European infrastructure for ocean and marine data management.

In the long run DEKOSIM aims to develop and to test analysis systems for the synoptic description of the environmental status of Turkish coastal waters including the Mediterranean, Marmara and the Black Seas. DEKOSIM also aims to provide knowledge transfer tools to assist authorities and other stakeholders to manage routine tasks and evaluate trends.

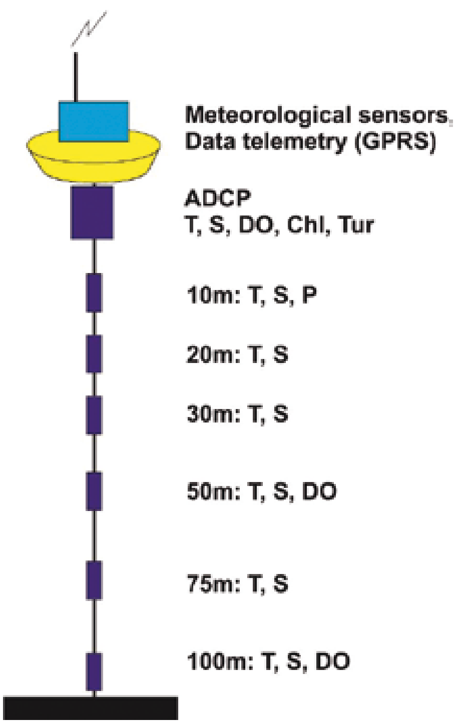


Fig. 2 - The elements of mooring stations.

POSTERS

A new Approach in Data Management for an Integrated Coastal Observation System

Gisbert Breitbach, Helmholtz-Zentrum Geesthacht, gisbert.breitbach@hzg,de (Germany)
Hajo Krasemann, Helmholtz-Zentrum Geesthacht, hajo.krasemann@hzg.de

An integration of different types of data into one common data management system is an ambitioned task for every observation system. In most cases a portal is used which links to the different data types even though they address finally the same parameter (e.g. temperature measured in-situ or as SST from space). In COSYNA (Coastal Observation System for Northern and Arctic seas) a rather new approach is chosen which allows to integrate different types of data into a structured set of parameters using a common view. The base concept of this new approach is a modular structure with web-services as central modules. The access to data and metadata is handled by these web-services which are largely transparent to the users.

In addition a strong focus is laid on standardisation and harmonisation to be interoperable and to simplify data management. Every observed property is mapped to CF standard names. All data are connected to a data level reflecting the maturing grade of the data. For data level 3 which includes world access via the data portal a further standardisation leads to only 2 data formats allowed in COSYNA: NetCDF for more-dimensional data and a relational database system for one-dimensional data and also for metadata. As quality flags, a subset of the quality scheme of SeaDataNet is in use.

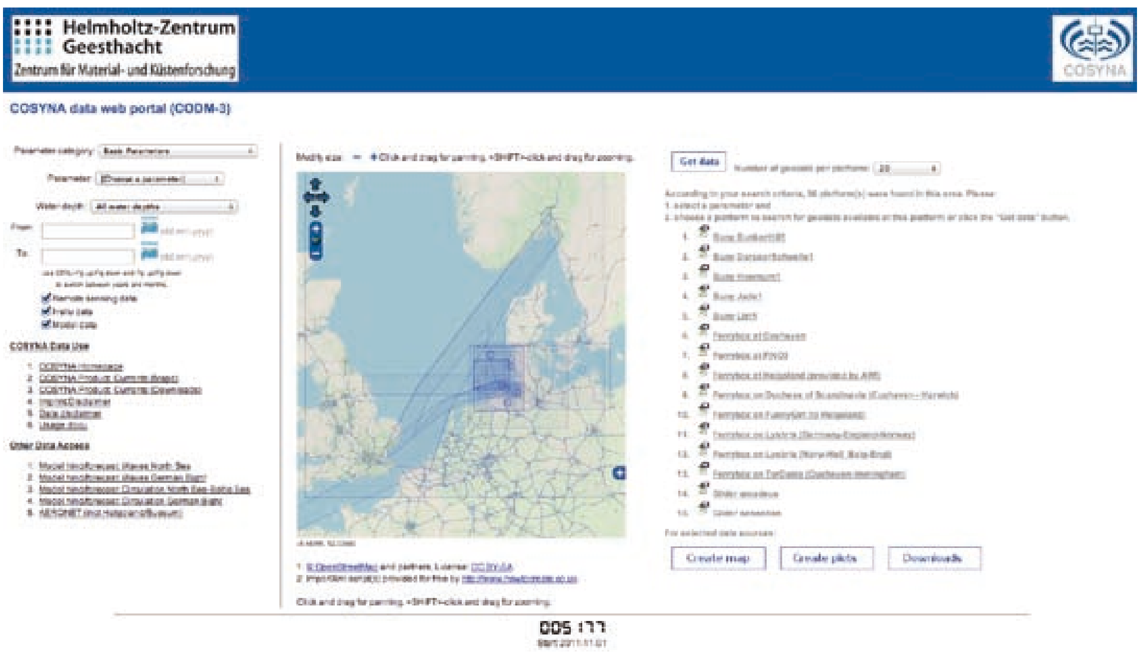


Fig. 1 - Initial view of COSYNA data portal http://codm.hzg.de/codm.

The metadata system of COSYNA is ISO14119/39 and INSPIRE compliant and includes an element 'observedProperty' which describes the parameters. Discovery of data is done using metadata. Search metadata are defined with the help of an user-adjustable map, a selectable time region and the wanted parameter selectable from a pop-up list. For all data information about the platform metadata including sensor information and the metadata of single measurements could be accessed. In addition for all data exists the possibility to download the data directly or to view the numeric data in ASCII. All data could either be visualised as colour coded map or as chart plot. Some data allow both types of visualisation.

Web-service URLs for download, map or plot are added to the descriptions of the observed properties. A dynamic parameterisation of these URLs is described with xml-syntax.

All maps are defined as WMS. For NetCDF files this is ncWMS in other cases WMS-servlets developed for COSYNA are used.

OPeNDAP is the web service for downloading NetCDF files whereas the SOS-GetObservation request is used for downloading other data.

Web-services for chart-plots are yet not standardised and are realised within COSYNA as a specific solution.

The management of maps from different data sources is done using OpenLayers. Maps of Ferrybox or glider transects could be shown together with maps from remote sensing or models.

With the realisation of this approach new possibilities for the visualisation and for comparing data are shown in this talk. The data are available in the COSYNA data portal in near real-time which means within 2-4 hours after measurement.

CORA: in-situ re-qualified dataset at the Coriolis Data Center

Christine COATANOAN, Ifremer/Sismer, christine.coatanoan@ifremer.fr (France)
Cécile CABANES, INSU/CNRS, cecile.cabanes@ifremer.fr
Antoine GROUZAEL, INSU/CNRS, antoine.grouazel@ifremer.fr
Victor TURPIN, MERCATOR Ocean, victor.turpin@mercator-ocean.fr

Coriolis, as part of the French operational oceanographic system, has been especially involved in gathering all global ocean in-situ observation data in real time, and developing continuous, automatic, and permanent observation networks. In this framework, it produces the COriolis dataset for Re-Analysis (CORA) on a yearly basis.

The CORA dataset contains in-situ temperature and salinity profiles from different data types:

- Argo
- GTS data
- VOS ships
- NODC historical data...

The latest release CORA3 covers the period 1990 to 2011. Several tests have been developed to ensure a homogeneous quality control of the dataset and to meet the requirements of the physical ocean reanalysis activities (assimilation and validation). Improved tests include some simple tests based on comparison with climatology and a model background check based on a global ocean reanalysis. Visual quality control is performed on all suspicious temperature and salinity profiles identified by the tests, and quality flags are modified in the dataset if necessary.

In addition, improved diagnostic tools have been developed – including global ocean indicators – which give information on the quality of the CORA3 dataset and its potential applications.

CORA3 is available on request through the MyOcean Service Desk (http://www.myocean.eu/).

An Information System for Monitoring Coastal Erosion at Colombia – COSTERO

David F. Morales, Marine Geoscience Research Program,
Marine and Coastal Research Institute - INVEMAR, david.morales@invemar.org.co (Colombia)
Jose Eduardo Fuentes, Information Systems Lab – INVEMAR, jose.fuentes@invemar.org.co
Julio Bohorquez, Information Systems Lab – INVEMAR, sinam@invemar.org.co
Daniel M. Rozo, Information Systems Lab – INVEMAR, drozo@invemar.org.co

The Information System for monitoring coastal erosion – "COSTERO", is defined as a technological platform to support and integrate scientific knowledge about coastal dynamic and to be used as a tool for public management. Contributing in the strategic implementation on coastal zones for sustainable development at long term in environmental and socioeconomically aspects is one of its purposes.

Using information and communication technologies, COSTERO promotes scientific knowledge and supports decision making related to prevention, mitigation and control of the spatiotemporal changes of the coastline focused on coastal erosion.

COSTERO has been developed by the Systems Laboratory belonging to INVEMAR. Its development had the support and supervision of technical advisors for the research program on «Analysis and evaluation of erosional processes in the Colombian Caribbean continental and insular coast », sponsored by:

- COLCIENCIAS – Administrative Department Science, Technology, and Innovation
- Magdalena Government
- INVEMAR – Marine and Coastal Research Institute

General aspects related with the development of this project can be consulted at the proposal for the information system – COSTERO (Fuentes et al., 2011).

Development and implementation

This system was designed and implemented through three stages. First, it was necessary the consensus and approbation of national experts, making clarify definitions considered relevant, followed by an overview of the developments and tools that may be relevant in coastal erosion as a national priority. In the second part, there was developed the conceptualization, user's definition, methodology, requirements, logical and software architecture. In the last stage an open access website was developed in order to provide different software tools, necessary to gather, organize, find and transfer geographic and alphanumeric information. These aspects are related to subjects such as climate, geology, geomorphology, physical oceanography and hydrography of Colombia's coastal areas.

Integrated data and information sources were organized to develop products as web services and geographical statistics, required for better understanding coastal erosion. This information is used as an input for different ways to manage the risk associated with coastal erosion. COSTERO also includes a metadata consulting tool, and establish standards and protocols to collect, exchange and spread different information of projects related to coastal dynamics.

Its development strategy will include some modifications depending of the precision on requirements, technology selected and resources available, mainly of experienced people on the research of coastal erosion. Further information available at www.invemar.org.co/erosioncostera or follow us on twitter @costerocolombia.

References

Fuentes, J.E., J. Bohorquez, D.F. Morales, J.F. Gómez, D. Rozo. 2011. *Diseño del Sistema de Información para el seguimiento de la Erosión Costera en Colombia – COSTERO. Informe técnico final.* INVEMAR, COLCIENCIAS, Gobernación del Magdalena. Santa Marta, 37 pp.

See the Sea - Multi-User Information System Ocean Processes Investigations Based on Satellite Remote

Olga Lavrova, Space Research Institute of RAS, olavrova@iki.rssi.ru (Russia)
Evgeny Loupian, Space Research Institute of RAS, evgeny@d902.iki.rssi.ru
Marina Mityagina, Space Research Institute of RAS, mityag@iki.rssi.ru
Ivan Uvarova, Space Research Institute of RAS, uvarov@d902.iki.rssi.ru
Tatiana Bocharova, Space Research Institute of RAS, tabo@iki.rssi.ru

The principal aim of multi-user information system "See the Sea" (STS) is to provide researchers with access and analysis tools for information derived from remote sensing data (online and archived) tailored for investigation of various processes in ocean and atmosphere above it. The main advantage of STS is its facilities for comprehensive description of various processes and phenomena, estimation of their quantitative and qualitative characteristics, and utilizing different analysis tools.

The following subsystems have been built within STS:

- data collection from various sources and data processing to obtain information products necessary for ocean processes/phenomena investigation;
- management of satellite data and information products archives providing versatile search and fast selection of datasets required for analysis and processing;
- data interface subsystem providing:

a) easy search and selection of data within predefined geographical regions and time periods for particular processes/phenomena observed;

b) data analysis facilities, including analysis of spatial and temporal distributions, different parameters and their interrelations required for ocean dynamics modeling purposes.

STS is designed to handle the following main types of data: satellite data, including those of various radar sensors, Russian Earth-observation sensors of high and low resolution, and meteorological sensors; data of ground meteorological stations; cartographic data. STS is an open system capable to incorporate in future any required information (altimeter, buoy, contact data, etc.).

The prototype STS has successfully undergone tests for the following case studies:

a) investigation of vortex structures and frontal zones in selected regions;

b) detection of areas worst affected by oil products and other anthropogenic pollution;

c) accumulation of statistics on spatial, seasonal and interannual variations of dynamics processes in coastal zone affecting pollution drift.

STS teaching classes have been held for students and doctoral students of the Oceanology Chair, Geography Faculty of the Lomonosov Moscow State University and doctoral students in related sciences of Space Research Institute RAS and Geophysical Center RAS.

Proposals are elaborated for building, based on the technologies and prototype STS developed, a multi-user environment providing remote sensing data handling facilities for experts and institutions engaged in world ocean studies.

The development of STS is supported by Russian Foundation for Basic Research project 11-07-12025-ofi-m-2011.

Bridging the gap between data and metadata (Part 1): a novel paradigm and technologies for data discovery, evaluation and access, emerging from addressing the issues of seismic data within the Geo-Seas project

Paolo Diviacco, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, pdiviacco@ogs.trieste.it (Italy)

Alessandro Busato, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, abusato@ogs.trieste.it (Italy)

Helen Glaves, British Geological Survey (BGS), hmg@bgs.ac.uk, (UK)

Dick Schaap, Mariene Informatie Service, dick@maris.nl, (NL)

Introduction

The mission of the EU FP7 Geo-Seas project is to provide a mean to deliver and access integrated sets of primary marine geological and geophysical data, which are among the most important elements in the process of scientific and applied marine research, economic activities, and sustainable environmental management, at regional, European and global scales.

Such data space requires an European-wide services infrastructure, standardised practices by the data repositories, and middleware so that end users can identify, locate and access the data they might be interested in. Being a sibling of the SeaDataNet project, Geo-Seas has adopted the technologies developed within the latter, extending them and introducing new paradigms.

Within this perspective specific attention has been paid to seismic data in comparison with other data types handled by the Geo-Seas initiative, as this data type has particularities that need to be carefully considered.

The problem

One very important point is data value. In fact, seismic data acquired by research institutes, being relevant for mineral prospecting can be an important resource for private companies. Commercial exploitation of data acquired by research institutes can vary a lot depending on the data policy of the specific research institute and therefore has to be handled with care in order to avoid the data owner losing control on the data itself. For example, permitting un-mediated data download does not allow for the protection of the ownership of the dataset. On the other hand data is one of the most important facilitators in fostering a collaborative attitude between research institutions. It is therefore very important to balance protection and dissemination of data.

From a technical point of view most of the issues are related to the size of the data files. Although the extent of a seismic survey can vary a lot from case to case, data files often exceed a gigabyte in size. It is therefore easy to imagine that moving back and forth such amount of data through the web has to be considered carefully. On the end user side, another very important issue is the time and resource requirements coming from the need to upload into processing and interpretation packages huge amounts of data that were not previously selected upon relevancy criteria. This information overload can be unbearable. Geo-Seas therefore had to develop a new paradigm able to support the end user in his/her search for the data that takes into consideration all the issues mentioned above. In the traditional view, data discovery is separated from data usage. Data discovery is based on metadata, which is a description of data based on semantic rich text supposed to be queried in order to filter all available datasets to hit the data an end user is interested in. Traditional discovery of data is considered in a Dublin Core or ISO19115 sense; meaning a core set of metadata elements that allows searching across domains or data types. On the opposite side data is simply the observation itself, without any semantics. In our opinion this duality between metadata and the actual data is detrimental to an effective data search and therefore we devised a means to bridge these two extremes.

A new paradigm

Within the Geo-Seas project, 5 steps were identified that allow an end user to fully assess the usefulness of the data he/she selected. These steps are:

1) Discovery - Core metadata (Dublin Core, ISO19115 sense): relying on minimal information as when, where, who, what (very approximate). Please note that this set of information can correspond to a single profile crossing all domains
2) Browsing - Domain specific metadata: domain specific parameters (e.g. Sampling frequency or trace length in the case of seismic data). It should be noted that considering that the set of metadata changes from domain to domain, it is not possible to have just one cross-domain profile
3) Data preview: once data of interest has been identified through discovery and browsing it is necessary to step into the actual data, but considering the technical and value related problems of seismic data mentioned above this is not straightforward; it is advisable, instead, to introduce several steps. The first of these is data preview where data is represented as a low resolution, and possibly watermarked, static image of the data. This allows large geological features to be resolved but not the finer detail.
4) Remote data access (viewer): if the data preview has revealed that the data could be of interest then the end user will need to have a look at subtle features and this cannot be done using static low resolution images but only with a data viewer enabled to directly access, while at the same time protecting, the data.
5) Download: once data of interest has been identified, the end user should then contact the data owner and negotiate the possible use of the data. Generally this corresponds to an agreement that may include a joint collaborative project on the exploitation of the data.

The implementation

The SeaDataNet infrastructure had already made available tools and services to allow discovery (1) and downloading (5) of data, the Geo-Seas project has had to develop a further three segments and integrate them in the infrastructure for use with seismic data.

Browsing was addressed introducing O&M and SensorML. A collaborative effort between all partners involved in seismic data management allowed the definition of a domain specific metadata profile and related controlled vocabulary (Diviacco et al. 2012) that linked from the CDI, allows detailed selection of data based upon specific parameters such as, for example, source type (and therefore penetration), sampling rate (and therefore vertical resolution) or acquisition geometry (and therefore horizontal resolution).

Data preview was implemented via open source software able to produce thumbnail images and enabling linking from the CDI.

Most of the efforts were spent on the development of a high-resolution seismic viewer. This tool allows end users to directly access the actual data while at the same time protecting the data by preventing direct download of the dataset; it, at the same time, decouples and connects the user and the data. Taking into account the technical and value-related issues mentioned above, the viewer was developed using a mixed server-side/client-side approach (Diviacco 2005, 2007) where the maps are based on SVG graphics rendered by a Batik applet running on the client side. Seismic data is interactively processed by the open source vertical software CWP/SU (Diviacco, 2008,2012,2012b) running server side and producing pyramidal TIFF files that can be zoomed and panned interactively using the server-side applet. Interactive requests from the client applet are managed by a JMS broker.

All of the five steps for data discovery and access are integrated into the Geo-Seas infrastructure, so that users are identified and tracked as part of the "shopping" process. Most of the main Geo-Seas partners managing seismic data are currently using a test configuration for the seismic viewer that has proved to be robust. The viewer will therefore move to a fully operational instance in the near future allowing all partners holding seismic data to be on-line.

References

Diviacco P.; 2005: *An open source, web based, simple solution for seismic data dissemination and collaborative research.* Computer & Geoscience, 31, 599-605.

Diviacco P.; 2007: *Data systems and the social aspects of scientific research. In: Proceeding of the European Geosciences Union (EGU),* Vienna, IGO 2007-A-02542, GI10-1FR4P-0328, poster.

Diviacco P.; 2008: *Col-laboratories, tools to foster the collaborative attitude within a scientific community while protecting data assets.* In: Proc. 33° International Geological Congress, Oslo, August 6-14, 2008.

Diviacco P.; 2012: *Towards a collaborative research data space in Geophysics.* Mediterranean Marine Sciences, Special Issue on Marine Data Systems.

Diviacco P., Lowry R., Schaap D.; 2012b: *"Marine seismic metadata for an integrated European scale data infrastructure: the FP7 Geo-Seas project",* BGTA, Vol. 53, n. 2, pp. 243-252.

Harmonization of Digital Terrain Model production in EU distributed infrastructures

Eric Moussat, Ifremer, eric.moussat@ifremer.fr (France)
Benoît Loubrieu, Ifremer, benoit.loubrieu@ifremer.fr
Thierry Schmitt, SHOM, thierry.schmitt@shom.fr

Bathymetric products results of measurements carried out by various organizations whose responsibilities differ significantly from one to another: oceanographic Institutions, universities, hydrographic offices and governmental agencies in charge of the safety of the navigation and of the Defense, private companies. Collecting soundings by these organizations to make bathymetric products is time consuming and expensive. Many of their data sets are not indexed in public catalogues. Policies of data providers might restrict their access especially in cross border areas or require long, and not always successful, negotiations. Lack of common approaches (metadata and data content made available, geometry, vocabularies and format) to generate bathymetric products makes processing complex and sometimes impossible.

Thanks to the experience acquired during the EU initiatives, SHOM and Ifremer developed a general mechanism to provide bathymetric data from multiple data providers. This was done within the context of distributed European geographic database and as part of EU initiatives as Emodnet hydrography , Geo-Seas and SeaDatanet projects.

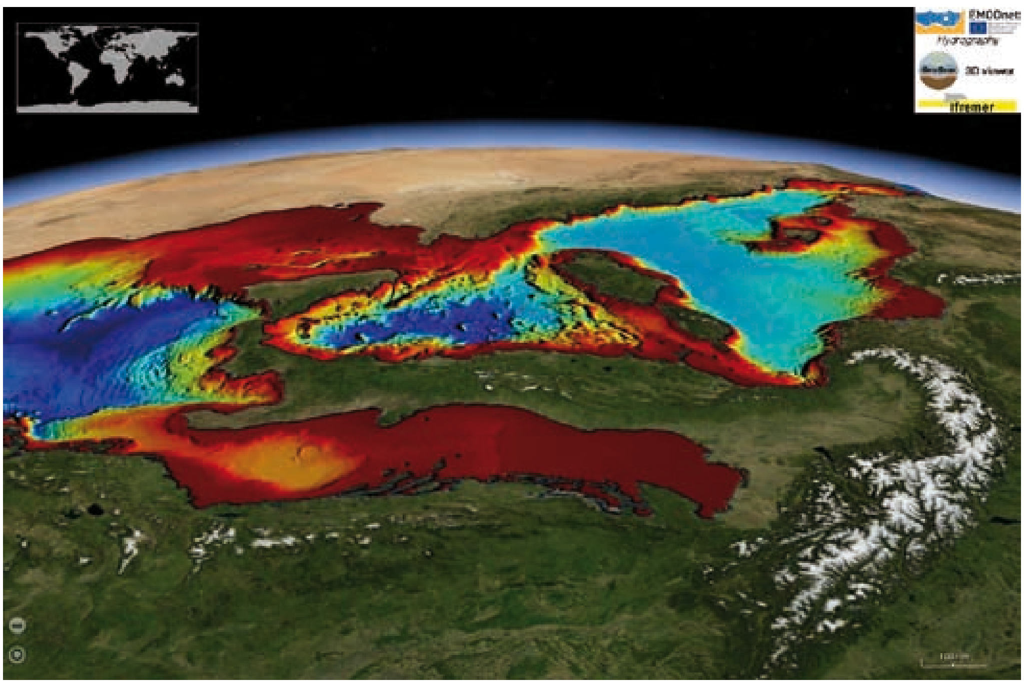


Fig. 1 - View of the Emodnet DTM with the Geo-Seas 3D DTM viewer.

The specifications of the demonstrator were based on a survey of end-users carried out by the Geo-Seas partners. The end-users can benefit of an integrated Digital Terrain Model (DTM) product that merges survey data at different scales from various data providers, while preserving the technical and administrative (diffusion policy) constraints of the data provider. New services can be offered such as 3D viewing using tools such as the Globe 3D DTM viewer, a freeware adapted to the Geo-Seas purposes. Both the DTM products and the corresponding services were designed with the aims of helping the end-users to access bathymetric products, metadata and other qualitative attributes and to allow them to assess the quality of the source data sets and their fitness of use.

The success of the EMODNET Hydrography project using similar principles and procedure shows that the proposed mechanism to provide data has been well accepted by many data providers as it preserves their interest while giving more visibility on their activities.

This mechanism allowed a decentralized cooperation for the production of large coverage synthesis, using decimated grid and leaving source datasets (at the highest resolution) held and managed by the data provider. This decentralization allowed also a closer interaction with local actors. Overall the mechanism contributed to create the EMODNET 15” DTM of the European Seas (Fig. 1) in a remarkable short time.

Marine Geology data accessibility in the european framework: the I.G.M.E. Database and access services after its participation in the GEO-SEAS project

Irene Zananiri, Institute of Geology & Mineral Exploration - EKBAA, izanan@igme.gr (Greece)
Dimitris Mitropoulos, IGME-EKBAA, mitropoulos45@gmail.com
Vaggelis Zimianitis, IGME-EKBAA, zimianitis@igme.gr
Chryssanthi Ioakim, IGME-EKBAA, ioakim@igme.gr
Vassilis Papadopoulos, DSD-EKBAA, rootmail@igme.gr
George Efthimiou, IGME-EKBAA, efthimiou@igme.gr

During more than 30 years of activity and participation in numerous research pro-jects, the Institute of Geology and Mineral Exploration of Greece has carried out extensive work in marine geology and a vast amount of data has been collected. In order for those data to be available and used by the international scientific community, a need for standardization and harmonization was imminent. Towards this scope IGME, along with 25 other marine geological and geophysical data centres, located in 17 European maritime countries, participated in the GEO-SEAS project: the implementation of an e-infrastructure.

The aim of the project was to enable users to identify, locate and access pan-European, harmonised and federated marine geological and geophysical datasets and derived data products held by the data centres through a single common data portal.

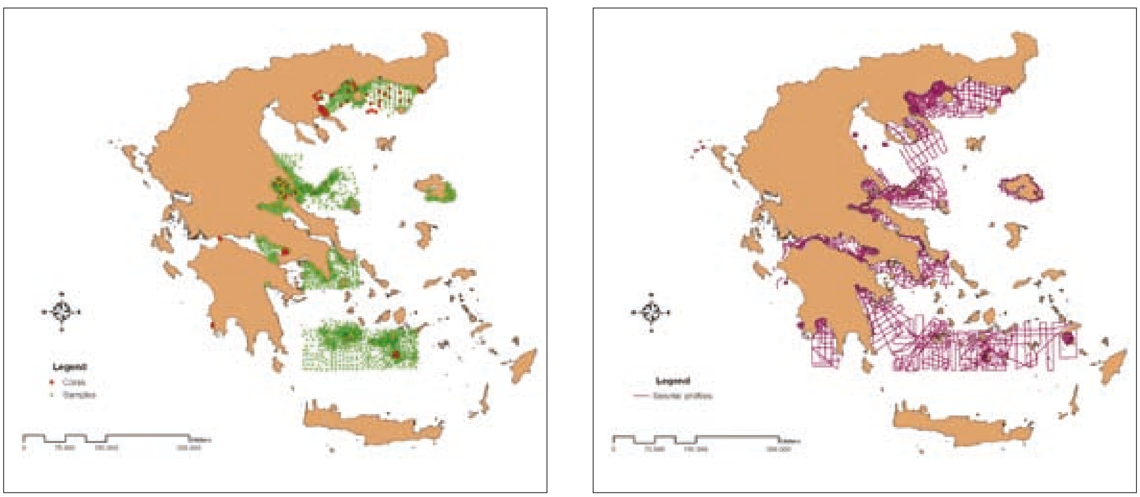


Fig. 1 - The dataset of IGME included in the Geo-Seas portal: sediment grab samples and sediment cores (left), lines of seismic profiles (right). These data have been collected by IGME from 1979-present.

The IGME Marine Geology database

For the scope of the GEO-SEAS project the IGME database was updated and completed with all the information required. Metadata and associated data files from a total of 2147 sediment grab samples, 79 sediment cores – comprising 437 specimens – and 547 seismic profiles (~20,000 km) – reaching a total of 2603 scanned images – were prepared for inclusion in the GEO-SEAS e-infrastructure (Fig. 1).

One of the primary goals of the project was to produce the CDI metadata records for inclusion in the GEO-SEAS infrastructure. The CDI population and maintenance activities at IGME included various sub-activities:

- Analysis of the formats and local availability of metadata
- Editing of existing metadata according to the GEO-SEAS standards and guidelines
- Compiling and validating a first test-batch of metadata
- Compilation and/or updating of the metadata and validation into a test site by MARIS
- Importing geological and geophysical XML metadata entries into the central CDI directory.

In total 2773 CDI records were prepared to describe the IGME geological and geophysical data, including all the mandatory and optional information –where applicable. Moreover, to provide additional information for the available seismic profile data, xml metadata files describing the sampling methodology and seismic section data record were created [2603 SensorML and 547 O&M (Observation and Measurements) metadata files]. Finally, to conform to the low-resolution viewing services pre-requisite, low-resolution thumbnails of the 2603 scanned seismic profile images were created.

Following the GEO-SEAS standardization and harmonization protocol, 2226 files – 2147 sediment grab samples and 79 sediment cores – of lithology and grain-size parameters were prepared.

The files were in Ocean Data View (ODV) format, which is a commonly used INSPIRE-compliant transfer format. Concerning the geophysical data, the transfer format that was applicable to IGME was high-resolution image files with raw data representation. To achieve this all the seismic profiles, still in analogue format, were scanned to the required resolution (~20,000 km of seismic lines, with a total of 547 seismic lines – amounting to a total of 2603 high-resolution images).

Future prospects

Detailed records of palynolology, palaeontological, radiochronological and other data on certain cores are being prepared in accordance with the GEO-SEAS policy and will be included in the e-infrastructure. Moreover, to facilitate the use of all the existing marine geology data, a GIS database is being prepared, with layers of existing information and mapping products (e.g. grain-size distribution, geological features etc.).

Through the implementation of the GEO-SEAS project the existing marine geological and geophysical information of IGME has been standardized, harmonized and is now available for search and download through a single portal. This has facilitated the search and use of the information from the scientific community and other users.

The standardization, harmonization and provision of data and metadata, carried out through the GEO-SEAS project, are aligned with European directives and recent large-scale framework programmes on global and European scales, such as GEOSS and GMES, EMODNET and INSPIRE. Thus, they allow IGME to fully conform to the EU directives for the provision of marine geology data.

Opening up Irish marine data

Trevor Alcorn, Marine Institute, trevor.alcorn@marine.ie (Ireland)
Eoin O'Grady, Marine Institute, eoin.ogrady@marine.ie

The Marine Institute (MI) Open Data Portal aims to provide simple self-service user access to Marine Institute open data assets, following the trends evolving in the open data community via the delivery of simple, clean, and easy to navigate data portals providing data in multiple commonly used formats. The objective of the data portal is not to act as a central data repository for the MI but to simplify access to open marine data that is available from various organisational applications and services. The portal facilitates public access to MI data, promoting reuse, and will reduce the overhead on internal staff in processing manual data request process.

The MI Open Data Portal includes a number of high level functions such as
- Search by Keyword
- Search by Geographic Extent
- Filter by Data Theme (Category)
- Data is presented for direct download in multiple common formats
- Data is presented for user defined download
- Data owners can add and edit dataset details and links (access controlled)
- Administrators have the ability to report on open data usage

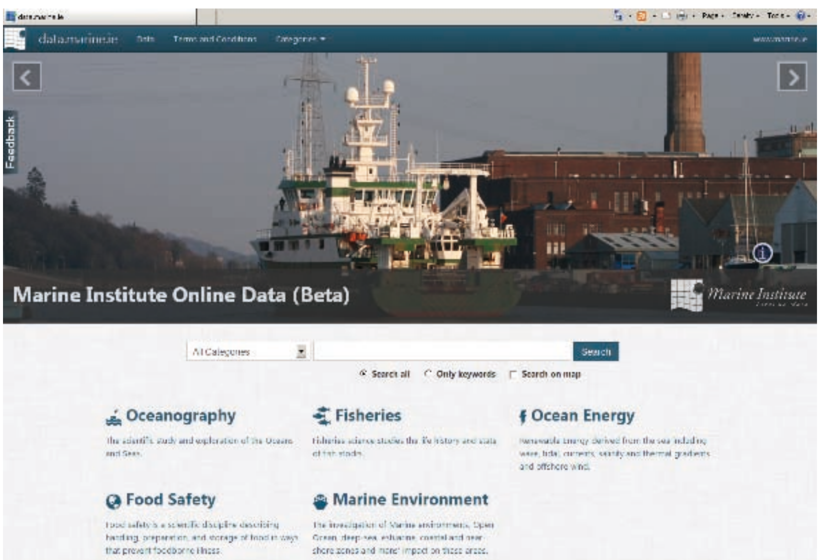


Fig. 1 - MI Open Data Portal.

Towards Open Monitoring Data within marine monitoring projects

Gerrit Hendriksen, Deltares, gerrit.hendriksen@deltares.nl (The Netherlands)
Gerard van der Kolff, Deltares, gerard.vanderkolf@deltares.nl (The Netherlands)
Kees Borst, Rijkswaterstaat, kees.borst@rws.nl (The Netherlands)

Introduction

Port of Rotterdam is carrying out the Project Mainport Development Rotterdam, PMR, expanding the port with 1,000 hectares net of industrial ground, located directly on deep water, through a suppletion of sand in the Voordelta area in the North Sea. As the Voordelta has the status of protected nature reserve (Natura2000 area), European laws dictate that the expansion of the port must be compensated for loss of habitat. As part of the compensating measures, the Dutch national government established a series of protected areas in the Voordelta region. Rijkswaterstaat selected a consortium of Deltares, Imares, CSO, Bureau Waardenburg, Arcadis/Alkyon, INBO and NIOZ to conduct a long term monitoring of the nature compensation in the Voordelta (PMR-NCV).

The Monitoring data gathered are stored in an online geodatabase. The data can be disseminated via OGC standards and be viewed in for instance Google Earth or QuantumGIS.

The data are available for analysis to answer the questions raised within the project. The main question is whether the seabed protection area and resting areas adequately compensate for the loss of habitat and the increased difficulty of obtaining food for nesting terns and wintering common scoters

Storage and dissemination

During start-up of the monitoring project OpenEarth (http://openearth.eu) methodology was chosen to store information of the monitoring and to disseminate information. OpenEarth is a philosophy which describes the use of open source products and generic tools to store and disseminate data. It is not a ready to use product, but the implementation is flexible and dependent on available techniques and demands. The nature of the data collected by the consortium partners pointed towards the use of databases. PostgreSQL in combination with the module PostGIS was the obvious choice to store data. The use of PostGIS enhances the standard storing and querying of data with so-called spatial queries.

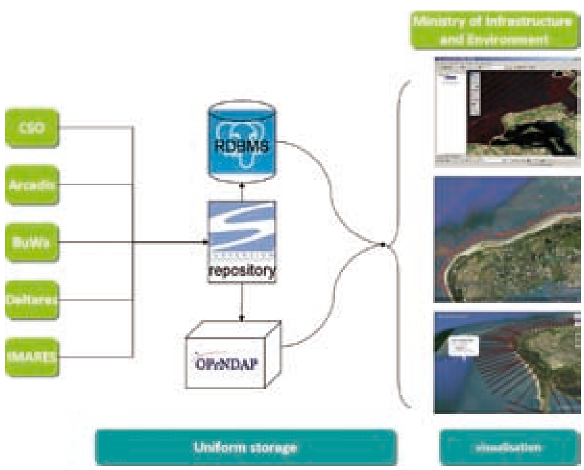


Fig. 1 - Data flow within the PMR-NCV monitoring project.

This ensures spatial aggregations and also enables easy visualisation through commonly used OGC standards like WMS, WFS and KML.

At this moment several millions of data entries have been stored and are available for analysis to answer the questions within the project. The data is structured in a generic data model.

Data can be uploaded through the internet. Various partners created their own solutions, mainly using Python as scripting environment. Quality assurance is arranged within the internal processes defined by the partners of the consortium. They use commonly accepted standards. The complete data stream is visualised in Fig. 1. The repository plays a crucial role in the total stream of data. This ensures version control on the data imported into the database. If necessary, data can also be converted into netCDF files. In practice this is only done in case large arrays of model output have to be available.

Standards

Parameter tables of the database are built using European standards for Marine Species and Landuse, respectively WoRMS (http://www.marinespecies.org) and the HILUCS classification system (http://inspire.jrc.ec.europa.eu/index.cfm/pageid/241/documentid/2627) are used. For geomorphological characteristics the CF-standard is used (http://cf-pcmdi.llnl.gov/).

Dissemination of the data can be arranged using OGC technical standards like WMS, WFS or KML. These standards are described on http://www.opengeospatial.org/standards . At the moment there are two ways to visualise data, that is either by using QuantumGIS (http://qgis.org) or disseminating via geoserver (http://geoserver.org). A stored procedure that converts basis metadata to ISO compliant XML is added to facilitate push of metadata to a Geonetwork node from Deltares enabling further dissemination of data.

Data systems of the Dutch government handle data in the Dutch standard called Aquo directly. To enable this direct handling a server side mapping to the datamanagement system is provided.

Visualisation

The nature of the database and the accessibilty gives the opportunity to visualise the data in various ways by for instance QuantumGIS or KML. The advantage of QuantumGIS is that it is able to contact the PostgreSQL database directly. This combines true GIS functionality with the use of PostgreSQL. Other means of visualisation is arranged by Geoserver. Geoserver plays a crucial role in disseminating data through the internet. Geoserver is able to convert PostGIS data (module on PostgreSQL) to web services like WMS, WFS and KML (check Fig. 2). This enables (at least for the Web services WMS, WFS) use of data in other software like ArcGIS.

Fig. 2 - Dynamic visualisation in Google Earth of occurrence of Common scoter (Melanitta nigra) in the study area.

Querying structured and unstructured data - combined approach

Asen Stefanov, Institute of Oceanology- BAS, a.stefanov@io-bas.bg (Bulgaria)
Atanas Palazov, Institute of Oceanology- BAS, palazov@io-bas.bg

Traditional data base management systems(DBMS) are highly structured, making them ill-suited to store the huge volumes of random unstructured information. Unstructured information does not fit well in the relational model. The problem is that this data is not well defined. In the past there were two options to store unstructured data file in the database: either to store the file in the database as a BLOB column, which was good for management, but very bad for performance, or to keep the file in the file system, and store a link in a database- good for performance, bad for managing the files and keeping them in sync with the database. Now there is a completely new architecture inside the present DBMS for handling file or unstructured data inside the database. The current DBMS offers the best solution for storing file content such as images, audio, video, PDFs, spreadsheets etc.

Bulgarian oceanographic data center (BGODC) has hands-on some experience in the still relatively new domain of unstructured data processing, and has helped the scientists to combine structured and unstructured solutions to deepen their knowledge. Initially the sources and file types of unstructured oceanographic information are identified – cruise reports, project documentation, contracts, maps, and cost sheets and so on. The various files and documents are bulk loaded, updated as well as managed in SQL in DBMS, and access is provided for standard applications as if they were stored in the file system. There are mostly text heavy, but can contain also scanned documents and images, presentations and etc.To access the unstructured part of data base the common network share for local users was implemented and an FTP portal was installed for internet users. To implement a combined approach in queries a set of drill-down reports from the structured parts of the database are developed. These reports are used as search filters over unstructured part of the database to set keywords for full text contents searching. The different queries over unstructured textual data associated with different sources are created (Fig.1).

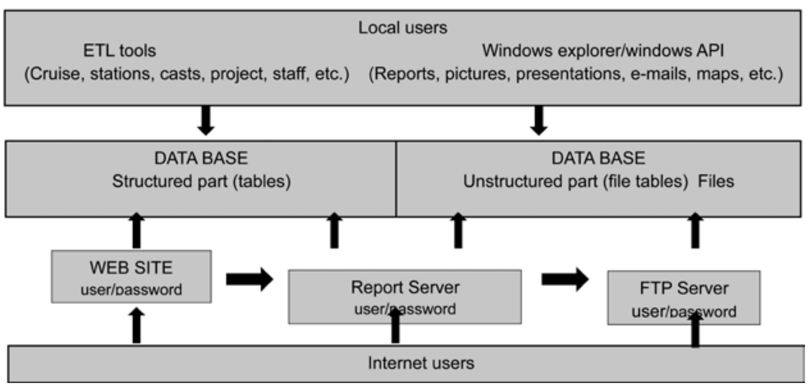


Fig. 1 - Data management and data flow.

The paper discusses the mechanism by which the current DBMS lets the files to be stored in the operating system file system, and keep them transactionally consistent with the database. This new mechanism brings support for the standard file namespace and compatibility with other applications to the file data stored in the DBMS. The paper proposes the ways an application integrates its storage and data management components, and explain full-text search and semantic search over structured, unstructured data and its metadata.

In summary the full text searching is best done in the structured environment; however, it is necessary to read, integrate, and precondition the unstructured data before it can be used for the purposes of analytics in the structured environment. The benefits of combining the structured and unstructured data are: The unstructured data can be analyzed; the unstructured text can be accessed by direct or indirect searches; the unstructured data can be linked to structured data and composite queries can be created; at a certain point the saved in data base unstructured information can be transferred, for instance, into an established oceanographic relationship management system.

BSH contribution to the Copernicus Marine Core Service "MyOcean": Northwest-Shelf in-situ observations

Tobias Gies, BSH Hamburg, tobias.gies@bsh.de (Germany)
Kai-Christian Soetje, BSH Hamburg, kai.soetje@bsh.de (Germany)
Mathias Luecker, BSH Hamburg, mathias.luecker@bsh.de (Germany)

Summary

Efficient in-situ based monitoring of marine conditions over time and space is an important and developing field to provide essential validated information for many applications such as: coastal and marine environment, marine safety, marine resources and weather, climate and seasonal forecasting.

A huge network of different ocean in-situ observatories facilitates the accessibility to an extensive range of ocean data. BSH co-operates with the EuroGOOS regional alliances to collect and integrate marine in-situ data (in real-time and delayed mode), make them interoperable with other national and international programmes and provide access to data and information products on a sustainable basis.

The infrastructure developed jointly by MyOcean and EuroGOOS ROOSs has set up a useful service both for operational oceanography in Europe, but also for the research community and the development of downstream services.

Northwest-Shelf in-situ observations

The Federal Maritime and Hydrographic Agency (BSH) contributes to the MyOcean Thematic Assembly Centre, a distributed system of regional components whose initial focus is to deliver in-situ observations from automatic observatories at sea which are transmitted in real-time to the shore. A real-time data base has been developed which allows to import and provide different data formats (e.g. NetCDF, ODV, ASCII or CDI Metadata format) and a interactive web portal was established at BSH to provide an primary access point for search, access, visualisation and download of real-time and historical data (Fig.1). Every hour the portal is updated with new data transmitted from:

- Fixed buoys and moorings
- Drifting buoys and gliders
- Tide gauges
- Drifting Argo floats
- Research Vessels
- Ferryboxes
- XBTs launched by research vessels and ships of opportunity

One of the main focal points was and is the definition of standardized Real Time Quality Control (RTQC) procedures and the provision of information about the quality of data for physical parameters. These procedures are including automated tests and their specific parameterization for the different European sub-basins and are standardized to reach an agreed level of processing and flagging scale for all products, coping with the aim of the project as a European initiative, to process the data in a distributed but common way.

Bio-geochemical (BGC) data quality control procedures are also under development following similar guidelines but the definition of automated RTQC procedures of BGC data has proven to be specially challenging due to the influence of external factors like sensor calibration, natural biofouling, the lack of historical climatologies and the fact that discrete sampling for validation purposes is taking long time, before becoming available.

Fig. 1 - Interactive web portal.

A contribution to the Marine Observatory network

Sara Almeida, Instituto Hidrográfico, sara.almeida@hidrografico.pt PORTUGAL
Rui Baptista, Instituto Hidrográfico, reino.baptista@hidrografico.pt
Ilmer Golde, Instituto Hidrográfico, ilmer.golde@hidrografico.pt
Ana Lopes, Instituto Hidrográfico, ana.lopes@hidrografico.pt
Ana Saramago, Instituto Hidrográfico, ana.saramago@hidrografico.pt

Framework

Over the years the Instituto Hidrografico (IHPT) has incremented the national observation network with the intent of better serving the community of sea users.

IHPT is the responsible for the publication of tide prediction tables which led to the first observation network – the installation of tide gauges covering the Portuguese country's coast. During the eighties the net was increased with wave buoys and meteorological stations. Over the last five years two new types of equipment were added: multiparametric platforms (buoys) and coastal HF-radars.

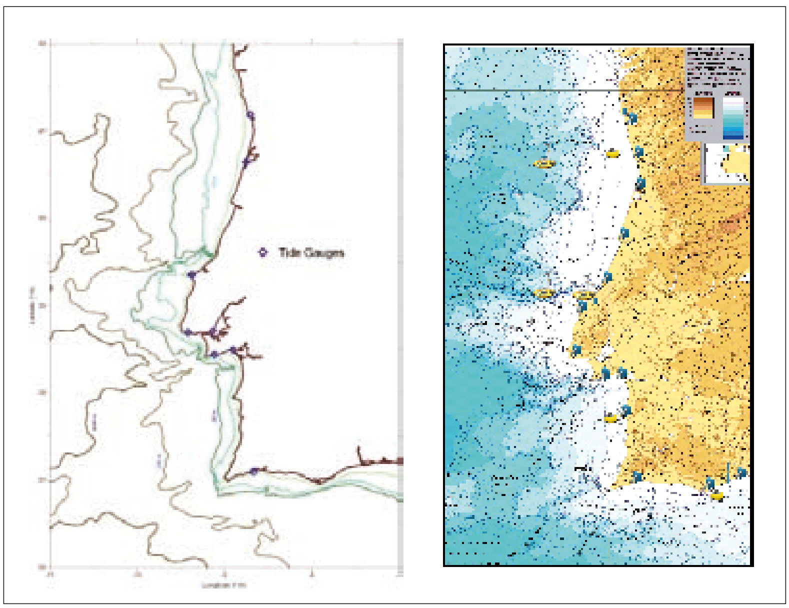


Fig. 1 - The network in the seventies and now.

Monitoring the system at sea

The platforms at sea and the capability to monitor their position in real time forced the installation of an alert system that allows the verification of their positioning. Limits were created for the normal spread of the movement for each mooring: for coastal moorings the acceptable circles are small, but on the oceanic ones the circle has a ray of up to 2 Km.

When the buoys stray away from these circular areas an e-mail is generated and sent to several persons. The monitoring of the movement may be followed using the implementation of the signal on a GIS available on-line. This has improved the efficiency of recovery and repositioning operations.

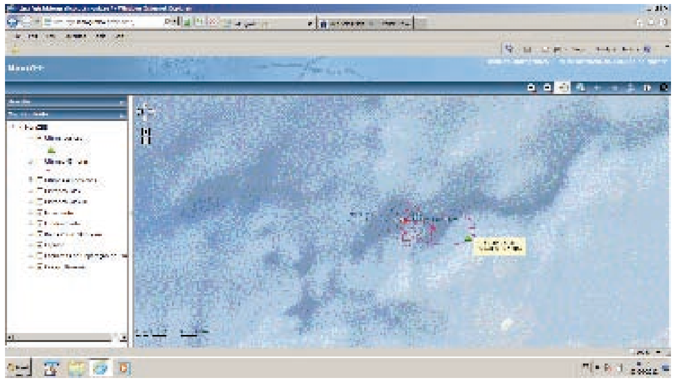


Fig. 2 - A drift situation in 2012.

Outputs

The data management systems, allied to a wide variety of competences, gave origin to a set of products that are available to a large community of interests of those that use the sea as their way of life or pleasure.

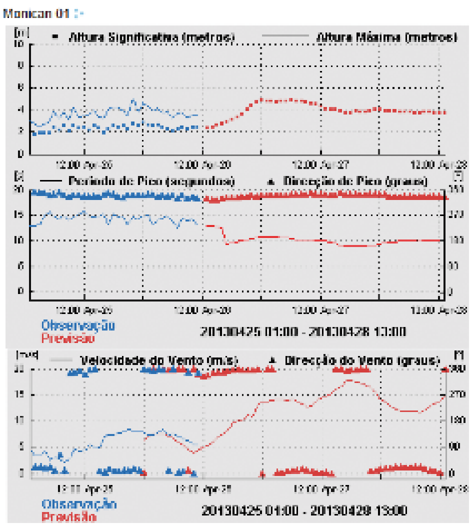


Fig. 3 - Example of a combined product.

Everyday, with the outputs of the forecast models, combined products are generated as the one displayed in Fig. 3. Here, parameters like significant wave height, peak period and direction, wind speed and direction are shown, in blue regarding the observations, and in red for the forecasts.

These products, combining observations from the multiparametric buoys and forecasts, are available at the Instituto Hidrografico website and send by e-mail to the municipality that serves communities of fishermen.

Other products exist, some important to the sports activity like "Surfing Quality Index" or the surface current drift useful to harbour management and that are available through the institute's portal.

Consolidation of marine datasets at Marine Biology Station, NODC for Slovenia

Damir Dezeljin, DEZO s.p., damir.dezeljin@dezo.org (Slovenia, EU)
Branko Cermelj, National institute of biology – Marine biology station Piran, branko.cermelj@mbss.org

As a response to an increased interest of the local community about the the environmental status and environmental issues of the slovenian coastal area, at National Institute of Biology (NIB), Marine Biology Station Piran (MBS) a decision was taken, to invest significant effort in educating the general public about the environmental status of the North Adriatic Sea. This fact and the participation in the SEADATANET project, strengthened the ongoing activities of centrally organize biological, ecological and oceanographic data. The required actions were put on the list of MBS's activities for the 2012/2013 period and they are listed as follows:

- Centralize and organize the existing environmental datasets;
- Execution of quality control procedures on centrally collected data;
- General public education.

Consolidation and organization of the existing environmental data

From the 1980s on MBS is regularly monitoring the southeastern part of the Gulf of Trieste (the northernmost part of the Adriatic Sea). This includes regular measurements of sea water temperature, conductivity, depth and other parameters (hereafter referred to as CTD) as well as chemical and biological analyses of sea water samples collected (hereafter referred to as EWN) during sampling cruises .

In 2009 a central relational database server was set up with the aim to centralize existing data as well as to organize metadata, such as information on expeditions - cruises for data and samples collection (hereafter referred to as Cruise Summary Report or CSR). In 2010 the existing CSR and CTD data were imported into the database. By the end of 2011 the processes and procedures for manually import of CSR and CTD data were developed too. In 2012 we started with the development of processes and procedures for semi-automatic import of CSR, CTD and EWN data. A task that will be finished in 2013.

Applying quality control procedures on centralized CSR, CTD and EWN data

At the beginning all collected data were initially quality checked before they were imported into the central database system. As central database system is beyond question becoming the main source of scientific research and other aplicative activities at MBS, it became indispensable to pass the data through a consistent set of well defined and documented (semi-) automatic data quality control mechanisms. At the moment the development, documenting and execution of mentioned quality control procedures is in its crutial phase and will be finished by end of 2013.

General public education

Thanks to the growing awareness of the importance about environmental issues, the ecology has become an important value for the local communiy. In line with this the sea environmental status and environmental issues became very important as well. For this reason MBS increased the activities in this field. Procedures for an automatic sea environment bulletin creation and publishing were developed and for the public of all ages an underwater camera live streaming (http://www.mbss.org) from one of the protected areas was set up.

The Aqualog profiler is measuring platform for an ocean observatory

Vladimir Solovyev, P. P. Shirshov Institute of Oceanology, Russian Academy of Sciences sol@ocean.ru

Alexander Ostrovskii, P. P. Shirshov Institute of Oceanology, Russian Academy of Sciences osasha@ocean.ru

Andrey Zatsepin, P. P. Shirshov Institute of Oceanology, Russian Academy of Sciences zatsepin@ocean.ru

This presentation concerns with research and development of an ocean observatory. The observatory consists of several moored profiler Aqualog, communication system and data warehouse OceanDB. This technique allows a user to obtain the regular time series of oceanographic data at fixed geographical locations by using conventional oceanographic probes, which are transported by special carrier that moves vertically between surface and bottom of the sea, see Fig. 1.

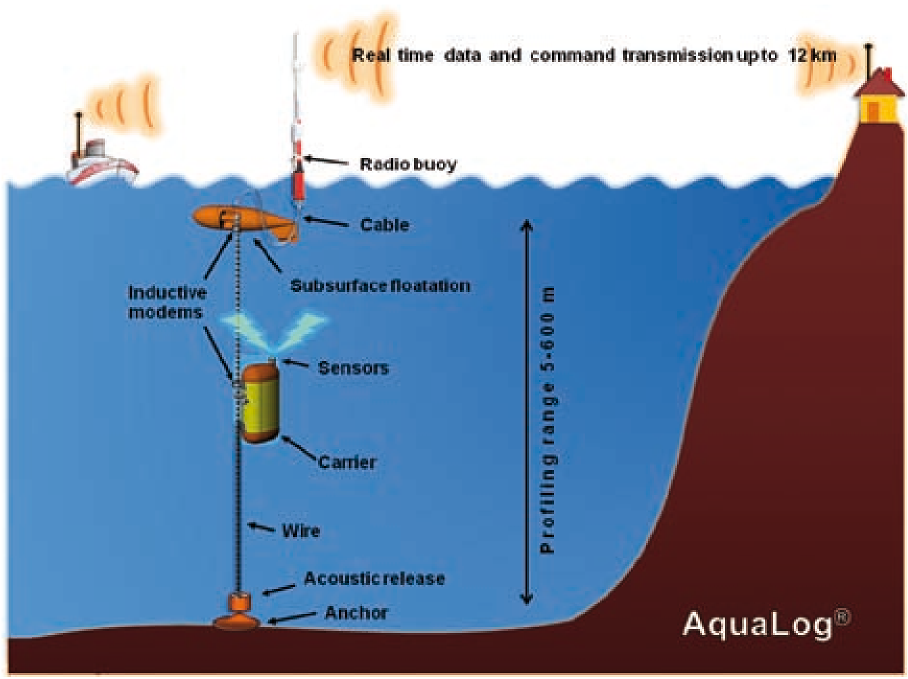


Fig. 1 - The Aqualog profiler.

A moored profiler is a useful tool for multidisciplinary investigations such as water exchange and deep current structure. In the framework of conducting oceanographic research the profiler is a useful tool for field investigations of the variability of both biotic and abiotic parameters of the sea environment from a few hours to a few months. In comparison with the free drifting profiling floats, the moored profilers take a smaller risk of loss and facilitate technical services including replacement of the power batteries and regular cleaning of sensors. The main advantage of a moored station is the data acquisition on a real time basis.

A new moored profiler named Aqualog was designed in P.P. Shirshov Institute of Oceanology in 2005-2012. It was tested at the field trials in the Black Sea, Baltic Sea, Caspian Sea, Japan Sea, Mediterranean Sea, Red Sea and Dead Sea. The profiler is built to carry a load of modern oceanographic instruments. It moves down and up along the mooring line as a sea 'lift' carrying various instruments including SBE 52MP CTD probe, Nortek Aquadopp current meter, and AANDERAA Oxygen Optode 4330F. The pay load of this sea elevator may also comprise other environmental probes e.g., fluorimeter and turbidimeter. The profiler mooring line is made of stainless steel wire that allows a user to extend the maintenance period of the mooring system up to several months. The programmable hardware of the profiler allows a user to set up the automatic operation algorithm (variable movement speed, time and period of the profiling, etc.).

So far the typical depth range of Aqualog's operation was 5–1000 m. With the titanium instrument housing the maximum operation depth reaches out to 3000 m. Vertical speed can be set up within 0.05 and 0.3 m/s. The Aqualog has an energy resource sufficient for profiling the water column in the programming regime for up to several months.

The communication system provides data transfer from the moored Aqualog profilers to data warehouse OceanDB in real time. It consists consists of two parts:

- inductive communication in sea water between the Aqualog carrier and the subsurface profiling buoy;
- mobile GSM / GPRS or satellite communication between the surface buoy and shore point;

The central element of observatory is the oceanographic data warehouse which includes both the Ocean DB itself and instrumental means to work with it. To describe the results of field research a data format in XML language was developed and proposed. The proposed standard for describing the oceanographic data was called OceanML.

The OceanDB was developed on three-tier architecture. It has the following three tiers: presentation tier, application tier, data tier. The data tier of the data warehouse was developed on Oracle Database Server 10G. The choice has been proved by significant progress of the given platform in a direction of integration XML with technologies of databases. The application and presentation tiers were developed on platform Oracle Application Server 10G. The business logic has been developed on Enterprise Java Beans technology.

OceanML describes the structure of the XML document of an expedition's report, possible descriptors and attributes, as well as data types. The description is given in the XML Schema language.

Autonomous profiling stations of the moored type located at a fixed geographical point represent the most preferable technical means for regional marine environmental monitoring. A moored profiler station inhibits a smaller risk of loss and facilitates technical services including power batteries and maintenance of the sensors

Open Source Software as an Efficient Distributed Infrastructure Stack for GIS Data Processing and Dissemination

Przemysław Kulesza, Maritime Institute in Gdansk, przemyslaw.kulesza@im.gda.pl, (Poland)
Michał Wójcik, Maritime Institute in Gdansk, michal.wojcik@im.gda.pl, (Poland)

Introduction

Maritime Institute in Gdańsk has been gathering data on South Baltic environment for many years. Those data comes from different sources (e.g. bathymetric, sonar, geological, oceanographic, spatial planning and magnetic field data) in different formats (e.g. vector and raster data including profiles and probes). The main goal of the BBB (BalticBottomBase) project (http://balticbottombase.eu/info) is to make these valuable data available to other institutions using contemporary standards for geographic data sharing. In order to make those data available in convenient way, they must be described in the uniform manner. Moreover an infrastructure for storing data should be compatible with SeaDataNet project (distributed Marine Data Management Infrastructure) and INSPIRE (Infrastructure for Spatial Information in the European Community) directive. This work presents a distributed platform for GIS (Geographic Information System) data processing and storing.

System architecture overview

Fig. 1 shows an overall architecture of the platform used in the BBB project. The whole was designed in a such way that the data, depending on the format, can be stored in multiple spatial database instances or file systems located on different machines. As a database PostgreSQL enchanced with PostGIS module is used. All the data are exposed through the dynamic web application which can be used by a number of users through a web browser. The web application can be connected to one ore more (to facilitate parallelization) business logic modules which have direct access to the storages. Both, web application and business logic module (EJB module) are deployed on Java EE GlassFish application server. Because GIS data processing can be very time and resources consuming, it is done asynchronously by specialized workers. Those are independent Java SE programs which

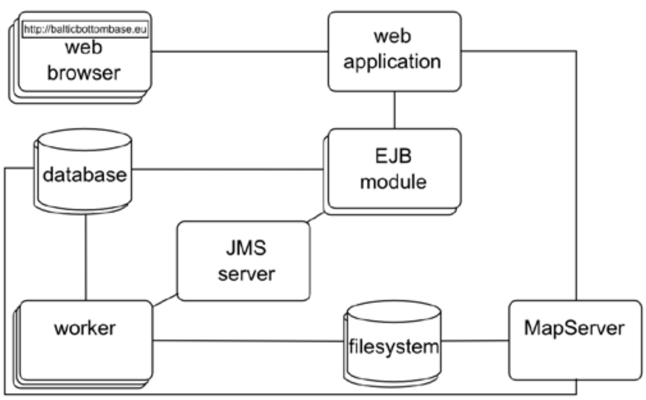


Fig. 1 - System architecture.

can run on different machines. Workers can access business logic synchronously by remote EJB interfaces, but tasks submissions to workers are done through asynchronous messages passed through central JMS (Java Messaging Service) server. Although all the data is stored in databases and file systems, to the user it is exposed through the MapServer using HTTP protocol. Modular design of the platform will allow to expose remote interfaces compatible with SeaDataNet project and INSPIRE directive.

Workers functionality

Implemented workers are responsible for acquiring data details from the provided files and processing it in order to put in the storages. For each file simplified contour in WKT (Well-Known Text) format must be generated. This allows to search for the data in the specified geographical region using uniform search engine and present previews of regions where those data have been collected.

In order to put a new object in storage, user uses the web interface. He must provide the type of the data and files in one of the supported formats. Data insertion is done within two requests. Firstly user requests for detailed data description. This process is usually small time consuming so the results are delivered to the user after short time. Then, if user accepts acquired data, the second request for processing is done. This step can be time consuming, depending on the data size. The requests are sent to the business logic where the worker with the lowest load is selected and the user is informed that the processing has started. Asynchronous message with data type, files localization and request type is put in the shared topic in the JMS server, from where selected worker acquire it.

Worker, in designed time intervals, checks if there are new requests in the topic. Each of the workers has specified number of threads (depending on the resources of the machine where they are running) in which it can process the data. If there are more requests than available threads, they are put in the worker's private queue.

As described above, first step of the processing is acquiring detailed data about the particular GIS object (e.g. bands, SRS - spatial reference system, no data value, size, info about need for optimization). All those data are acquired with an usage of GDAL (Geospatial Data Abstraction Library) which is call through bindings in Java language.

The second step is more complicated and consists of the following steps:

1. if request consists of a number of raster files, those are merged into one with a *gdal_merge.py* tool,
2. the new raster with lower resolution (5 % of input raster size) is created with a gdal_translate tool,
3. from the low resolution raster the simplified outline is created with a gdal_trace_outline tool,
4. the outline in WKT format is stored in the spatial database through the JDBC,
5. source raster is processed by *gdal_translate* tool which is responsible for converting it to

GeoTIFF format with defined tiles:

1. if raster is small (threshold is set in worker's configuration) then it is stored without modifications,
2. if raster is big, the optimized version is created by adding overviews with *gdalddo* tool;

6. object registration in the MapServer is done by adding appropriate layers in two mapfiles responsible for exposing contours and raw data, to style those layers, corresponding SLD (Styled Layer Descriptor) files are created by GeoTools library,
7. temporary files clean up,
8. object registration in business layer through remote EJB interfaces.

Tools for which there are no Java language bindings, are called by the workers through shell scripts.

Conclusions

It can be seen that the automated and distributed platform for processing and storing GIS objects can be build in cost efficient way with an usage of free software without proprietary solutions. At this moment processed files can be provided in the most popular raster formats which can be handled by GDAL, but support for additional formats like ESRI shapefile is currently being developed.

Acknowledgments. Project co-financed by the European Regional Development Fund under the Operational Programme Innovative Economy.

MaNIDA: An integrated national approach to scientific information management and innovative use of information assets' relationships

Roland Koppe, Alfred Wegener Institute, Roland.Koppe@awi.de (Germany)
Ana Macario, Alfred Wegener Institute, Ana.Macario@awi.de (Germany)
MaNIDA Scientific Team

The Marine Network for Integrated Data Access (MaNIDA) aims to build a sustainable e-Infrastructure to support discovery and re-use of data archived in a distributed network of data providers in Germany. One of MaNIDA's long-term goals is to be able to offer an integrative "one-stop-shop" framework for management and access of ship-related information based on international standards and interoperability. This access framework will be freely available and is intended for scientists, funding agencies and the public.

One of the primary themes of MaNIDA is the underway data acquired on board of German research vessels. We will focus this presentation on recently developed solutions for R.V. Polarstern. For this sake, we are going to show the web interface EXPEDITON (http://expedition.awi.de) which offers an innovative approach for discovering and accessing various types of scientific content along with its complex relationships.

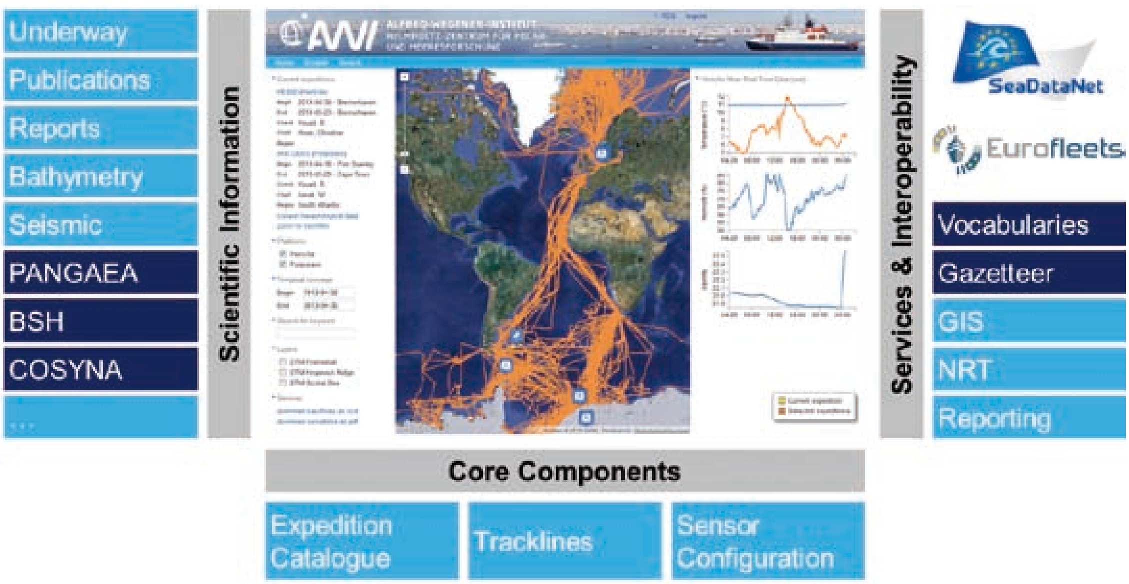


Fig. 1 - EXPEDITION architectural elements.

Fig. 1 depicts the high-level architecture of EXPEDITION with its integrated scientific information, core components as well as interoperability and core services. Scientific information is harvested via the OAI-PMH and OGC CSW standards and indexed with Apache Lucene. This allows access to metadata with good performance and enables several search functionalities. Map-supported browse, facetted search and relationship graphs are some of the features offered by EXPEDITION. In addition to discovery of archived datasets, publications and their relationships, EXPEDITION provides visualization of activities and near real-time (NRT) data for some of the vessel-mounted sensors.

MaNIDA is committed to international and interdisciplinary interoperability, indicated here by the intended use of SeaDataNet gazetteer and vocabularies.

We are currently extending our expedition catalogue to fulfil concrete requirements from the scientific coordinator of AWI's research vessels Polarstern and Heincke. These are, among others, the need to track scientific output associated with various instruments on board and map-support to plan tracks of future cruises in order to optimize resources (e.g. new tracks should be placed in areas where the vessels have not been yet or follow old tracks so as to be able to search for trends). In addition, EXPEDITION will offer soon semi-automatic generation of SeaDataNet Cruise Summary Reports (CSR) to be delivered to BSH. Together with AWI's Directorate and ships' scientific coordinator, we are also planning more intensive co-operation with EUROFLEETS 2 in polar regions.

We demonstrate recent efforts towards visualization of various data products and discuss a typical underway data management lifecycle from cruise approval to data publishing.

A data storage for generic and heterogenous scientific data

Ketil Malde, Institute of Marine Research, ketil.malde@imr.no (Norway)
Tomasz Furmanek, Institute of Marine Research,
Esmael M. Hassen, Institute of Marine Research

Systems for storing and managing scientific data are often targeted at specific data types, and often have specific applications in mind. This focus results in integrated systems that are tailored to important problems and with implementations that cater to the experts in the field.

The downside is that monolithic end-to-end systems often are expensive and difficult to extend or modify. In quickly developing fields (like molecular biology) there are large volumes of data and a plethora of formats and technologies, and future uses can be hard to predict. As these data are scientifically important, there is a need for lightweight systems that can incorporate these data with a minimum of effort, yet maintain enough structure to make the data reusable and interpretable in the future.

To address this, we have designed and implemented a lightweight, generic data store, and are currently using it to manage a variety of molecular biology data.

Goals

The system is designed with the following goals in mind:

The system should be *generic*, or data agnostic. It should not be limited to specific data types, and adding new data types and formats should be as effortless as possible.

The system should be simple to understand. It should be easy for data managers to handle data without explicit understanding of the data, and it should be equally easy for domain experts to use data without much knowledge about data management structures and systems. This implies a *separation* of data management skills from domain knowledge.

Also, scientists must be able to easily obtain copies of relevant data, in formats they are used to and with sufficient metadata to ensure *integrity* and *provenance*. Similarly, it must be a minimum of effort for the scientist and data manager to prepare new data sets for inclusion.

The system should be long term sustainable. This means it should be as *technology independent* as possible, and not tied to any particular language or software.

It should be modular and *forwards compatible*, so that existing functionality will not be affected by new developments in data or metadata formats.

Structure and implementation

Each data set is a stored in a unique directory, and the data is stored in files using native file formats. This makes it simple to extract or transfer datasets (with a file copying operation, or by exposing the directory to an HTTP server and using e.g. curl or wget). Receiving data in standard formats also relieves the scientist from learning about any structure imposed by the storage system.

Metadata is stored in a separate file (simply named meta.xml) using a simple XML schema. The metadata consist of a list of files contained in the data set, including checksums and file types. In addition, there is a free text description, which allows markup of specific terms (e.g. names of species, dates, locations, or references to other data sets). Using XML markup lets the system enforce standard lexical formatting and reference notations (e.g. using TSN numbers to uniquely identify species). An important part of the metadata is provenance of the data, describing how the origin of the data set and its relationship to other data sets.

Many central functions (including: search, visualization, automated analysis, data conversion, data extraction, user management) are intentionally omitted from the core specification of the system. By implementing these as separate, external modules, these can be tailored to specific needs or domains, and ignore data and metadata that are irrelevant for their purpose.

Use cases and sample application modules

Verification and validation of data sets include checking metadata contents and data formats. An XML validator is used to ensure that the metadata file is well-formed and that the lexical format of formal values is correct. Additional checks, like md5 checksums, are performed by ad-hoc scripts.

Downloading a copy is performed with standard software operating on files, for remote access using HTTP (wget or curl), SSH (scp, rsync).

Searching metadata is performed by scanning metadata files and building a database (currently using xapian). A CGI front end provides a web interface.

Data submission is currently manual, and requires some intervention by a data manager in order to construct an appropriate metadata file. This can also be achieved through a web front end that allows the user to upload files and assists in constructing the metadata file.

Automated analysis is performed by tools (typically simple scripts that wrap standard analysis tools) that consume existing data sets in order to generate new data sets. The metadata file is produced automatically, and incorporates provenance information, including the use of data sets, tools and version.

Data search and data extraction by criteria not encoded in the metadata must typically be tailored to specific data types. Search front ends scan the metadata files, and incorporates all appropriate data in an application-specific search database.

Conclusions

We have implemented a lightweight framework for generic data storage. This allows scientists to submit and make use of data sets with a minimum of effort, while making it easy for data managers to incorporate and manage data, as well as facilitating independent and modular implementation of applications that make use of the data store.

National Sea Data System: The Argentine strategy

Gabriela SILVONI, Biblioteca y Servicio de Documentación INIDEP,
Mar del Plata, gsilvoni@inidep.edu.ar (Argentina)

Paola A. AZRILEVICH, Ministerio de Ciencia, Tecnología e Innovación Productiva,
Ciudad Autónoma de Buenos Aires, República Argentina

The National Sea Data System (SNDM) from Argentina, is an initiative of the Ministry of Science, Technology and Innovation framed within a Program for different Database Systems. It was created by Ministerial Resolution in 2010, this is a government policy also supported by a bill sent to the Parliament still to be approved. The mission of this project, is ensuring the accessibility of marine data and information from the southwestern Atlantic Ocean and Antarctica, addressing actions to coordinate initiatives and adoption of plans, programs and projects in this context. Different types of data are included such as environmental, biological, fisheries, oceanographic ones and those collected from Research Vessels. Among the main objectives are: -facilitate the exchange of marine data and information to improve scientific knowledge and decision-making, -provide basic information about the marine environment, under agreed standards, -providing international projection of the Sea Data produced in the country through its dissemination in virtual networks.

MAPAMED, the database on Marine Protected Areas in the Mediterranean Sea

Bruno Meola, Database manager, MedPAN, bruno.meola@medpan.org (France)
Chloë Webster, Scientific officer, MedPAN, chloe.webster@medpan.org (France)

In 2010, on MedPAN's initiative and with the support of RAC/SPA[1], a partnership approach was launched in order to develop a common database on Marine Protected Areas (MPAs) in the Mediterranean Sea.

Objectives of the project

This database – called MAPAMED (Marine Protected Areas in the Mediterranean) – addresses the need for an up-to-date and evolving tool to aggregate, structure, search and make available data on Mediterranean MPAs. More specifically, it aims to:

- Support the analysis and assessment of the progress of the Mediterranean MPA system towards international objectives,
- Facilitate access to data related to Mediterranean MPAs,
- Identifiy ecological and management issues at a supra-MPA scale.

MAPAMED

http://www.mapamed.org

Project steering

With regards to the strategic aspects, this project was followed and advised by a Steering Committee composed of several regional, national and local MedPAN partners. Besides, the MedPAN Scientific Committee and an associated group of experts provided recommendations on scientific and technical aspects of the project. The MAPAMED database was developed in 2011-2012 and is now available online on: http://www.mapamed.org/. It is maintained jointly by MedPAN organisation and RAC/SPA and it will be regularly updated and enriched.

Available data

The MAPAMED database contains both spatial (i.e. boundaries) and attribute (i.e. information) data about MPAs:

- **Spatial data:** points and/or polygons showing the location and/or the boundaries of MPAs.
- **Core attributes:** MPA basic information (identity sheets). These attributes are based on the "core attributes" defined by UNEP-WCMC[2] in the data standard for the World Database on Protected Areas[3]. This data standard has been created to facilitate the sharing of information on Protected Areas. Some of

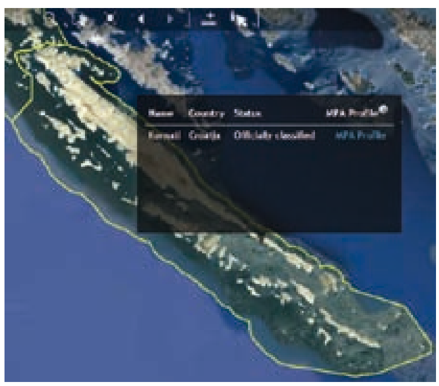


Fig. 1 - MAPAMED viewer.

the core attributes have been slightly adapted to better meet the needs of MedPAN network and its partners.

- **Specific attributes:** these attributes supplement the core attributes with more detailed information on MPAs (Governance, objectives, management plan, staff, equipment, budget, uses, pressures, regulations, monitoring activities…).

Challenges

The database currently contains 170 MPAs representing 208 records[4] (excluding the 507 Natura 2000 sites at sea). Spatial data and core attributes are available for almost all these records. However, specific attributes are more difficult to collect and a lot of information remains to be gathered and completed. Thus, there is a clear need to continue the data collection effort in order to provide a comprehensive view of the Mediterranean MPA system and also be in a position to assess management effectiveness.

Interoperability is also a major issue to address. Indeed, MAPAMED database is expected to be linked to other databases related to Protected Areas in order to facilitate data sharing and to ensure data consistency between databases. In this perspective, its core attributes are based on the standard developed by UNEP-WCMC and metadata should be designed so to be in tune with metadata standards.

Further information about this project is available in the FAQ section of the MAPAMED website.

References

[1] Regional Activity Center for Specially Protected Areas

[2] United Nations Environment Programme – World Conservation Monitoring Centre

[3] UNEP-WCMC, 2010. Data standards for the World Database on Protected Areas. UNEP-WCMC, 9p.

[4] One record represents one designation established on one particular area. As a result, there can be several records for one MPA, according to the number of related designations (e.g. one MPA can be designated both as a national park and as a SPAMI. In this case, two records are created).

Integration of oceanographic databases into the Geographic Information System (GIS) as support for Marine Spatial Planning. Case of Romanian Littoral

Alina Spinu, National Institute for Marine Research and Development "Grigore Antipa" Constanta, alina_daiana_ct@yahoo.com; aspinu@alpha.rmri.ro

PhD. Luminita Buga, National Institute for Marine Research and Development "Grigore Antipa" Constanta, lbuga@web.de

PhD. Razvan Mateescu, National Institute for Marine Research and Development "Grigore Antipa" Constanta, razvan_doru@yahoo.com

PhD. Laura Alexandrov, National Institute for Marine Research and Development "Grigore Antipa" Constanta, laurenta05@yahoo.com

PhD student Mihailov Emanuela, National Institute for Marine Research and Development "Grigore Antipa" Constanta, mihailovme@gmail.com

The Black Sea is a border sea of Europe considered an enclosed sea in the planetary ocean with a very small link with the Mediterranean basin. The Romanian coast is an area with a high instability, substantial amount of natural processes, Danube huge load and impact and physical and economic activities. It is a zone of great significance concerning agriculture impact, business, tourism and recreational activities take place here, old and new settlements have been established in the coastal areas.

The marine and coastal data management for Romanian Black Sea coast, developed towards implementation of Coastal GIS standards, is supported by complementary systems in order to provide a modern scientific and technical instrument to support Marine Spatial Planning and the development of the policies, coastal and marine strategies and management.

A Geographic Information System (GIS) is a computer-based information system used to digitally represent and analyze geographic features. It is used to input, store, manipulate, analyze and output spatially referenced data (Burrough and McDonnell, 1998). A GIS can be distinguished from database management systems or from visualization packages through its specialized capability for spatial analysis.

The integration of data from various sources within a consistent framework, as provided by a GIS, improves their accessibility and their availability. GIS also provides efficient functions for rapid production of thematic or synthetic maps and other visualizations. The main aim of GIS use is to support the sustainable development, the rational use of resources having an integrated view of the area and its activities-objectives and to avoid the different possible conflicts risen between the main coastal or marine activities.

EMODNET-Geology: assembling and disseminating marine data for European seabed mapping

Helen Glaves, British Geological Survey, hmg@bgs.ac.uk (UK)
Alan Stevenson, British Geological Survey, agst@bgs.ac.uk (UK)

The EMODNET- Geology project is one of six projects that have been initiated by the European Commission in response to the EU Green Paper on Future Maritime Policy. The overall objective of the European Marine Observation and Data Network (EMODNET) is to assemble fragmented and inaccessible marine data into interoperable, contiguous and publicly available datasets covering whole ocean basins.

The specific objective of the EMODNET-Geology lot is to provide public access to interoperable geological datasets and make further progress towards a harmonised terrestrial-marine geological data resource for Europe. To achieve these objectives, EMODNET-Geology is compiling marine data and information from a variety of European organisations to create 1:250, 000 scale map products covering a range of themes including seabed sediments, seafloor geology, geological events (submarine slides, earthquakes etc.) and minerals. EMODNET-Geology project will expand the geographical scope of the initial preparatory phase of the project to include wider spatial coverage.

A customised instance of the OneGeology-Europe portal will be used for the delivery of the harmonised map products, and will also act as an access point for the raw data that underpins each of the derived products developed by the EMODNET-Geology project. Both the map layers and the associated data will be delivered using a federated model with each data centre providing the necessary local web services to facilitate the delivery of the data and data products via the OneGeology-Europe portal. The metadata for the data and the associated map products will make use of a standard ISO19115 profile and will be fully INSPIRE compliant to conform to current European legislation.

The long-term sustainability of the EMODNET-Geology infrastructure within the overall EMODNET framework will also be addressed. This will be achieved by ensuring that the EMODNET-Geology infrastructure is integrated with those of existing national mapping programmes wherever possible and also with European initiatives such as the EU-funded European Geological Data Infrastructure (EGDI) project, a scoping study for establishing a pan-European geological data infrastructure.

The Wadden Sea Long Term Ecosystem Research (WaLTER) project: Using the SeaDataNet infrastructure to provide access to integrated environmental and socio-economic data from the Dutch Wadden Sea

Taco de Bruin, NIOZ Royal Netherlands Institute for Sea Research, Taco.de.Bruin@nioz.nl (The Netherlands)

Ronald de Koster, NIOZ Royal Netherlands Institute for Sea Research, Ronald.de.Koster@nioz.nl

Peter Thijsse, Marine Information Services (MARIS), Peter@maris.nl

The Wadden Sea, an UNESCO World Heritage Site along the Northern coasts of The Netherlands, Germany and Denmark, is a very valuable, yet also highly vulnerable tidal flats area. It is noted for its ecological diversity and value, being a stopover for large numbers of migrating birds. The Wadden Sea is also used intensively for economic activities by inhabitants of the surrounding coasts and islands, as well as by the many tourists visiting the area every year.

A whole series of monitoring programmes of both ecological and socio-economic parameters is carried out by a range of governmental bodies and institutes, to study the natural processes occuring in the Wadden Sea ecosystems as well as the influence of human activities on those ecosystems. Yet, the monitoring programmes are scattered and it is difficult to get an overview of those monitoring activities or to get access to the data resulting from those monitoring programmes. The Wadden Sea Long Term Ecosystem Research (WaLTER) project aims to:

1. To provide a base set of consistent, standardized, long-term data on changes in the Wadden Sea ecological and socio-economic system in order to model and understand interrelationships with human use, climate variation and possible other drivers.
2. To provide a research infrastructure, open access to commonly shared databases, educational facilities and one or more field sites in which experimental, innovative and process-driven research can be carried out.

This presentation will introduce the WaLTER-project (2011-2015) and explain the rationale for this project. The presentation will focus on the distrubuted data access infrastructure which will be used for WaLTER. It is based on and makes use of the existing data access infrastructure of the Netherlands National Oceanographic Data Committee (NL-NODC), which has been operational since early 2009. The NL-NODC system is identical to and in fact developed by the European SeaDataNet project, furthering standardisation on a pan-European scale. The presentation will focus on the use of a distributed data access infrastructure to address the needs of different user groups such as policy makers, scientists and the general public. Benefits and pitfalls of using this infrastructure will be addressed.

Development of ITSASGIS-5D: seeking interoperability between Marine GIS layers and scientific multidimensional data using open source tools and OGC services for multidisciplinary research

Yolanda Sagarminaga, Marine Research division, Azti-Tecnalia, ysagarminaga@azti.es. SPAIN
Ibon Galparsoro, Marine Research division, Azti-Tecnalia, ysagarminaga@azti.es
Iratxe Orbe, Estudios GIS, iorbe@estudiosgis.com
Jose Antonio Sánchez, Estudios GIS, josean@estudiosgis.com

Since 2000, an intense effort was conducted in AZTI's Marine Research Division to set up a data management system which could gather all the marine datasets that were being produced by different in-house research projects.

For that, a corporative GIS was designed that included a data and metadata repository, a database, a layer catalog and search application, and an internet map viewer. Several layers, mostly dealing with physical, chemical and biological in-situ sampling, and basic and thematic cartography including bathymetry, geomorphology, habitat maps, and human pressure and activities maps, were successfully gathered in this system.

Very soon, it was realised that new marine technologies yielding continuous multidimensional data, sometimes called FES (Fluid Earth System) data, were difficult to handle in this structure. The data affected, mainly included numerical oceanographic and meteorological models, remote sensing data, coastal RADAR data, and some in-situ observational systems such as CTD's casts, moored or lagrangian buoys, etc.

A management system for gridded multidimensional data was developed using standardized formats (netcdf using CF conventions) and tools such as THREDDS catalog (UNIDATA/UCAR) providing web services such as OPENDAP, NCSS, and WCS, as well as ncWMS service developed by the Reading e-science Center.

At present, a system (ITSASGIS-5D) is being developed, based on OGC standards and open-source tools to allow interoperability between all the data types mentioned before. This system includes, in the server side, postgresql/postgis databases and geoserver for GIS layers, and THREDDS/Opendap and ncWMS services for FES gridded data. Moreover, an on-line client is being developed to allow joint access, user configuration, data visualisation and query and data distribution.

This client is using mapfish, ExtJS – GeoEXT, and openlayers libraries. Through this presentation the elements of the first released version of this system will be described and showed, together with the new functionalities to be developed in new versions that include among others, the integration of geoNetwork libraries and tools for both FES and GIS metadata management, and the use of new OGC Sensor Observation Services (SOS) to integrate non gridded multidimensional data such as time series, depth profiles or trajectories provided by different observational systems.

The final aim of this approach is to contribute to the multidisciplinary access and use of marine data for management and research activities, and facilitate the implementation of integrated ecosystem-based approaches in the fields of fisheries advice and management, marine spatial planning, or the implementation of the European policies such as the Water Framework Directive, the Marine Strategy Framework Directive or the Habitat Framework Directive.

Are the Data Really Available?

Dicky Allison, Woods Hole Oceanographic Institution, mallison@whoi.edu (USA)
Robert Groman, Woods Hole Oceanographic Institution, rgroman@whoi.edu
Cynthia Chandler, Woods Hole Oceanographic Institution, cchandler@whoi.edu
Shannon Rauch, Woods Hole Oceanographic Institution, srauch@whoi.edu
Danie Kinkade, Woods Hole Oceanographic Institution, dkinkade@whoi.edu
Peter Wiebe, Woods Hole Oceanographic Institution, pwiebe@whoi.edu
David Glover, Woods Hole Oceanographic Institution, dglover@whoi.edu
Charlton Galvarino, Second Creek Consulting, LLC, charlton@2creek.com

Making data available for re-use/re-purposing has recently become a driving force of data management. However, having data deposited into an archive or repository does not guarantee the data will be usable. Can the data be found and are they in a form that makes them useful? Data must be discoverable and understandable to be considered available and metadata are key to making data understandable and ultimately usable. Funded scientific projects are becoming more and more interdisciplinary and access to well-documented, metadata-rich data is crucial for a project's success.

Furthermore, as data curators, how do we design a user interface that fosters successful data discovery and promotes data comprehension? We recognize that a solution must include a data access system for human and machine clients that facilitates interoperability through standards adherence and vocabulary control. This is now an active area of research.

The Biological and Chemical Oceanography Data Management Office (BCO-DMO), created to serve data from research projects funded by the Biological and Chemical Oceanography Sections and the Office of Polar Programs Antarctic Organisms and Ecosystems Program of the US National Science Foundation (NSF), has been in existence since 2006. The BCO-DMO data system is open access and uses public domain software for its geospatial access. Metadata, which includes descriptions of acquisition and processing methodology, instrumentation, vocabulary, and mapping parameters, are stored in a MySQL database. The data themselves are made available via the JGOFS/GLOBEC[1] object-oriented, relational-like data management system and geospatial access (see Fig. 1) is provided by MapServer software (originally developed by the University of Minnesota[2]).

Since 2006, and in response to the evolving scientific and informatics community expectations, BCO-DMO has

- Enlarged and expanded its metadata complexity and content,
- Solidified the long-held philosophy of flexibility in data acceptance and partnership with the

contributing scientist[3],

- Developed an ontology[4], which, in conjunction with vocabulary control, is providing the underpinnings of interoperability as well as providing for an advanced faceted search capability, and
- Continued its intense effort into making all the data discoverable from a geospatial interface, including those data with no geospatial emphasis.

These improvements were designed to make the data more discoverable/available than they were before, but user feedback is essential to evaluating their effectiveness.

This presentation will feature examples of making diverse data types from a broad spectrum of ocean research disciplines both accessible and available, and, at the same time, solicit comments on the effectiveness of the endeavor.

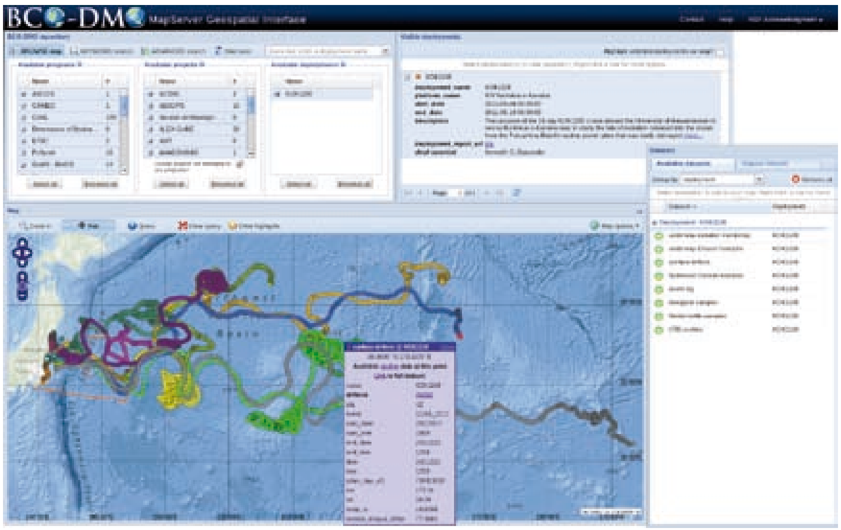


Fig. 1 - A screen shot of the BCO-DMO MapServer interface showing drifter tracks from one cruise following the March 2011 earthquake, tsunami and Fukushima Daiichi nuclear power facility disaster.

Acknowledgements. BCO-DMO is funded by the NSF Division of Ocean Sciences and the Office of Polar Programs and we are grateful for their support and encouragement.

References

[1] Fierl, G.R., James K.B. Bishop, David M. Glover, and Satish Paranjpe, 2004. JGOFS Data System Overview, http://globec.whoi.edu/globec-dir/doc/datasys/jgsys.html

[2] http://www.mapserver.org

[3] Glover, D.M., C.L. Chandler, S.C. Doney, K.O. Buesseler, G. Heimerdinger, J.K.B. Bishop, and G.R. Flierl, 2006. The U.S. JGOFS Data Management Experience, Deep-Sea Res. II, **53**: 5-7.

[4] Chandler, C.L., M.D. Allison, R. Groman, P. West, S. Zednick, and A. Maffei (2011), "Developing an Ontology for Ocean Biogeochemistry Data", Geophysical Research Abstracts, Vol. 13, EGU2011-8578, presented in Session "ESSI1: Informatics in Oceanography" at EGU General Assembly 2011, Vienna, Austria. http://bcodmo.org/sites/default/files/event/BCO_ontology_110329_prePrint.pdf.

Management of taxonomic species and observational data – case Baltic Sea phytoplankton

Kimmo Tikka, Finnish Meteorological Institute, kimmo.tikka@fmi.fi (Finland)
Guy Hällfors, (Finland)
Seija Hällfors, seija.hallfors@gmail.com, (Finland)
Maija Huttunen, maijahut@gmail.com, (Finland)
Riitta Olsonen, riitta.olsonen@gmail.com, (Finland)
Mikko Virkkunen, Simsoft Ltd., mikko.virkkunen@simsoft.fi (Finland)

Background

Phytoplankton occupies a central role in monitoring the state of lakes, estuaries, seas and oceans. In the context of the Baltic Sea, monitoring has developed into a comprehensive system since the Baltic Marine Environment Protection Commission (the Helsinki Commission, HELCOM) was founded in 1974. Regular assessments have been based on measurements made according to common standardized procedures. In order to be able to evaluate possible changes in the environment, long-term measurements of different environmental parameters are of utmost importance, among them especially phytoplankton. Phytoplankton species numbers and relative abundance, i.e. diversity, reflect the state of the environment, and have been used in HELCOM monitoring since 1974 (in ICES monitoring since the beginning of the 1900's).

History

Good taxonomical knowledge is a prerequisite for successful and meaningful monitoring. Crucial for this task is the management of taxonomy and the mutual understanding of species concepts and definitions, i.e. that everybody gives the same species the same name. The phytoplankton taxonomist community around the Baltic Sea was aware of the essence of the common species lists. Several actions have been taken since 1970's. RUBIN code (late 1980's) listed easy hierarchical 8 character codes. A huge step was the collection of G. Hällfors' Checklist (2004) for HELCOM containing the names of thousands of different phytoplankton species with their alias names (synonyms) and author information. In 1991 HELCOM's Phytoplankton Expert Group (PEG) was established with the main aim of standardizing methods of collection, counting and identification of phytoplankton species in the Baltic Sea. During annual meetings a common understanding of phytoplankton species and volume lists were compiled based on the lists already in use at the different laboratories participating in monitoring.

Workflow

For database systems the management of taxonomy is a demanding and interesting challenge. One should be able to trace changes in the taxonomy and at the same time save the original information of the observations. The user should get the information even for long time series in a uniform format, so that possible changes in the taxonomical information are correctly taken into account.

In our solution (SUMPPU system) we are able to insert both old and new data into the database using the species list valid for the counting date, and still the user gets all information according to the current species list (provided the determination was correct, in the first place).

PEG experts revise and make decisions of changes to the Baltic Sea phytoplankton list in their annual meeting. We update these changes into the database using an excel-worksheet, with the current species list as the starting point and describe the changes with the following operations:

- UPDATE: update non-key value attributes of a taxon
- INSERT: insert a new species e.g. a taxon with a new combination of key values
- FOLLOW: form a new taxon by changing some of the key values of an existing one
- MERGE: merge two taxa into one
- BRANCH: split a taxon into two different taxa e.g. combinations of key values
- DETACH: release the definition of a taxon so that the same combination of key values may be used later by another taxon

With these operations updated specieslist form a network of succeeding changes in the taxonomy, and the sequence of changes in taxonomy are precisely described and stored into a database. Our practical experience has shown us that a close cooperation with a taxonomist and a data manager is essential for success.

Our system ensures that all of observations, irrespective which were the taxa definitions used when the observation saved for the first time, can be brought out according to the current specieslist.

Discussion

Though we use here phytoplankton as an example, the same system is applied to zooplankton and benthos in the Sumppu database. Sumppu has its own, Finnish species list, which includes all PEG species and all G. Hällfors' Checklist species. The system has also information on whether a taxon occurs in other lists, as most are, e.g. WoRMS and ITIS

The framework we built into this system may be easily adapted to a variety of problems with changing definitions of key values.

For the environmental policy making it is most important that decisions are made relying on correct information about the changes in the nature. One must be certain about the correctness of the occurrence and abundance time series of indicator species even if new information on genera and species appear. Taxonomical knowledge, however, varies greatly between individual analysts and countries, and in time, with the advancement of taxonomy. "New species" are not necessarily new to the area, or science, but mostly just previously overlooked. Going far back in time, actually reduction of the data sets to their least common denominator may be required.

References

Hällfors, G. 2004. *Checklist of Baltic Sea Phytoplankton Species (including some heterotrophic prostistan groups)*, HELCOM Baltic Sea Environment Proceedings No. 95:1-208.

An interoperable infrastructure for the Italian Marine Research

Paola Carrara, CNR-IREA, carrara.p@irea.cnr.it (Italy)
Alessandro Sarretta, CNR-ISMAR, alessandro.sarretta@ismar.cnr.it
Alessandra Giorgetti, OGS, agiorgetti@ogs.trieste.it
Maurizio Ribera D'Alcalà, SZN, maurizio@szn.it
Alessandro Oggioni, CNR-IREA, oggioni.a@irea.cnr.it
Elena Partescano, OGS, epartescano@ogs.trieste.it

Background

The Italian maritime cluster represents for Italy an important economic sector contributing to 2.6% of national GDP, 11% of production in the industry of transport, and using almost 1% of the units of work identified in the country, share as high as 2% including the impact of upstream and downstream, for a total of around 480,000 employees (Cluster and maritime development in Italy and in the regions CENSIS - September 2011). In the European scenario Italy keeps the 1st place in Europe in terms of imports across the sea (185.4 million tonnes of freight), and the 3rd in terms of exports (47 million - not far from Germany and the Netherlands). Moreover Italy is in the first place in the passenger transport sector with 6.7 million people as a base and cruise destinations.

RITMARE (la Ricerca Italiana per il MARE – Italian Research for the sea) Flagship Project is one of the National Research Programmes funded by the Italian Ministry of University and Research, involving the whole marine research sector, including various public research bodies (CNR, OGS, INGV, ENEA, ISPRA, Stazione Zoologica) and Inter-university consortia (CoNISMa, CINFAI)l, as well as many private companies working in the sector and, as a result of technological transfer, enhancing the competitiveness of Italian industry.

Objectives

In its Blue paper (COM2007/575 of October 10th 2007) the European Commission highlighted the need to implement an integrated maritime and marine policy in order to "enhance Europe's capacity to face the challenges of globalisation and competitiveness, climate change, degradation of the marine environment, maritime safety and security, and energy security and sustainability." It stated further that such a policy "must be based on excellence in marine research, technology and innovation".

The aim of RITMARE is to implement what is suggested in the Blue Paper in terms of research and innovation, by means of a national programme of scientific and technological marine research.

More specifically, RITMARE has been structured around the following three objectives:

- to support integrated policies for the safeguard of the environment (the health of the sea);

- to enable sustainable use of resources (the sea as a system of production);
- to implement a strategy of prevention and mitigation of natural impacts (the sea as a risk factor).

Structure

The RITMARE project is organised into seven sub-projects:

- SP1. Maritime Technologies for the development and construction of a Demonstration Vessel.
- SP2. Technologies for Sustainable Fishing.
- SP3. Planning of the Maritime Space in Coastal Waters.
- SP4. Planning of the Deep Marine Environment and the Open Sea.
- SP5. Observation System for the Marine Mediterranean Environment.
- SP6. Research, Training and Dissemination Structures.
- SP7. Interoperable Infrastructure for the Observation Network and Marine Data.

Interoperable Infrastructure for the Observation Network and Marine Data

This work presents the sub-project 7 of RITMARE, that aims at designing and developing an infrastructure that allows coordinating and sharing data, processes and information produced in the various different sub-projects, without forcing the pre-existing practices and the enabling technologies already adopted by the RITMARE scientific communities.

The great variety of actors is reflected in the coexistence, within the project, of different data, formats, practices, approaches, needs and scopes.

The RITMARE information infrastructure will consist of a network of interconnections and tools, to help the data flows, information and service management, to meet the needs of the involved scientific communities and to support their growth.

The design of the information Infrastructure will take place through some steps, which are connected to the objectives of four work packages (WP):

- a robust preliminary analysis of requirements of the project participants and on their goals as regards the infrastructure, and an analysis of the solutions found not only by marine research (WP1);
- a survey on data, technologies, processes and standards used, to share multidisciplinary observations, data, metadata and products as well as solution adopted by the participants (WP2);
- the definition and application of a data policy, in order to support the reuse of data and resources, respecting the constraints, the needs of the participants and the existing ruling conventions (WP3);
- the design, development and implementation of the infrastructure, tested on and complemented by prototypes and demonstrators (WP4), besides technological support tools; with the envision of maintenance and sustainability plans.

The approach to the infrastructure design is intended as user-driven, while based on a constant interaction with the experts' requirements, on the cooperation in setting goals and testing solutions, through the use of participative tools and methods. This will allow to achieve the fulfilment of the requirements coming from all the involved communities and to facilitate a common growth, while preserving the peculiarities of individual collaborators.

Scalability, flexibility, distribution and decoupling are the conditions to ensure the progressive growth of the infrastructure during the working period and to facilitate its sustainability after the project end. Maintenance and sustainability plans will be defined in order to reach these aims.

The RITMARE infrastructure will not only focus on Italian capabilities and experiences on the marine research area; it will also integrate in the wider European and international domain. The development of the sub-project must therefore be consistent with the technological context and international regulations. To make it possible, international initiatives (GEO / GEOSS), European directives (INSPIRE, SPI, MFSD, WFD, Fishery Directive) and previous experiences in European projects (SeadataNet) will provide basic references and will be considered as sources of valuable information for the infrastructure design and maintenance.The focus of this work will be on the first two Work Packages, presenting and describing the results from the preliminary analysis of requirements of the project participants and from the survey on data, technologies, processes and standards used. First steps toward the design, development and implementation of a prototype of the infrastructure will also be presented.

Information system of marine environment indicators in Croatia

Vlado Dadic, Institute of oceanography and fisheries, Split, Croatia, dadic@izor.hr
Damir Ivankovic, Institute of oceanography and fisheries, Split, ivankovic@izor.hr
Josip Bašic, Institute of oceanography and fisheries, Split, basic@izor.hr
Ivan Vucic, Institute of oceanography and fisheries, Split, vucic@izor.hr

Institute of oceanography and fisheries (IOF) has very succeseful collaboration with Croatian environmental agency (AZO) on reporting of marine indicators for almost ten years. Therefore, Croatian Government determined IOF as Croatian Marine referral centre (RCM) on March 15, 2012. RCM provides technical and scientific support to the AZO, collects, analysis and reports of national indicators of the marine environment, fisheries and aquaculture, as well as EEA marine core indicators (EIONET) and MSFD (DIKE) tables. In addition, IOF has developed and performed web oriented Marine information system (MIS) as part of the Environmental Information System (EIS) for the purposes of fulfilling reporting obligations on the state of the marine environment of national and international requirements (e.g. implementation of MSFD in Croatia).

For this purposes web oriented Oceanographic database of the Adriatic Sea (MEDAS: Http:// www.izor.hr/medas) has been upgraded with a geospatial layers and two sub-databases: Database of marine indicators and (Http:// www.izor.hr/azo) and database of Water quality at the Croatian beaches (Http://www.izor.hr/kakvoca).

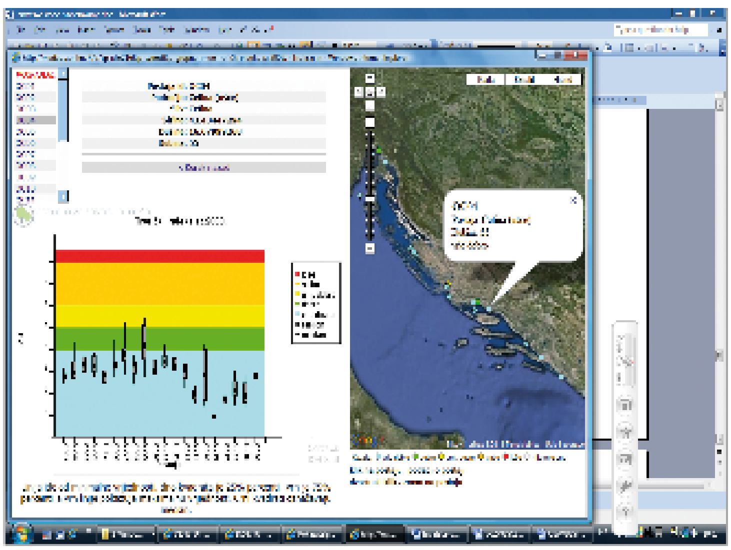


Fig. 1 - Trophic index at the monitoring stations and the ecological status of the station at the mouth of the Cetina river.

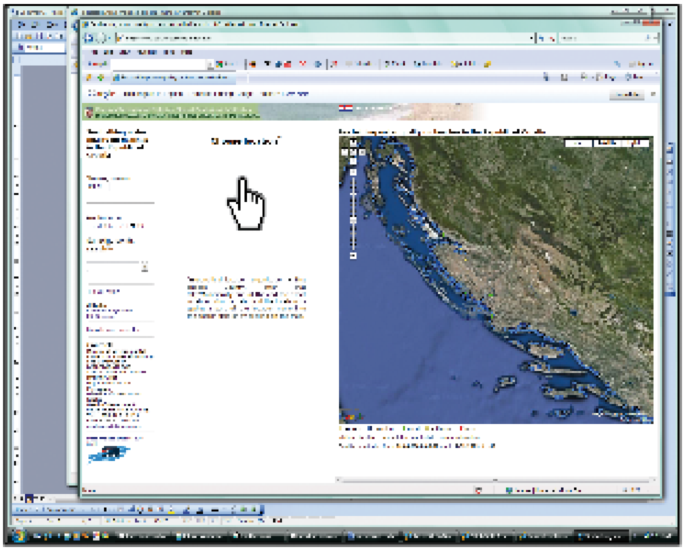


Fig. 2 - Croatian beaches with sea bathing water quality monitoring.

The databases are currently working on Database management systems (DBMS) Oracle 10.2 with the automatic storage management (ASM), Oracle Application Server 10.2, Geoserver and the CentOS 4.4 operating system (Linux).

As in databases about 40 different marine indicators and more than 800 Croatian beaches with monitoring of water quality have been set up, about 50 experts from different fields have been responsible for uploading data and reports in databases. As they uploaded data and reports in databases through Internet, special attention on security of databases, data and output results has been taken in account. For this purposes multiple user groups with different rights was created for access and manipulate with data and information. Fig. 1 shows example of output results from Database of marine indicators, and Fig. 2 shows list of Croatian beaches with sea bathing water quality monitoring from Database of water quality.

GO-SHIP: building a time series of a suite of ocean properties

Stephen Diggs, Scripps Institution of Oceanography / UCSD, sdiggs@ucsd.edu
Bernadette Sloyan, Centre for Australian Weather and Climate Research, bernadette.Sloyan@csiro.au

In 2013, the international ocean community will complete another hydrographic survey of the global oceans and begin the next global survey. Despite numerous technological advances over the last several decades, ship-based hydrography remains the only method for obtaining high-quality, high spatial and vertical resolution measurements of a suite of physical, chemical, and biological parameters over the full water column. Ship-based hydrography is essential for documenting ocean changes throughout the water column, especially for the deep ocean below 2 km (52% of global ocean volume not sampled by profiling floats).

The time-series of ocean properties derived from the global survey do provide significant insight into the importance of the ocean in climate and climate variability. For example, documentation of substantial changes in the oceanic inorganic carbon content, driven by both the uptake of anthropogenic CO2 and natural variability ; evidence of large-scale changes in oceanic oxygen concentrations; near global-scale warming of abyssal waters of Antarctic origin, and freshening of these waters in deep basins adjacent to Antarctica; intensification of the global hydrological cycle; reduction in the lower limb of the Meridional overturning circulation; and estimates of water mass formation rate.

This presentation will discuss the success of the data collection and rapid availability of the data to the international community.

SESSION 3

Data Services in ocean science

- Standards and quality-assurance issues
- Services and Visualisation tools
- User oriented services

ORAL PRESENTATIONS

IODE Data Quality Flag Scheme

Cynthia L. Chandler, Woods Hole Oceanographic Institution, cchandler@whoi.edu (USA)
Sergey Konovalov, Marine Hydrophysical Institute, sergey@alpha.mhi.iuf.net (Ukraine)
Hernan Garcia, US National Oceanographic Data Center, hernan.garcia@noaa.gov (USA)
Reiner Schlitzer, Alfred Wegener Institute for Polar and Marine Research, Reiner.Schlitzer@awi.de (Germany)
Laure Devine, Department of Fisheries and Oceans, Laure.Devine@dfo-mpo.gc.ca (Canada)
Gwen Moncoiffé, British Oceanographic Data Centre, gmon@bodc.ac.uk (UK)
Toru Suzuki, Marine Information Research Center, suzuki@mirc.jha.jp (Japan)
Alex Kozyr, Oak Ridge National Laboratory, kozyra@ornl.gov (USA)
Greg Reed, Australian Ocean Data Centre Joint Facility, greg@metoc.gov.au Australia)

The objective of oceanographic data quality control is to ensure data consistency within a single data set or within a collection of data sets, and to ensure that the quality and any errors within the data are apparent to the user. Data quality information is essential to assess the suitability of the data for a task. Ideally, data quality flags provide the user of the data with clear information about actions taken to assess the quality of the original data and any subsequent actions performed to modify the data (e.g., interpolation to replace missing data values). Data quality assessment procedures such as flagging data values to indicate their quality, reliability, or checks that have been carried out, or altering values after checking, filling in data gaps, etc., can vary from project to project, between different laboratories and across data centers.

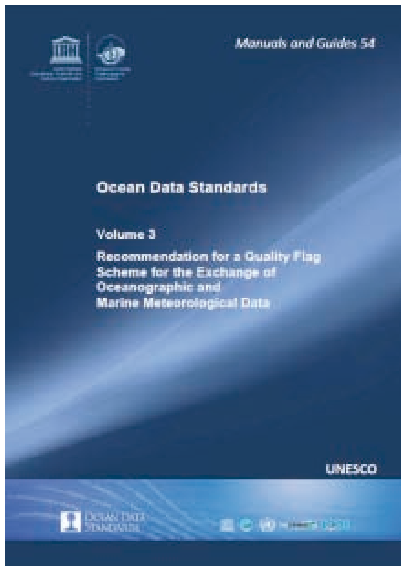


Fig. 1 - The QC flag scheme is available as Manuals and Guides 54, Ocean Data Standards volume 3.

The purpose of the recommended IODE Quality Flag (QF) scheme is to define a common set of quality flags that could be used by data centers and projects. The scheme is recommended to facilitate the exchange of data between systems, or for groups that do not already use a quality flag scheme. There is no suggestion that current schemes must be replaced by this new scheme.

The IODE QF scheme defines a common set of quality flags that can be used by individual researchers, collaborative groups, projects or large data centers. The scheme comprises two levels designed to facilitate the exchange and integration of multi-disciplinary oceanographic and marine meteorological data. The first, or primary, level (see Tab. 1) defines the data quality flags only, while the secondary level complements the first level by providing the justification for the quality

Value	Primary-level flag short name	Definition
1	Good	Passed documented required QC tests
2	Not evaluated, not available or unknown	Used for data when no QC test performed or the information on quality is not available
3	Questionable/suspect	Failed non-critical documented metric or subjective test(s)
4	Bad	Failed critical documented QC test(s) or as assigned by the data provider
9	Missing data	Used as place holder when data are missing

Tab. 1 - Primary level quality flags are unambiguous and limited and fixed in number.

flags, based on quality control tests, data processing history or other important data provenance details.

The design of the IODE QF scheme was based upon an extensive review of existing quality flag schemes. None of the reviewed schemes met all advantages stated below. The advantages of this two-level scheme are described below.

- A small and fixed number of unambiguous flags at the primary level can be justified by the details in the second level.
- Primary level flag values are numeric and ordered such that increasing quality flag values indicate a decreasing level of quality. This supports the identification of all data that meet a minimum quality level and assignment of quality flags to calculated parameters.
- The scheme is universal; it can be applied to all types of data enabling exchange and integration of multi-disciplinary data.
- Existing QF schemes can be mapped to the proposed scheme with no information loss. This is especially true when information on the applied tests is delivered by data providers.
- If data sets with different QF schemes are merged into one data set, all pre-existing quality flags can be preserved and it is possible to apply new quality tests to the merged data and reflect those results in the two-level IODE QF scheme as well.

The IODE QF scheme was recommended for adoption as a standard at the IODE XXII conference in March 2013 (Recommendation IODE-XXII.5; IOC/IODE-XXII/3 Report, 2013). While the main intent of the IODE QF scheme is to facilitate the exchange and integration of oceanographic and marine meteorological data, there is also the possibility that data managers, curators and archivists who do not already use a QF scheme may choose to adopt this one.

Acknowledgements. The authors gratefully acknowledge the support for several workshops on the Quality Control of Chemical Oceanographic Data Collections that was provided by the International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission (IOC) of UNESCO.

References

IOC/IODE-XXII/3 Report. (2013) IOC Committee on International Oceanographic Data and Information Exchange, Twenty-second Session, Ensenada, Mexico, 11-15 March 2013, Paris, UNESCO, 28 March 2013. http://www.iode.org/index.php?option=com_oe&task=viewDocumentRecord&docID=10583

Ocean Data Standards volume 3: Recommendation for a Quality Flag Scheme for the Exchange of Oceanographic and Marine Meteorological Data, available as UNESCO/IOC Manuals and Guides No. 54 - Volume 3, Version 1, 18 April 2013, http://www.iode.org/mg54_3

Workshop reports and related information

IOC Workshop Report No. 228 (2010) First IODE Workshop on Quality Control of Chemical Oceanographic Data Collections, IOC Project Office for IODE, Oostende, Belgium, 8-11 February 2010, Paris, UNESCO, 26 March 2010, (English, http://www.iode.org/index.php?option=com_oe&task=viewDocumentRecord&docID=5109)

IOC Workshop Report No. 253 (2012) Second IODE Workshop on Quality Control of Chemical and Biological Oceanographic Data Collections, IOC project Office for IODE, Oostende, Belgium, 22-24 October 2012. Paris, UNESCO, 15 November 2012, (English, http://www.iode.org/index.php?option=com_oe&task=viewDocumentRecord&docID=9757)

Manual of Quality Control Procedures for Validation of Oceanographic Data. UNESCO-IOC. Manual and Guides 26 (1993), link: http://www.jodc.go.jp/info/ioc_doc/Manual/mg26.pdf

Marine Regions: towards a standard for georeferenced marine names

Simon Claus, Flanders Marine Institute, simon.claus@vliz.be, Belgium
Francisco Souza Dias, Flanders Marine Institute, francisco.souzasias@vliz.be, Belgium
Bart Vanhoorne, Flanders Marine Institute, bart.vanhoorne@vliz.be, Belgium
Nathalie De Hauwere, Flanders Marine Institute, nathalie.dehauwere@vliz.be, Belgium
Francisco Hernandez, Flanders Marine Institute francisco.hernandez@vliz.be, Belgium

Geographic Information Systems have become indispensable tools in managing and displaying marine data. However, a unique georeferenced standard of marine placenames and areas is not available, hampering several marine geographic applications, such as the linking of these locations to databases for data integration. In order to improve the current situation, we developed "Marine Regions", a standard, hierarchical list of geographic names, linked to information and maps of the geographic location of these names, freely available at http://www.marineregions.org. The objectives of Marine Regions are to improve access and clarity of the different geographic marine names such as seas, sandbanks, ridges and bays and to display univocally the boundaries of marine biogeographic or managerial marine areas.

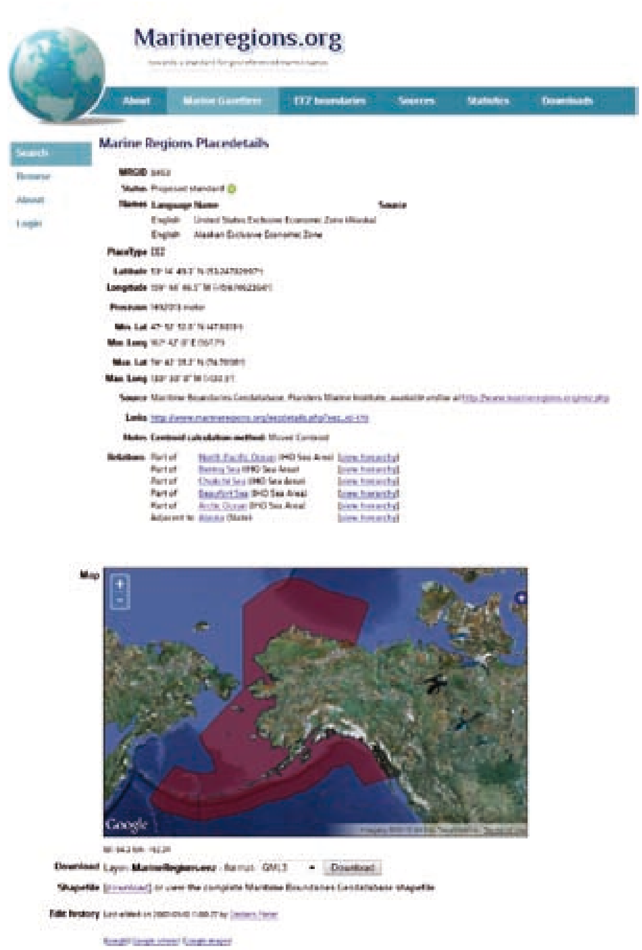


Fig. 1 - Screenshot of the "Marine Regions", displaying the placedetails, relations, boundaries and download links of the Alaskan part of the US Exclusive Economic Zone.

Structure & technology

All geographic objects of the Marine Regions database have a unique identifier, called the MRGID (= Marine Regions Geographic Identifier), used for locating the geographic resources on the web. The different geographic objects are determined by a placetype and coordinates. While the coordinates can be represented as different vector data types (being a point, a multipoint, polyline or a polygon), a placetype provides contextual information to the geographic objects, for example a sea, a bay, a ridge, a sandbank or an undersea trench. Not only physical placetypes are considered, but also administrative placetypes, like countries, EEZ's, fishing zones or territorial seas can be stored in the database.

The actual name of the geographic objects is stored as a different entity, allowing thus multiple naming for one geographic object (i.e. dealing with different languages). It is also possible to define different relations between the geographic objects (part of, partly part of, adjacent to, similar to, streams through or flows out). Such a structure allows the user to group joint geographic units and to create a hierarchical classification of different places. Once logged in, geoobjects can be edited through the webinterface of Marine Regions. If a point, a line or a polygon is available for a geographic object, the geographic position of the object will be visualised on an interactive web mapping interface (Fig. 1).

The geographic webinterface is based on the OpenLayers technology. All shapefiles containing the polylines and polygons are uploaded to a local Geoserver installation, allowing to distribute the geographic objects as different Web Mapping Services (WMS) and Web Feature Services (WFS). The polygons of the different geographic classifications can be downloaded from the website as individual shapefiles.

Content of Marine Regions

At the moment Marine Regions lists and provides geographical information and relations on more than 32,604 placenames, representing 25,487 marine geographic places. The number of records is an approximate value based on the database on January 15th, 2013. We can distinguish three different categories of information: regional checklists, global checklists and global or regional spatial marine classifications.

The regional checklists include detailed information from the North Sea, the Black Sea and the Antarctic region. The main global checklists integrated in Marine Regions are the IHO-IOC GEBCO Gazetteer of undersea feature names, marine placenames from the Aquatic Sciences and Fisheries Abstracts (ASFA) thesaurus and several distribution records from the World Register of Marine Species.

Marine Regions gives access to 12 marine geographic regional or global marine classifications including the boundaries of the major oceans and seas of the world, defined by the International Hydrographic Organisation (IHO), the Large Marine Ecosystems of the World, the Longhurst Biogeographical Provinces or the Marine Ecoregions of the World. The database contains also 5,597 polygons of geographic places. Marine Regions gives also access to the database of the Exclusive Economic Zones of the world making them available to the scientific community. As this information was not freely available, two global GIS covers, containing the lines of the maritime boundaries of countries and the polygons of the EEZ's have been calculated. The first step in the creation of the geodatabase was the integration of information already available. In a second phase the database of negotiated treaties from the United Nations Convention on the Law of the Sea (UNCLOS) was con-sulted and imported into a GIS. The geographic coordinates from the documents were converted to decimal degrees and imported into a database. If no treaty was available from UNCLOS, the 200 nautical miles buffer around a country was calculated. If the distance between two countries was less than 400 nautical miles, the maritime boundary was calculated as the median line between both countries.

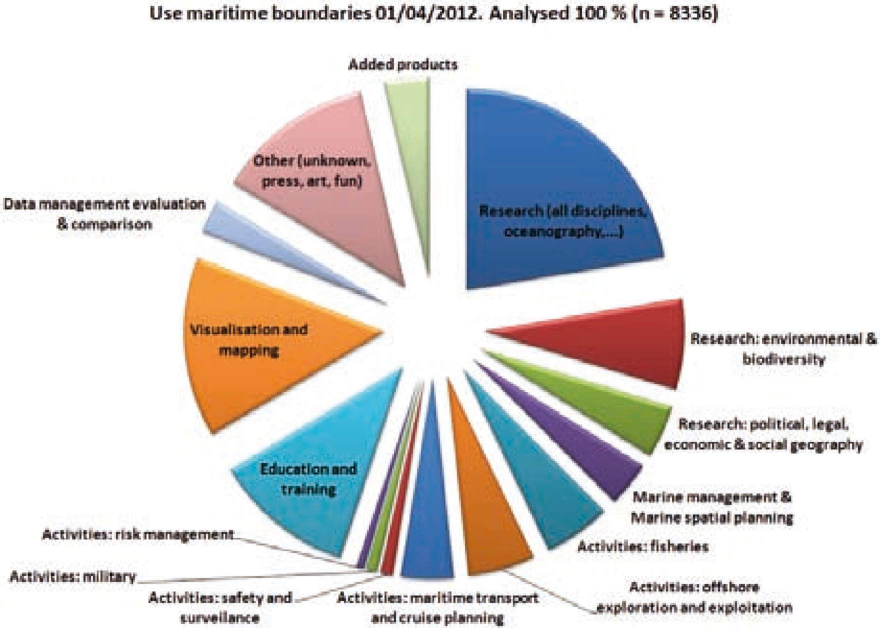


Fig. 2 - Table indicating different usages of the Marine Regions data.

Users

The web statistics and downloads of the system have been monitored since 2008. Between 2008-11-19 and 2013-01-05 18,967 shapefiles were downloaded, with 2,584 downloads in 2009, and 6,294 downloads in 2012. These statistics do not include the downloads or consultation of the geographic objects through the available WMS or WFS services, as these services are more difficult to monitor. In January 2013 the website received 90,284 hits from 3,602 unique visitors. We analysed the purpose of download of 8,336 downloads between 2008-11-19 and 2012-03-30 (Fig. 2). This represents 44% of the total recorded downloads between 2008-11-19 and 2013-01-05. The reason for download is a required but free text field when a GIS layer is downloaded from Marine Regions. We grouped the various reasons for download into different categories. Most of the downloads were performed for research purposes (33.9%) with main disciplines oceanography (22.2%), environmental and biodiversity sciences (7.3%) and political and economic geography (4.2%). In 2.7% of the cases the data was used for standardization purposes in marine data management. These two categories constitute the reasons for which Marine Regions was developed but represent less than half of total downloads. Over 21% of the geographic information was downloaded for specific purposes related to different marine and maritime activities including fisheries (5.4%), offshore exploration (5.4%), maritime transport and cruise planning (4.2%), marine management and marine spatial planning (3.7%), surveillance (1.0%), military use (0.9%) or risks assessments (0.6%). The data was also extensively used for educational or visualization purposes (26.0%). In 3.6% of cases, added products were created out of the geographic data, for example incorporation of the geographic data in an online coastal or marine atlas. Finally unknown reasons and other specific reasons like press releases or artist impressions constituted 12.7% of the downloads. The user analysis indicates that a very large user community uses the system for different applications. It is by integrating feedback from this community, that Marine Regions intends to reach his overall objective, which is evolving towards a globally accepted standard for georeferenced marine regions.

Trendylyzer: a Long-Term Trend Analysis on Biogeographic Data

Ward Appeltans, Peter Pissierssens, IOC-UNESCO,
{w.appeltans, p.pissierssens}@unesco.org (Belgium)

Gianpaolo Coro, Angela Italiano, Pasquale Pagano, ISTI-CNR,
{gianpaolo.coro, angela.italiano, pasquale.pagano}@isti.cnr.it

Anton Ellenbroek, FAO, anton.ellenbroek@fao.org

Tom Webb, University of Sheffield, t.j.webb@sheffield.ac.uk

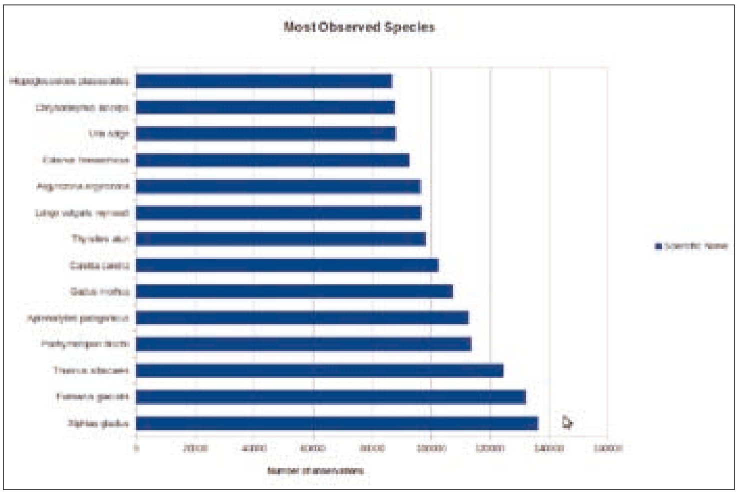


Fig. 1 - Number of records for the 15 most observed species in OBIS.

The United Nations General Assembly (A/RES/63/111)[1] expressed its serious concern over the current and projected adverse effects of climate change, pollution, overfishing and habitat destruction on the marine environment and marine biodiversity. As a result, the UN established a Regular Process for Global Reporting and Assessment of the State of the Marine Environment Including Socioeconomic Aspects, and will publish its 1st World Ocean Assessment in 2014[2]. An entire section will be devoted to the status of marine biodiversity. A basic metric of biodiversity is species composition. High species diversity is essential for maintaining ecosystem functioning and could be regarded as one of the most important indicators on ocean's health. Unfortunately, assessing the global status of marine species is difficult because historical data and accurate long-term time series on species occurences are scarce. One third to two thirds of marine species are yet undiscovered[3], and those species that are described are often known from a single observation or are so rare that a scientific basis to make a meaningful assessment of their status is lacking.

The Ocean Biogeographic Information System[4] (OBIS) is an important source of information on marine species diversity. OBIS is world's largest global online open-access database on the diversity, distribution and abundance of all marine life and can provide an important baseline, against which future change can be measured. OBIS is one of the most used resources in marine science and has been involved also in ecological modelling[5] as well as in marine monitoring systems[6]. OBIS is the data legacy of the decade-long Census of Marine Life, a US$ 650 million foundation-led project that gathered data from 2,700 scientists and 540 expeditions, and now operates under UNESCO's Intergovernmental Oceanographic Commision as part of its International Oceanographic Data and Informatoin Exchange (IODE) programme. OBIS has continued to grow and has established a network of hundreds of data providers around the globe. In total, it now integrates 1,130 datasets and holds 35 million observations of 120,000 marine species.

In this paper we present Trendylyzer, a new marine species trend analysis tool using data from OBIS. The aim of Trendylyzer is to provide indicators for use in marine biodiversity assessments. Trendylyzer is a tool developed within the D4Science e-Infrastructure[7] (currently co-funded by the EU iMarine project[8]).

This is a Hybrid Data Infrastructure that aims at supporting large-scale resource sharing (both hardware and software) and allows data to be processed with distributed computing. Data can also be enriched with data coming from multiple sources, which are accessible through the e-Infrastructure. Furthermore, the e-Infrastructure allows for the creation of Virtual Research Environments (VREs), which are fully equipped web-based cooperation environments. Trendylyzer will allow users to access the data through the e-Infrastructure, specify filters and groupings, select a statistical analysis model, adjust the algorithm parameters (if any), and monitor the progress. The results will be represented in a variety of formats and the selection of data, algorithms and results will be summarised in a report.

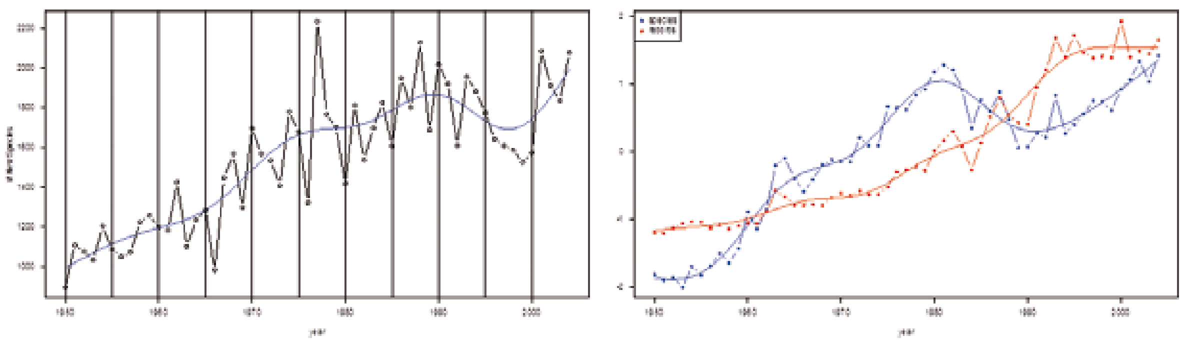


Fig. 2 - Number of new species reported in OBIS per year (left)
and number of globally reported species (blue line) and occurrence records (red line) per year (right).

Trendylyzer will aim to answer questions such as:

1. Which are the most common marine species globally or by region and has this been changing through time?

2. What are the current species gaps and how fast is this gap been filled with records of new species in OBIS?

3. Do threatened or endangered species (based on the IUCN red list) occur in one or more Marine Protected Areas (MPAs) and do the MPAs occur in the center or in the outer limits of the species population range?

One of the challenges will be to define « common species » and to take into account sampling effort. Fig. 1 ranks the most observed species in OBIS. The list was produced by counting the observations for each species. We merged the records that reported the same values for longitude, latitude, depth, collector and recording time to filter out potential duplicates. Fig. 2 shows the number of new species recorded in OBIS per year (left) and the global trend of the number of species (blue line) and observations (red line) per year (right). Interestingly, the number of records increases steadily, until it begins to level off around 1990 and the number of species declines through the 1980s, but then increases subsequently.

A trend graph like the global number of species observed in each year is obviously biased by the occurrence of rare species, but also by new species recorded in the database. Hence, a better approach to discover shifts in biodiversity is to look at the pool of most common species. The presented charts constitute preliminary analysis and future work will concentrate on more complex analysis involving (i) a study of the trends per year for the most observed species, (ii) interactive visualization facilities to fit data with proper representation metaphors, (iii) representations of new species reports per geographical area and taxonomic group, and (iv) species status investigation based on IUCN indications for MPAs. Trendylyzer is going to be released as an interactive web application on the i-Marine project portal in a specific VRE, which will allow users to produce charts and automatic reports and to share them with the i-Marine community of practice.

References

[1] http://www.un.org/Depts/los/general_assembly/general_assembly_resolutions.htm

[2] http://www.un.org/Depts/los/global_reporting/global_reporting.htm

[3] Appeltans et al (2012). The magnitude of global marine species diversity. Current Biology, 22(23): 2189-2202

[4] Intergovernmental Oceanographic Commission (IOC) of UNESCO. The Ocean Biogeographic Information System. Web. http://www.iobis.org.

[5] Owens, Bentley, and Peterson. (2012). Predicting suitable environments and potential occurrences for coelacanths (latimeria spp.). Biodiversity and Conservation, 21: 577-587

[6] Grothues (2009). A review of acoustic telemetry technology and a perspective on its diversification relative to coastal tracking arrays. In Tagging and Tracking of Marine Animals with Electronic Devices, pp 77-90

[7] Candela, Castelli, and Pagano. (2009). D4Science: an e-Infrastructure for Supporting Virtual Research Environments. In Agosti, Esposito, and Thanos (Eds), Post-proceedings of the 5th Italian Research Conference on Digital Libraries - IRCDL 2009, pages 166-169

[8] The i-Marine European Project, 2011. http://www.i-marine.eu.

Variational data analysis for generating ocean climatologies (DIVA) and web-based distribution of data products (OceanBrowser)

Alexander Barth, ULG/GHER, a.barth@ulg.ac.be (Belgium)
Charles Troupin, IMEDEA, ctroupin@imedea.uib-csic.es
Aida Alvera-Azcárate, ULg/GHER, a.alvera@ulg.ac.be
Jean-Marie Beckers, ULg/GHER, jm.beckers@ulg.ac.be

Variational interpolation in an n-dimensional space

DIVA-nd (Data-Interpolating Variational Analysis - *n*-dimensional) is a tool that allows to generate gridded data products from *in situ* observations. DIVA-nd extends the 2d capabilities of the DIVA tool and allows the interpolation of observations on curvilinear orthogonal grids in an arbitrary high dimensional space by minimizing a cost function. This cost function penalizes the deviation from the observations, the deviation from a first guess and abruptly varying fields based on a given correlation length (potentially varying in space and time). Physical constraints can be added to this cost function such as an advection constraint, diffusion or source terms. One major advantage of the method is that it naturally decouples basins that are not connected and where water masses often have very different properties. Individual elements of the *a priori* and *a posteriori* error covariance matrix can also be computed, in particular expected error variances of the analysis. The benefit of this multidimensional approach over the traditional 2d analysis (stacking of horizontal layers) is that the analysis is smooth in all dimensions. However the computational cost is increased compared to the stacking of horizontal layers.

Primal (problem solved in the grid space) and dual formulations (problem solved in observation space) are implemented using either direct solvers (based on Cholesky factorization) or iterative solvers (conjugate gradient method). In most applications the primal formulation with the direct solver is the most efficient method, especially if an *a posteriori* error estimate is needed. However, for correlated observation errors the dual formulation with an iterative solver is more efficient.

Validation of the method and improvement relative to 2d analysis

The method is tested by using pseudo observations from a global model. The distribution of the observations is based on the ARGO float positions. The benefit of the 3-dimensional analysis (longitude, latitude and time) compared to 2-dimensional analysis (longitude and latitude) and the role of the advection constraint are highlighted.

The tool DIVA-nd is developed in the frame of the SeaDataNet 2 project and distributed under the terms of the GPL license (http://modb.oce.ulg.ac.be/mediawiki/index.php/DIVAnd).

Distribution of data products

Data products generated by DIVA and DIVA-nd are distributed by the OceanBrowser portal (http://gher-diva.phys.ulg.ac.be/web-vis/). This web interface is developed to visualize gridded oceanographic data sets. Those data sets are typically four-dimensional (longitude, latitude, depth and time).

OceanBrowser allows one to visualize horizontal sections at a given depth and time to examine the horizontal distribution of a given variable. It also offers the possibility to display the results on an arbitrary vertical section. To study the evolution of the variable in time, the horizontal and vertical sections can also be animated. The user can customize the plot by changing the color-map, the range of the color-bar, the type of the plot (linearly interpolated color, simple contours, filled contours) and download the current view as a simple image or as Keyhole Markup Language (KML) file for visualization in applications such as Google Earth. The data products can also be accessed as NetCDF files and through OPeNDAP.

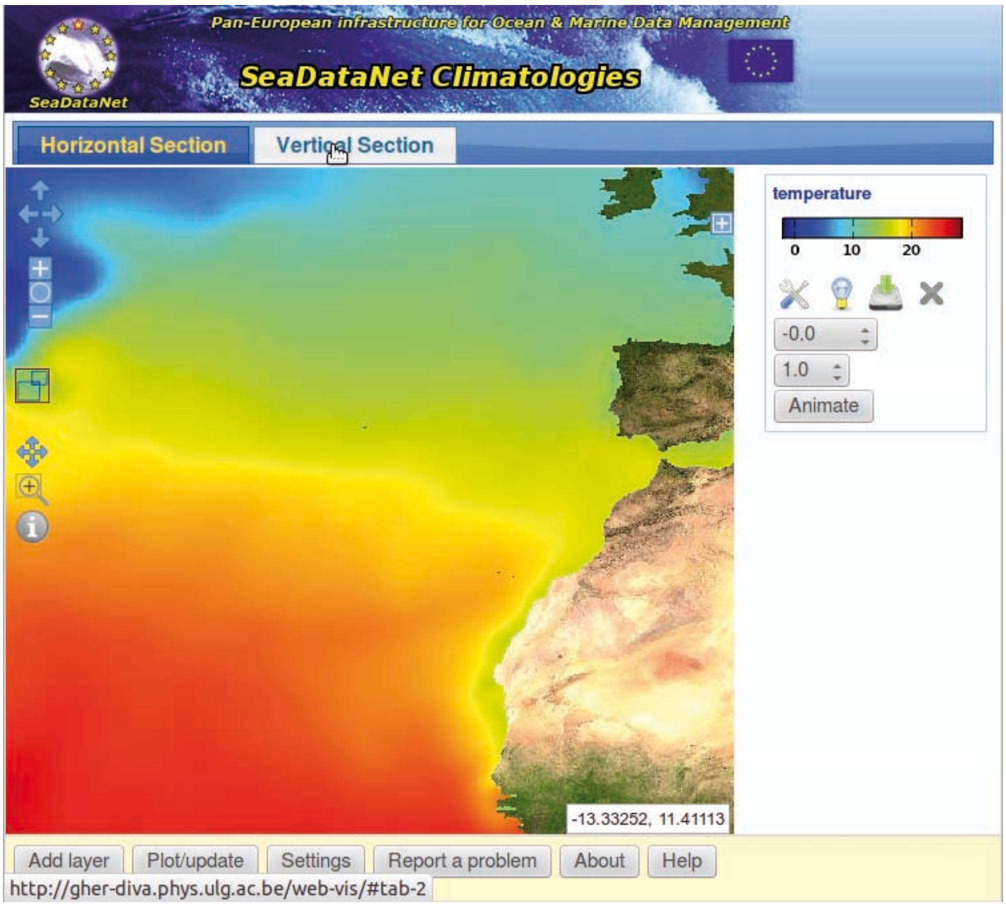


Fig. 1 - A temperature climatology visualized in OceanBrowser.

A new API for accessing ODV data collections from C++ and Java

Reiner Schlitzer, AWI, Reiner.Schlitzer@awi.de (Germany)
Michael Menzel, AWI, Michael.Menzel@awi.de
Stephan Heckendorff, AWI, Stephan.Heckendorff@awi.de

Many large and important oceanographic datasets are commonly distributed as Ocean Data View (ODV) data collections. Examples are the World Ocean Atlas 2009 created by the U.S. NODC containing standard depth hydrographic data on a global 1x1° grid, the SOCAT v1.5 collection of more than 6 million CO2 data in ocean surface water as well as the overlying atmosphere, or the Coriolis Ocean Database for ReAnalysis - CORA-3.4 containing more than 6.2 million original temperature and salinity profiles for the 1990 – 2012 time period. These and many more ODV data collections can be found at http://odv.awi.de/en/data/ocean/. Advantages of the ODV collection format are dense storage, easy extensibility and fast access to arbitrary data records, even in very large collections. The disadvantage of the ODV collection format has been in the past that ODV itself was the only available software that could open such collections and access the data inside.

While ODV is freely available, widely used and providing a broad range of analysis and visualization options, there are still many application-specific use-cases not supported by the software. Therefore it is desireable if users could develop their own software for opening ODV data collections and processing the data according to user specific needs. To satisfy this frequently expressed user demand we have developed a new ODV application programming interface (ODV API). This API is released at the IMDIS 2013 conference and is presented for the first time. In this talk we describe the capabilites of the API and discuss a number of example use-cases in C++ and Java.

GLOBE 3D: an new software for Geosciences data processing and 3D/4D viewing

Jean-Marc Sinquin, IFREMER, NSE/ILE, jmsinqui@ifremer.fr (France)
Jordi Sorribas, Unidad de Tecnología Marina, (CSIC), sorribas@utm.csic.es (Spain),
Paolo Diviacco, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, pdiviacco@ogs.trieste.it (Italy),
Matthias Baeye, MUMM - Management Unit of the North Sea Mathematical Models, matthias.baeye@mumm.ac.be (Belgium)
Gael Quemener, IFREMER, NSE/ILE, Plouzane, gael.quemener@ifremer.fr (France),
Benoit Loubrieu, IFREMER, NSE/ILE, Plouzane, benoit.loubrieu@ifremer.fr (France)

Within EUROFLEETS project, and linked to EMODNet and Geo-Seas European projects, GLOBE (GLobal Oceanographic Bathymetry Explorer) is an innovative and generic software combining all necessary functionalities for cruise preparation, for collection, linking, processing and display of scientific data acquired during sea cruises, and for export of data and information to the main marine data centers and networks.

The first version was delivered by the end of 2012 and was dedicated to MBES (Multi Beam Echo Sounder) data processing. The new releases are designed to accept further functionalities

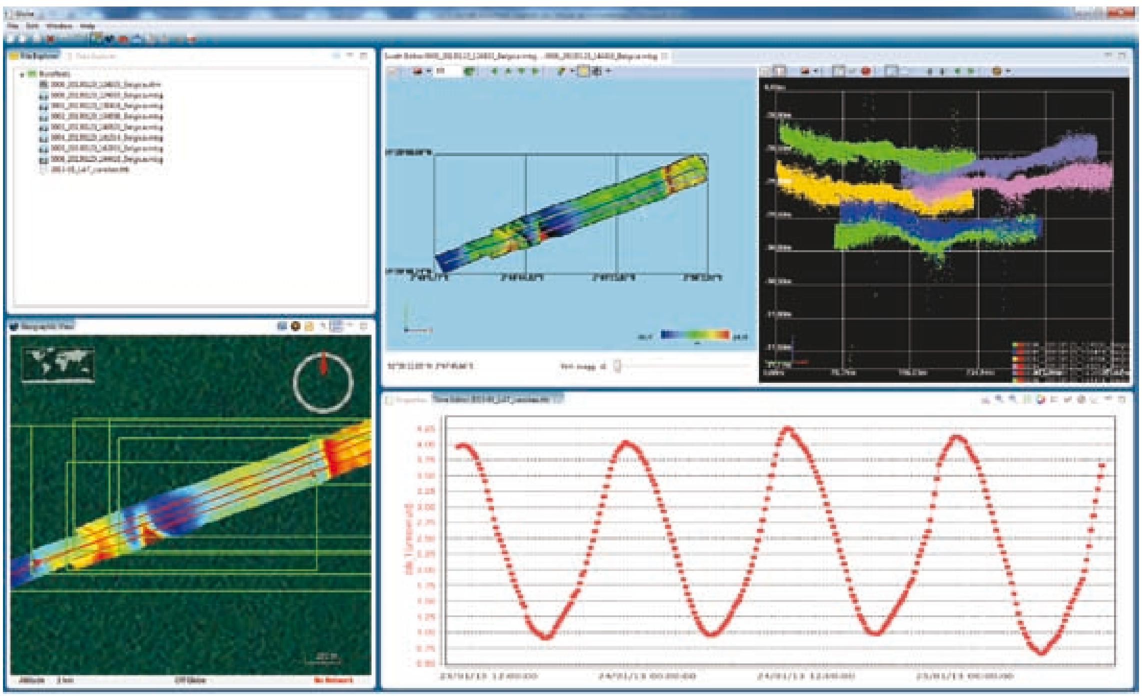


Fig. 1 - GLOBE software.

such as image and video. It can be used onboard during the survey to get a quick view of acquired data, or later, to re-process data with accurate environmental data.

Technically, the concept of the software relies on Eclipse RCP for the hosted client, Java and Nasa World Wind for the 3D views.

The user interface offers several perspectives to split several domains: swath editor, navigation editor, imagery editor, ...All views are dynamically relied

The version shown at IMDIS-2013 will present several key items :

- 3D visualization: DTM multi-layers from EMODNet, Water Column echogram, Seismic lines, ...
- Bathymetry Plug-In: manual and automatic data cleaning, integration of EMODNet methodology to introduce CDI concept, filtering, spline, data gridding, ...
- Photo/Video Plug-In
- Navigation 3D including tide correction, MRU corrections, GPS offsets correction,
- WMS/WFS interfaces.

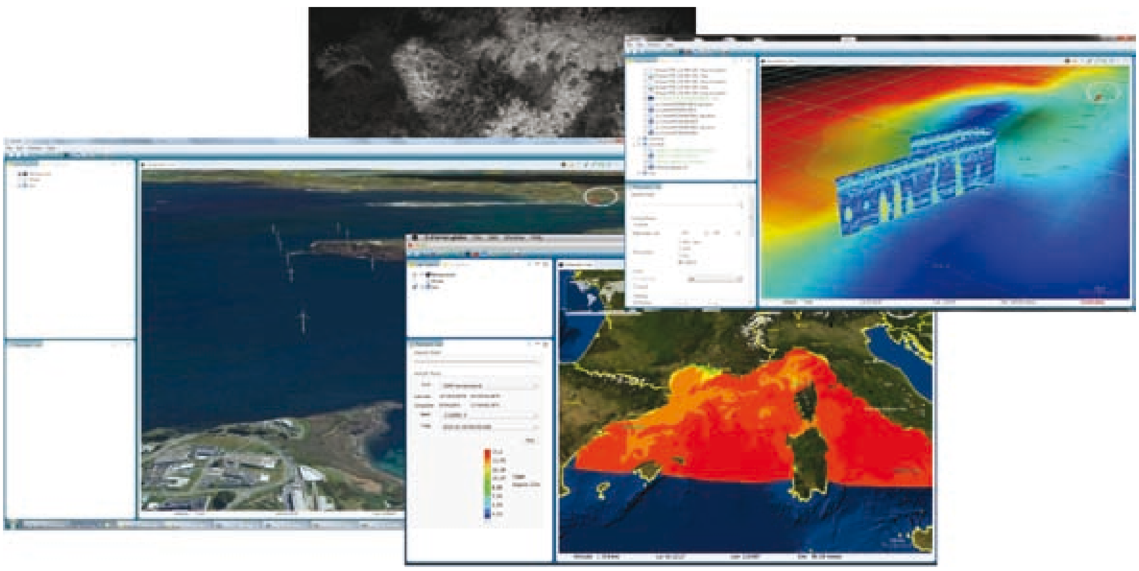


Fig. 2 - GLOBE software key items.

Oriented to data center and networking activities

A main objective of the EMODNet European project is to elaborate a common processing flow for gridding the bathymetry data and for generating harmonized digital terrain model (DTM) : this flow includes the definition of the DTM characteristics (geodetic parameters, grid spacing, interpolation and smoothing parameters…) and also the specifications of a set of layers which enrich the basic depth layer : statistical layers (sounding density, standard deviation,…) and an innovative data source layer which indicates the soundings origin and which links to the associated metadata.

GLOBE Software provides the required tools for applying this methodology and is offered to the project partners.

Sextant: the French Spatial Data Infrastructure for Marine Environment

Catherine Satra Le Bris, Ifremer, Catherine.Satra@ifremer.fr (France)
Erwann Quimbert, Ifremer, Erwann.Quimbert@ifremer.fr
Mickael Treguer, Ifremer, Mickael.Treguer@ifremer.fr
Abdelaziz Louarit, Ifremer, Abdelaziz.Louarit@ifremer.fr

Access to ocean and marine data is a major issue both for marine related researches and for marine environment management and spatial planning in application of Water Framework Directive (WFD) and Marine Strategy Framework Directive (MSFD).

In the framework of these initiatives, advances have been achieved both in the national and transnational cooperation: data policies, share of environmental data and also in the technical implementation of distributed spatial data systems (metadata management, portal accessing distributed repositories,…). These systems are now operationally robust and support the European Commission DGl-Mare actions such as the European Marine Observation and Data Network (Emodnet).

Ifremer has developed a spatial data infrastructure for marine environments, called Sextant. Sextant is a portal to manage, share and retrieve geographical marine information in order to support marine studies and decision making in fields such as biodiversity, marine renewable energy, integrated coastal zone management, fisheries, coastal and deep-sea environment, exploration and exploitation of the seabed.

Technologies used have been chosen to be compliant to Inspire directive both for discovery, data access and data visualization services. Standardization has been considered as a key point for the success of the integration of very large partnerships in data management

The main functions of Sextant are:

- the ‘Catalogue’ allows end-users to search one or more metadata according to different criteria (geographic extent, theme,…). To promote the share and distribution of datasets we have normalized the data description (ISO 19115), the XML encoding (ISO 19139) and access to catalogs (CSW). Moreover 95% of the metadata are Inspire compliant;
- the Web GIS interface allows end-users to create maps including several layers from internal and external sources. The portal access to local or remote OGC services displaying data by WMS;
- the download basket allows end-users to download the selected data in various formats and projections for use in GIS software;
- the possibility of creating a web portal for each thematic catalogue.

Some projects using Sextant

To promote the dissemination of marine environmental data sets, some European programs are using Sextant for its catalogue fonctionnality, such as Seadatanet, Emodnet or MyOcean.

We can also mention two Interreg projects, localized on the Channel marine area, wich are using Sextant as web GIS tool : CHARM and PANACHE.

The multidisciplinary integrated approach of the CHARM project (CHannel integrated Approach for marine Resource Management) between France and England offers decision makers a status report of the English Channel ecosystem and a range of tools based on scientific knowledge for the sustainable management of living marine resources. Through a personnalized Sextant webportal, the whole set of georeferenced data produced during the CHARM project is made accessible to managers, agencies and any stakeholders.

The main aim of the PANACHE project (Protected Area Network Across the Channel) is to develop a stronger and more coherent approach to the management, monitoring and involvement of stakeholders for marine protected areas in the Channel. This project is closely aligned with the ValMer project and together both projects will focus on better management, sustainable use and protection of the Channel marine area, using the same webGIS tool based on Sextant.

One Approach to Ensuring Long-term Access to Ocean Research Data

Cynthia Chandler, Woods Hole Oceanographic Institution, cchandler@whoi.edu (USA)
Robert Groman, Woods Hole Oceanographic Institution, rgroman@whoi.edu
Dicky Allison, Woods Hole Oceanographic Institution, dallison@whoi.edu
Peter Wiebe, Woods Hole Oceanographic Institution, pwiebe@whoi.edu
David Glover, Woods Hole Oceanographic Institution, dglover@whoi.edu

Oceanographic research requires access to data from an extensive range of disciplines. Data access, whether providing or attaining access, depending on one's role (e.g., data manager or oceanographic researcher) presents additional challenges that must be met by the research community. To meet these challenges the United States National Science Foundation (US NSF) is trying a different approach to management of ocean biogeochemistry research data. The Biological and Chemical Oceanography Data Management Office (BCO-DMO) was created in late 2006, and is a combination of the formerly independent data management offices for the United States Global Ocean Ecosystems Dynamics (US GLOBEC) and US Joint Global Ocean flux Study (US JGOFS) programs.

Efforts at BCO-DMO focus on comprehensive data management activities that span the full data life cycle from "proposal through preservation". BCO-DMO staff members work in partnership with NSF-funded investigators to ensure that data resulting from their respective research projects are archived at the appropriate US National Data Center. In addition to ensuring final archive of US NSF Division of Ocean Sciences funded research data, efforts undertaken by BCO-DMO data managers foster community building, establishment of trust and mutual respect between collaborative partners, and capacity building through outreach and education efforts. With support from BCO-DMO, NSF-funded ocean science researchers in the US have been contributing data from recently funded projects to the BCO-DMO data system, and it has evolved into a rich repository of data from ocean, coastal and Great Lakes research programs.

Providing comprehensive data management for ocean research data is a complex task. The BCO-DMO is just one node in an extensive network that is being funded to ensure long-term access to research data that are supported by sufficient metadata to enable accurate use, reuse and ideally, unanticipated use of those data. The BCO-DMO team has established strategic partnerships to enhance and extend our services. Some examples of these partnerships, identified

Fig. 1 - Data life cycle phases from the BCO-DMO perspective.

by the relevant life cycle phase (Fig. 1, Chandler et al., 2012) are listed below.

- Beginning with the proposal phase, BCO-DMO staff members provide recommendations to Principal Investigators (PIs) for creating the two-page data management plan that must accompany all US NSF proposals.
- Extensive coordination between the Rolling Deck to Repository (R2R) and BCO-DMO has affected substantial change during the data acquisition phase. This partnership enables reciprocal exchange of data and metadata thus improving not only efficiency but also the accuracy of curated content at both repositories and provides opportunities for connecting related information. Development of a shipboard event logging system increases the accuracy of event recording during research cruises with the final event log being reported as a dataset. Cooperation on controlled vocabularies describing sampling instrumentation and publication of content as Linked Open Data improves interoperability between the two repositories (Arko et al., 2013).
- During the data analysis and synthesis phase frequent communication with PIs provides opportunities to develop improved data structures and agreement on the answer to the ubiquitous question, ‘What constitutes a dataset?’. Such communication also results in more robust metadata, thus facilitating subsequent contribution and improving the utility of data for colleagues, especially those in different research disciplines.
- Frequent communication with US NSF program managers to align and confirm expectations, and active engagement of those program managers who follow up with PIs on data reporting requirements has yielded significant improvements during the data contribution phase. The ability to harvest metadata from ‘upstream’ complementary sources (e.g., trusted, authoritative repositories such as NSF and R2R that curate related information at earlier stages in the life cycle) enables BCO-DMO to work more efficiently with PIs to complete the catalog of data and information resulting from a research project.
- The discovery and access and data use and reuse phases are very closely related. Collaboration with computer scientists at Rensselaer Polytechnic Institute’s Tetherless World Constellation (RPI/TWC) resulted in the development of an ontology for oceanographic data. Mapping of local BCO-DMO terms to community-wide (e.g., SeaDataNet) terms available from the Natural Environment Research Council (NERC) Vocabulary Server (Leadbetter et al., 2013a) improves discoverability of the data and also improves interoperability with other data systems. The results of these two efforts are implemented in the semantically-enabled faceted search extension to the map-based data access system.
- In addition to serving the data (http://bco-dmo.org/data), a partnership with the MBLWHOI Library provides the ability to publish a dataset (data publication phase). Extracting data and metadata from the BCO-DMO system and assigning a Digital Object Identifier (DOI) renders that data object a ‘citeable reference’ acceptable to journal publishers (Leadbetter et al., 2013b).
- Data managed by BCO-DMO are submitted to a US National Data Center for permanent, long-term archive and the relevant accession number recorded in the BCO-DMO database. Data preservation in a permanent archive is the final phase in the data life cycle. Preserving the dataset and standards-compliant metadata in a long-term archive ensures that the data life cycle does not culminate in death, but rather creates the very real possibility of ‘data immortality’.

Acknowledgements. BCO-DMO is funded by the US NSF Division of Ocean Sciences and Office of Polar Programs. In addition to the authors, the BCO-DMO staff includes Nancy Copley, David Dubois, Steve Gegg, Danie Kinkade, Terry McKee, and Shannon Rauch. BCO-DMO acknowledges Robert Arko (Lamont-Doherty Earth Observatory, Columbia University) of the R2R project and Adam Leadbetter (British Oceanographic Data Centre) for making significant contributions to development and use of controlled vocabularies and progress made in using Linked Open Data. The ontology and semantically-enabled faceted search capability were developed in collaboration with Patrick West, Eric Rozell, Stephan Zednick and Peter Fox of RPI/TWC. The data publication and citation research project was supported by funds from Woods Hole Oceanographic Institution and the Jewett Foundation in the United States, the International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission (IOC) of UNESCO and the Scientific Committee on Oceanic Research (SCOR).

References

Arko, R., C.L. Chandler, P.D. Clark, A. Shepherd, C. Moore and S. Beaulieu (2013) *Rolling Deck to Repository (R2R): Collaborative Development of Linked Data for Oceanographic Research.* Geophysical Research Abstracts 15: EGU2013-9564

Chandler, C.L.; R.C. Groman; M.D. Allison; P.H. Wiebe; D.M. Glover (2012) *BCO-DMO: Improving Access to Ocean Research Data throughout the Data Life Cycle,* Poster IN43B-1506 presented at 2012 Fall Meeting, AGU, San Francisco, CA, 6 December 2012. http://www.bco-dmo.org/2012-fall-agu-meeting

Leadbetter, A., R. Lowry and O. Clements (2013a) *Putting meaning into NETMAR - the Open Service Network for Marine Environmental Data.* International Journal of Digital Earth. DOI:10.1080/17538947.2013.781243

Leadbetter, A., Raymond, L., Chandler, C., Pikula, L., Pissierssens, P., Urban, E. (2013b) *The Ocean Data Publication Cookbook,* UNESCO/IOC Manuals and Guides No. 64. v1, March 2013, http://www.iode.org/mg64

Coastviewer: a tool to enable the visualization of marine and coastal data

Giorgio Santinelli, Deltares, giorgio.santinelli@deltares.nl (The Netherlands)
Fedor Baart, Deltares, fedor.baart@deltares.nl (The Netherlands)
Gemma Ramaekers, Rijkswaterstaat, gemma.ramaekers@rws.nl (The Netherlands)
Erik-Jan Vos, Nelen & Schuurmans, erikjan.vos@nelen-schuurmans.nl (The Netherlands)

Introduction

In coastal and marine environments the availability of simple, interoperable and efficient access to scientific data is of primary importance. Community managers and researchers need information at different levels of detail and with different goals. Building a single standard-based and open source toolbox that allows open data access and collection and fulfill the needs of a diverse range of people is possible and has been proven successful with the *OpenEarth* initiative. As final step of this process, a real-time web based platform for visualization that includes time varying data has been provided in order to help and encourage the dissemination of open and high-quality data (Fig. 1).

Fig. 1 - Graphical User Interface of the Coastviewer. Map of 2d bathymetry and cross-shore transects along the Dutch coast. Image credit: Google Earth^{TM} mapping services.

Data standards and Web Services

Dealing with scientific and geospatial data has grown from handcrafted file formats, exchanged through ftp servers for small communities to a web of open data sources that provide data in easy consumable services. International groups such as the OGC (*Open Geospatial Consortium*) and many global scientific organizations (Unidata, OpenDAP) have created the standards and software that made this transition possible. Especially 3D and time enhanced maps, curvilinear and unstructured grids are not yet properly supported by these protocols. Other challenges include real-time generated visualizations (big datasets, running models), and taking into account other extra dimensions (wave

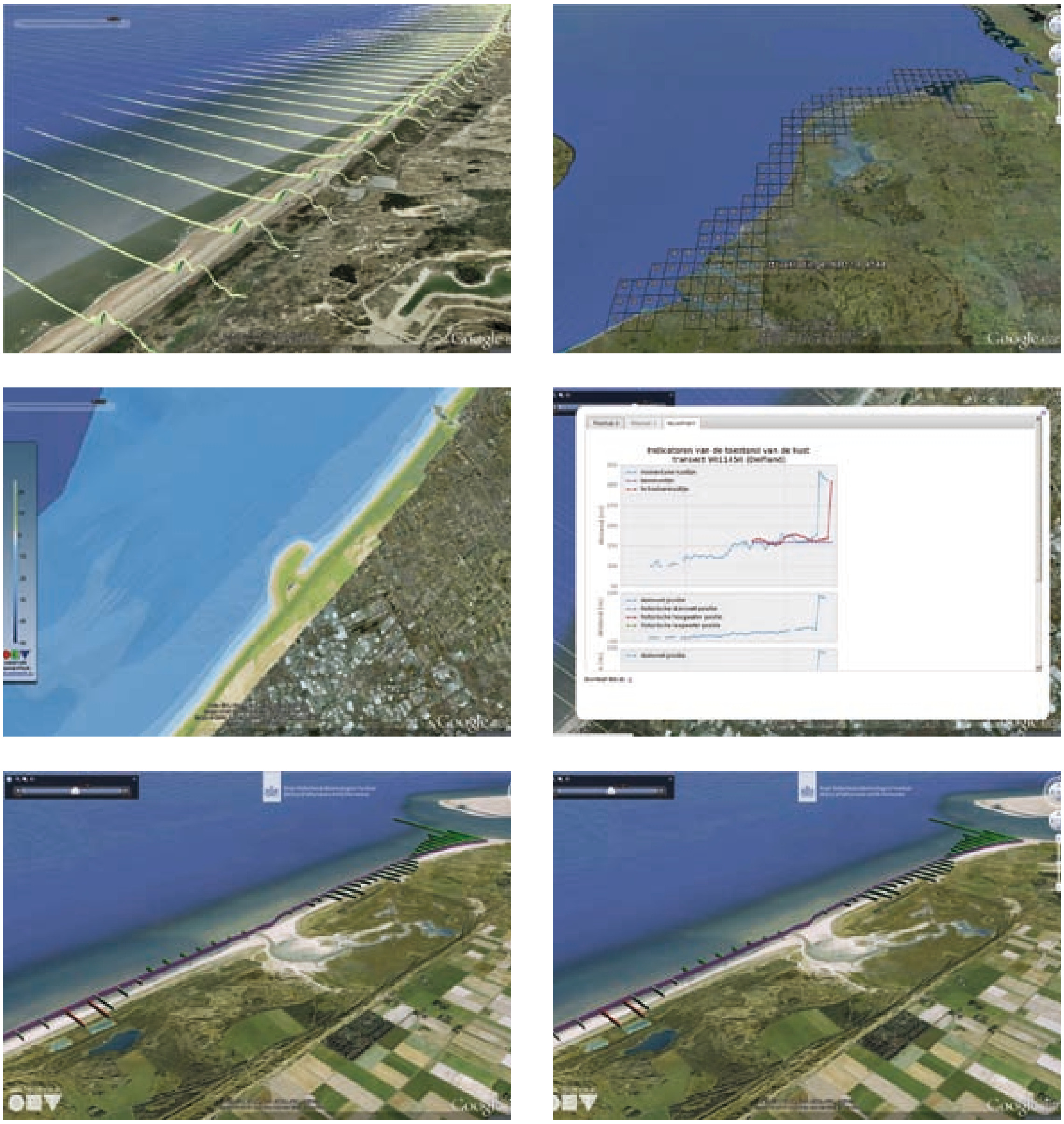


Fig. 2 - Screenshots of the Coastviewer. Image credit: Google Earth™ mapping services.

directions, spectral frequencies). The KML (*Keyhole Markup Language*) is used in the example below. Although KML language was developed for use with *Google Earth*, from 2008 it is an open standard for all geobrowsers, therefore can be used in viewers different than *Google Earth*.

The Coastviewer

The Coastviewer leverages KML and represents a complete solution for information services, being a web based platform that enables the user to publish visualizations of measurements, coastal indicators and model results for different time intervals and different areas. The concept behind the viewer is to support the decision making of coastal managers and the building of expert advice by scientists and engineers. The source of information is standard-based data, accessible via an OPeNDAP server and queried from the viewer using a web interface. Making use of the same idea behind Real-Time Control systems, the viewer is kept up-to-date by a direct link to a database server, which serves numerous types of datasets. The simultaneous visualization of different available data is enabled by the KML protocol which allows to present all kind of information, e.g. long term, large scale morphological changes due to sea level rise and large scale interventions and short term evolution of coastal indicators, safety levels and sediment distributions. Many marine and coastal datasets have already been implemented in the viewer, such as maps of grainsize distribution and geo-referenced sample properties, transect timestacks along the Dutch coast, map timestacks of multibeam eco-soundings of bathymetry and Lidar topography, water quality measurements, coastline indicators and sand nourishments carried out along the coast and offshore. Specific datasets have been augmented so the user can interact with the viewer, and modify the rendering by raising the elevation of 3D KML objects, changing the scale, choosing a colormap, setting the transparency, and save plots and download the visualized data in various formats (Fig. 2).

Acknowledgements. The Coastviewer and database is developed as part of Deltares project KPP-B&OKust (Knowledge for coastal management) commissioned by the Dutch Ministry of Infrastructure and the Environment.

References

Google Earth, http://earth.google.com;

KML, http://www.opengeospatial.org/standards/kml;

Coastviewer, http://test.kustviewer.lizard.net/kml/viewer/;

OGC, http://www.opengeospatial.org/standards;

OpenEarthTools, http://openearth.eu;

The MyOcean Product Database Management

Laurence Crosnier, Mercator Ocean, lcrosnier@mercator-ocean.fr (France)
Joël Dorandeu, Joel.Dorandeu@mercator-ocean.fr (France)
Remi De Dianous CLS, rde-dianous@cls.fr (France)
Vincent Claverie, CLS, vclaverie@cls.fr (France)
Thomas Loubrieu, Ifremer, thomas.loubrieu@ifremer.fr (France)

As a marine core service, the MyOcean (www.myocean.eu) objective is to provide to users the best generic information available on the state of the ocean. MyOcean (2009-2012) and now MyOcean2 (2012-2014) are committed to develop and run a European service based on a worldwide capacity for ocean monitoring and forecasting, using observations data, modelling and assimilation systems. The MyOcean service consists in providing “core” information on the ocean and can be defined as i) An integrated Service; ii) An Open and Free service; iii) Providing access to a single Catalogue of products (Fig. 1) ; iv) A reliable service; v) A sustainable service. We will describe the methodology, human processes and tools used within the MyOcean Information System in order to reach the above objectives. The MyOcean Network of product managers feeding the Product Database and gathering metadata for the 107 MyOcean products will be described. Moreover, MyOcean is now able to assess the impact of the upstream observing network on all the MyOcean products. This will be described along with the tools allowing monitoring such upstream dependencies.

The MyOcean service

The pre-operational marine Environment Monitoring service of COPERNICUS is currently provided through the EU-funded MyOcean2 project and delivers regular and systematic reference information on the state of the oceans and regional seas. The MyOcean Service www.myocean.eu addresses four main domains:

Fig. 1 - The MyOcean catalogue search engine www.myocean.eu.

- Marine safety (e.g. marine operations, oil spill combat, ship routing, defense, search & rescue...)
- Marine resources (e.g. fish stock management...)
- Marine and coastal environment (e.g. water quality, pollution, coastal activities...)
- Climate, seasonal and weather forecasting (e.g. climate monitoring, ice survey, seasonal forecasting...)

After data acquisition from the ground segment of the space-based (with processing levels L1 and L2) and in situ networks as well as acquisition of atmospheric forcing data (winds, temperatures, fluxes) from National Meteorological Services, MyOcean processes these data into quality-controlled datasets (with processing levels L3 and L4) at Thematic Data Assembly centers (sea surface temperature, ocean colour, sea level, sea ice, winds and in-situ data) and runs numerical ocean models in near real time assimilating the above thematic data to generate analyses and forecasts for the Arctic Ocean, Baltic Sea, Atlantic European North West shelf Ocean, Atlantic Iberian Biscay Irish Ocean, Mediterranean Sea, Black Sea and Global ocean. Reprocessing and reanalysis are also performed.

MyOcean Information System and MyOcean Web Portal

MyOcean products are available through the Interactive Catalogue provided on the www.myocean.eu web portal, allowing in particular downloading products in a "one stop shop" way. The MyOcean catalogue of products displays the following parameters: Ocean Temperature, Ocean Salinity, Ocean Currents, Sea Ice, Sea Level, Ocean Winds, Ocean Optics, Ocean Chemistry, Ocean Biology and Ocean Chlorophyll. There are 107 products delivered by MyOcean as of April 2013, 31 products are coming from Ocean Numerical Models and are sorted by geographical area (Global Ocean, Arctic Ocean, Baltic Sea, European-North-West Shelves, European South Shelves, Mediterranean Sea, Black Sea), and 76 products are coming from satellites or in-situ based observations and are sorted by parameter: sea surface temperature, ocean colour, sea level, sea ice, winds and in-situ data. MyOcean catalogue includes Real Time products, Forecast products as well as Multi Year Time Series products (i.e. Reanalysis, Hindcast or Reprocessing).

MyOcean Information System gathers the MyOcean web portal, the product database and the product download mechanisms infrastructure. It is managed centrally for some of its components but also locally for some other components. Indeed, the file data is stored locally in various locations spread all over Europe called "dissemination units". There are 25 MyOcean "disseminations units" distributing the products produced by the 44 MyOcean so-called "production units". The same set of MyOcean Download mechanism tools are installed within the 25 MyOcean "disseminations units" hence allowing MyOcean to deliver its data in a homogenous way from a user point of view, even though the file data is physically stored into different locations.

MyOcean has set up a data flow monitoring system which is able to monitor all the interfaces used within MyOcean in order to acquire the ground segment of the space-based data, the in situ networks data as well as the atmospheric forcing data (winds, temperatures, fluxes) from National Meteorological Services. Moreover, the internal interfaces used within MyOcean are also monitored. Hence, MyOcean is now able to assess the impact of the upstream observing network on all the MyOcean products.

The MyOcean Product Database management

The tools allowing to manage the catalogue of products are developped and maintained by the MyOcean Information System. Metadata information about products feeding those tools are managed by a central product manager as well a network of about 35 local product managers disseminated all over Europe. This network provides with up to date and consistent information about products and theirs characteristics, providing users with accurate product information and documentations.

Implementation of a live access server for the management of wave energy data in the Eastern Mediterranean

Stavros Stylianou, Oceanography Centre, University of Cyprus, stylians@ucy.ac.cy (Cyprus)
George Zodiatis, Oceanography Centre, University of Cyprus, gzodiac@ucy.ac.cy
George Galanis, Department of Physics, University of Athens, ggalanis@mg.uoa.gr
Andreas Nikolaidis, Oceanography Centre, University of Cyprus, nikolaidis.andreas@ucy.ac.cy

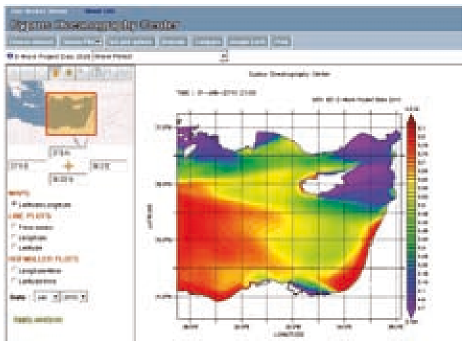


Fig. 1 - Plot showing the Wave Period of January 2010.

The Cyprus Oceanography Centre (OC_UCY) has implemented a Live Access Server (LAS) for the management, visualization and efficient distribution of wave energy data in the Eastern Mediterranean. The LAS server has been developed at NOAA/Oceanic and Atmospheric Research's (OAR) Pacific Marine Environmental Laboratory (PMEL) and consists of a software packages variety. For the successful implementation of LAS at the OC-UCY the Tomcat and Apache Servers, the Thredds Data Server and Ferret was installed. The LAS server provides the ability to the users to navigate through a structured list of options organized hierarchically with explanatory titles, to visualize data with on-the-fly graphics, request and export the spatial and temporal portion of the dataset they need in other software compliant forms. The visualization of the data is accomplished with Ferret which is an interactive visualization and analysis environment developed by the PMEL

Through the Live Access Server the Cyprus Oceanography Centre provides access to the E-wave project database. This database includes information and statistical analysis regarding the energy potential and climatological characteristics of sea waves (such as Significant Wave Height, Wave Frequency, Wave Period, Wave Energy Potential, etc.) for a period of time between 2001 and 2010 in the wider Eastern Mediterranean with 3 hourly and monthly interval. Users can access freely wave energy data individually per month (as maps, Fig. 1) or aggregated over time (line plots, Fig. 2). The use of the time series function in OC-LAS provides a continuous set of data covering each year. The use of the OC-LAS server enables those interested in the potential of wave energy and the climatology of the waves to quickly visualize, collect subsets and derived products on the fly regarding wave energy data in the Eastern Mediterranean, Levantine Basin.

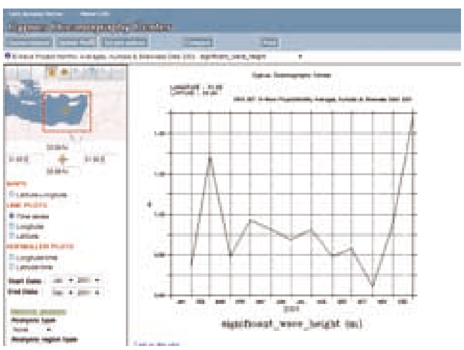


Fig. 2 - Time series plot of the significant wave height (m) from year 2001.

A Proactive System for Oil Spills and Marine Environment Monitoring

Marco Tampucci, Institute of Information Science and Technology, Marco.Tampucci@isti.cnr.it (ITALY)
Massimo Martinelli, Institute of Information Science and Technology, Massimo.Martinelli@isti.cnr.it
Davide Moroni, Institute of Information Science and Technology, Davide.Moroni@isti.cnr.it
Gabriele Pieri, Institute of Information Science and Technology, Gabriele.Pieri@isti.cnr.it
Ovidio Salvetti, Institute of Information Science and Technology, Ovidio.Salvetti@isti.cnr.it
Paolo Villa, Institute of Information Science and Technology, Paolo.Villa@isti.cnr.it

Oil pollutions impact the environment, the economy and the quality of life for coastal inhabitants. The increasing importance of petroleum products and the list of oil spill events in last two decades raised the concern on maritime safety and environmental protection. Indeed, annually, on average 1,300,000 t of oil entered the oceans during the 1990s with tanker vessel spills accounting for 100,000 t, run-off 140,000 t, and pipeline leaks just 12,000 t. Aside from natural seeps, which contribute an estimated 600,000 t or ~45% of total emissions, other important sources include vessel operational discharges.

For these reasons, there has been an increasing interest in frameworks for remotely detect oil spill at sea and several technological advances were made, especially under the propulsion of catastrophic events, like the Deep Water Horizon spill, during which several remote sensing technologies moved up to the technological readiness scale.

Current limits

Nevertheless, most of the approaches have been focused on the detection of large oil spills, demanding for the intervention of ad hoc task forces, while smaller oil spills and operational discharges in regional area for the routine work of local authorities have received somewhat less consideration, despite their importance especially in protected areas of great environmental value such as marine parks.

In addition, usual remote sensing systems can be enriched by adding information collected in situ thanks to static and mobile sensors. Only properly devising algorithms and methods for correlating and fusing both in situ and remote sensed information, it is possible to fully deploy the potential of integrated systems for sea monitoring.

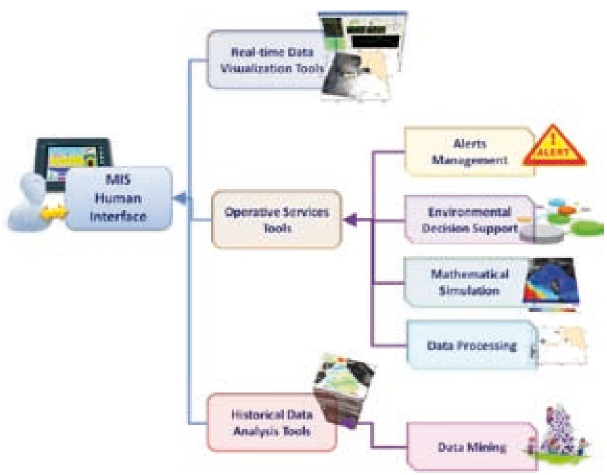


Fig. 1.

Contribution of this work

In this work, we aim at addressing these issues by proposing an integrated and interoperable system able to monitor ship traffic and marine operators. The system uses advanced sensing capabilities from a variety of electronic sensors, along with geo-positioning tools, and through a communication infrastructure, yet it is suitable for local authorities and stakeholders. Our model is capable of transferring data, freely and seamlessly, between different elements of the information systems (and their users). In this way different data are brought together, easily and in a consistent and usable form, in order to facilitate dynamic links between different models and analytical processes.

In particular, the overall information stream is managed by a modular framework named Marine Information System (MIS), based on PostgreSQL-GeoNetwork-Openlayers architecture. The MIS integrates multispectral aerial data, SAR satellite processed data, environmental data from in situ monitoring stations (e.g. buoys), dynamic data acquired from in situ mobile sources (e.g. volunteers, Autonomous Underwater Vehicles ...). The MIS is enriched with a collection of environmental decision support services, for i) automatically screening the overall situation, ii) quantitatively representing risk factors and iii) proactively notifying events that deserve the consideration of end users. A model for the computation of dynamic risk maps has been included, by aggregating the available heterogeneous data source ranging from maritime traffic density to water quality parameters sensed by electronic noses. Visualization of the risk map provides a quick yet effective way to have an outlook of the situation in the monitored area, while its automatic analysis – performed by intelligent agents – allows the delivery of proactive alerts to local authorities in charge of monitoring.

Demonstration activities and future developments

The proposed system has been demonstrated during extensive test exercises held at the National Marine Park of Zakynthos and at National Park of Tuscany Archipelago in the framework of FP7 Project Argomarine (FP7-234096). The modularity and flexibility of MIS framework makes it a suitable candidate not only for oil spills monitoring systems, but also for more general geospatial data management for marine applications, and its approach is now contributing as one of the starting references for the implementation of the interoperable data infrastructure of RITMARE project, promoted an funded by Italian Ministry of University and Research, coordinated by CNR and running from 2012 for 4 years.

Marine Social and Economic Data – UK Perspective

Leoni Adams, Marine Biological Association, leoams@mba.ac.uk (United Kingdom)
Dan Lear, Marine Biological Association, dble@mba.ac.uk
Becky Seeley, Marine Biological Association, bese@mba.ac.uk
Charly Griffiths, Marine Biological Association, chgr@mba.ac.uk

Marine social and economic data management in the UK - the current situation

The current level of use of marine social and economic data is low. The first ever dedicated chapter on the productive use of the sea in the UK was seen in Charting Progress 2 (2010), a review of the progress of the UK's vision of clean, healthy, safe, productive and biologically diverse oceans and seas. Chapter 5 – Productive Seas, included analysis of all use and value resulting from the marine environment. Whilst the collation of this information has been a forward step, it is still summarised with a need for better centralised data management ;

> *'We need better centralised collation of data on the distribution of pressures associated with activities such as aquaculture, leisure and recreation, and shipping.'*

Initiatives have been set up in the UK to assess the value of the marine environment. Both the Valuing Nature Network and the UK National Ecosystem Assessment have further cited a lack of data management in a marine social and economic context, and are pushing to increase the use of ecosystem goods and services approaches within impact assessments, even furthering the requirement for good practice in this area.

The importance of marine social and economic data

Following the introduction of the Marine and Coastal Access Act (2009) and the requirements under the Marine Strategy Framework Directive, the UK Government and devolved administrations established the Marine Management Organisation (MMO) and Marine Scotland. The purpose of the MMO and Marine Scotland is to promote the UK Government's goals for for clean, healthy, safe, productive and biologically diverse oceans and seas. To achieve these visions, the MMO and Marine Scotland must provide an evidence base to support marine planning and to help select a number of Marine Conservation Zones.

Marine data management in the UK is co-ordinated through MEDIN (Marine Environmental Data and Information Network). MEDIN supports a data network for the marine environment, through a number of of recognised, accredited Data Archiving Centres (DACs). Each DAC deals with specific data themes;

- Archaeology Data Service, for marine historical data;
- British Geological Survey, for geological data;

- British Oceanographic Data Centre, for oceanograhic data;
- DASSH, for species and habitats data;
- Marine Scotland and Cefas and AFBI, for fisheries data;
- The Met Office, for meteorological data;
- The UK Hydrographic Office, for hydographic data.

DAC's provide a centralised store for UK data, and also exist to assist UK marine data providers with guidance and assistance in good data management practices. Good practices have been established for the majority of marine data wihtin the UK. However, despite the need for good social and economic evidence, there has historically been a limited focus on this type of data, and no currect DAC exists.

In order to assess the need for social and economic marine data management, and to begin to formulate a strategy, in 2012 the MMO commissioned a marine social and economic data and tools review, co-ordinated by MEDIN. This review reported on existing datasets, key data providers, existing data strategies and reviewed available tools for incorporating social and economic data. The review of social and economic data presented a number of weaknesses in current data management practices. Data access, and ability to create useful, standardised discovery metadata was impeded by a number of issues. Data were held by a large number of disparate data holders and with changes in marine management, much of these data had recently changed hands, making it difficult to track down the new owners and the new location of data online. Metadata records were generally poor with few meeting full MEDIN compliance. In particular, there were often poor spatial and temporal records, and little information on the protocols and standards used to collect and analyse the data and consequently poor understanding of the quality of the data.

Marine social and economic data management – the future

These issues need to be addressed for effective marine planning and MCZ designation. DASSH are curretly working as part of a large consortia on the international, VALMER project, a marine valuation project which will look at site specific impacts, scenarios and visions in the UK and France. This project creates a useful platform from which, as part of the MEDIN UK data management framework, DASSH are able to pro-actively begin the implementation and testing of marine social and economic data management strategies, and to integrate the findings with the national marine data management strategy, in order to improve overall management practices UK-wide.

Guidelines for marine social and economic data are currently emerging. We have the opportunity to move forward and implement new data management strategies, with DASSH and MEDIN placed in a leading role. It is important that this is done correctly in order to be effective, with full engagement of key partners and close liaison with the terrestrial sector.

POSTERS

Analytical tools for ESIMO

Evgeny Vyazilov, RIHMI-WDC, vjaz@meteo.ru, (Russia)
Denis Melnikov, RIHMI-WDC, melnikov@meteo.ru, (Russia)
Nikita Chunyaev, RIHMI-WDC, chunyaev@meteo.ru, (Russia)

It is not the lack of information that poses current problems in information technologies but difficulties in data processing and analyzing. The applications created are mainly aimed at visualization of original data or at their wide-scale statistical processing. Standard reports that are widely used in different computer-aided systems fail to provide interactivity. To make decisions, users need to obtain prompt responses based on distributed and heterogeneous data held in the Unified State System of Information on the Global Ocean (ESIMO).

Nowadays, large volumes of data became accessible in the on-line mode. It is now possible to use different data processing techniques (determination of critical values of environmental parameters and deviations from climatic normal's, rapid evaluation of hydrometeorological conditions, decision support, etc.). The capability of identifying useful information in a large volume of data accumulated is a critical goal of analytical systems. To make decisions more effective, it is essential that decision-makers can have information at any time, in the format and volume required, and in the form correctly presented. To that end, data need to be made accessible for decision-makers and be presented as maps, diagrams, graphs, tables, texts and interactive schemes of information display on the instrument dial displayed on the screen. A decision-maker cannot constantly sit at a computer and expect for the critical values of marine environmental parameters to be exceeded, he/she needs to inform of these by using software-agent, working on the decision-maker computer. Observed, diagnostic, prognostic and climatic data are to be analyzed.

To analyze data, different eventful (tools of hardware and software operation monitoring), interactive (DataManing, OLAP) and regulatory (reports, databases, analytical and forecasting means) tools are widely used. The most advanced is innovative analytical system QlikView by QlikTech Company; it provides users with tools for independent research and visual data presentation. Content-addressable retrieval tools are complemented with comparative analysis means that allow vivid comparison between the indicators chosen by a user.

The analytical component is developed in ESIMO to computerize data analysis (http://ak.esimo.ru). An interactive simple-interface tool and integrated data handling means are created. Analysts do not use complicated SQL-queries; they add new information resources, if necessary, and change analytical representation with no help from IT-experts. A map is included in the analytical representation by using means of indicating the URL address of internal or external open cartographic (Web Map Service standard). ESIMO information resources allow a prompt preparation of representation based on new types of data. An analyst can transform the data representation form, create map animation, prepare a report for decision-makers, etc.

More than twenty applications are identified for the development of analytical representations; each application may have a variety of analytical representations. Each representation includes from two (table, graph or map, or text) to several tens of representation elements (a variety of graphs, tables, geographic layers). Representation elements are derived by tentative data filtration from one or several key attributes. For example, graphs of vertical distribution of temperature, salinity and hydrochemical parameters can be constructed by data filtration from country, ship, cruise numbers, oceanographic stations and parameters. A variety of thematically similar representations are arranged as applications, e.g. “Dynamics of Marine Activity”, “Observation Systems Monitoring”, “Storm Analysis”, and “Climatic Characteristics of Russian Seas”. The representations created are also put on the ESIMO portal (http://portal.esimo.ru) in the form of an application and representation catalogue.

A software agent is developed to inform of and provide real-time information on natural hazards. This programme is installed in the decision-maker’s computer, becomes active in specific situations and provides minimum of the information required. When a natural hazard emerges, the agent informs a decision-maker of it, provides information on the current situation and makes it possible to obtain information on the hazard in more detail through ESIMO portal. The agent is able to be adjusted to a specific object – port, drill tower or region and to critical values of hazard indicators that are dangerous for the object selected.

Future development of ESIMO analytical capabilities is creation of a scheme designer and development of a decision support system.

A current synoptic picture scheme used in hydrometeorology can be taken as the basis of the scheme designer. Each indicator is to have its own instrument (thermometer, barometer-aneroid, weather vane, hygrometer, velocity sensor, depth gage, wave-height gage). The analyst is to be able to adjust the appropriate instruments and critical values of parameters and after that the scheme is to depict weather conditions as a light signal in the instrument (green, yellow, red, crimson).

In making decision, a decision-maker should realize possible consequences from a natural hazard and rely on recommendations provided by a decision support system. Information on possible impacts and recommendations rest on the knowledge base created from the previous experience. In this case it is not only critical values for each object that are taken into account but also the type of information (observations, forecast, climate after the event), decision-making level, season, and climatic zone where the object is located.

The decisions proposed rest on the ESIMO information resources represented as the integrated database and containing marine environmental monitoring data. The data are automatically updated on a periodic predetermined base from several minutes to a week depending on the observation frequency.

Storm activity over North Atlantic with ESIMO analytical tools

Natalia Viazilova, RIHMI-WDC,nav@meteo.ru, (Russia)

The goal of study is to show the indices of the extratropical storm activity over North Atlantic, that are available on the Unified State System of Information on the Global Ocean (ESIMO) portal [http://portal.esimo.ru]. The storm activity indices calculated based on automated cyclone detection/tracking algorithm and the 6-hourly SLP from the NCEP/NCAR reanalyses from 1948 to present. The list of indices includes cyclone tracks, cyclone frequency, the cyclone intensity. The cyclone frequency and the cyclone intensity are calculated by counting the cyclone occurrences and sum of cyclone centers MSLP anomalies at each grid point over North Atlantic during the month. The study of the storm activity changing over time is avilable from storm activity integral indices for some selected region of North Atlantic.

The maps with patterns of cyclone activity for last calendar month and graphics with storm activity changing over time are visualisated based on the analytical component ESIMO [http://ak.esimo.ru]. Application with storm activity indices have some of analytical representations: “North Atlantic storm track”, “Maximum storm wind”, “Integral month storm indices for Baltic and Barenz seas” and other. Each representation includes several elements (maps and graphics).

The representation: “North Atlantic storm track” includes map with storm tracks over region and graphics of cyclone frequency distribution for cyclones with different intensity (weak strong, moderately strong, extreme cyclones) for the last month. The representation: “Maximum storm wind” includes map with pattern of maximum wind near cyclone centres and graphics of distribution for cyclones with different maximum wind intensity (weak strong, moderately strong, extreme wind). The representation “Integral month storm indices for Baltic and Barenz seas” includes graphics of region mean cyclone frequency and the cyclone intensity from 1948 to present and climatic value indices for every calendar month. The representation “Integral season storm indices for North Atlantic” includes graphics of region mean cyclone frequency and the cyclone intensity changing in midlatitude [30°N-80°N, 55°W-70°E] and high- latitude zone [55°N-80°N, 50°W-70°E].

Analyses shows, that in high-latitude North Atlantic extratropical cyclone frequency for moderately strong cyclones (cyclone with an MSLP more 970 hPa) significantly decreased from 1948 to early 1990. At that time, strong-cyclone frequency (cyclones with an MSLP of 970 hPa or less) significantly increased, especially last two decada. The most notable historical changes in cyclone activity in high-latitude are found to be associated with strong-cyclone. In midlatitude North Atlantic cyclone activity in main is associated with moderately cyclones. Baltic region is outstanding by most increasing of cyclone activity in North Atlantic.

A numerical based DSS as tool for local planning of Oil spill emergencies

Matteo Sinerchia, IAMC CNR Oristano, matteo.sinerchia@cnr.it
Andrea Cucco, IAMC CNR Oristano, andrea.cucco@cnr.it
Leopoldo Fazioli, IAMC CNR Oristano, leopoldo.fazioli@cnr.it
Antonio Olita, IAMC CNR Oristano, antonio.olita@cnr.it
Alberto Ribotti, IAMC CNR Oristano, alberto.ribotti@cnr.it
Roberto Sorgente, IAMC CNR Oristano, roberto.sorgente@cnr.it

A standalone DSS (Decision Support System) , part of an integrative operational forecasting system of the marine circulation, as support for the local Coast Guard in the Bonifacio Strait (Western Mediterranean) for the management of maritime operations during oil spill emergencies. The area has an extraordinary environmental value but also extreme meteorological conditions due to a particular orography. Since its maritime status of International Strait recognised by International Maritime Organization as the first Particularly Sensitive Sea Area in the Mediterranean, every year it is crossed by about 3500 vessels as oil, gas and chemical tankers (Sorgente et al., 2012). In case of accident the quickness of action is essential to avoid or mitigate an environmental and economic impact.

The system can facilitate the planning of emergency operations of the local authorities through the forecast of the physical-chemical processes determining the fate of the oil at sea. The system is composed by chain of nested operational models, from the basin to the coastal scale, whose outputs are used by the Java-based Graphical User Interface (GUI) to run the system. The system was qualitatively validated in January 2011 when a heavy oil spill occurred in a harbour in north Sardinia, then moving along the coast for 10 days.

The integrated system

The integrated system aims at providing a prognostic tool for managing the oil-spill emergencies giving an estimate of the chemical and physical evolution, both in space and in time, of a possible oil spill in the study area up to a maximum of 72 hours thus giving the opportunity to know, well in advance, its movement (Oilta et al., 2012). It draws the area of probability of the source of pollution when the event occurred a few days before through a backward investigation (Cucco et al. 2012; Cucco et al., 2012), then supporting in the identification of those vessels that illegally discharge their waste water into the sea. The integrated Oil-Spill forecast coastal system is named Bonifacio Oil-spill Operational Model (BOOM) and is composed by a hierarchy of different types of nested hydrodynamic numerical models, based on structured and unstructured grids, and a GUI to facilitate the set-up and analysis of simulations of transport and dispersion of the oil-spill. The core of BOOM is composed by a set of finite element numerical models, including a three dimensional coupled hydrodynamic-wave model (the SHYFEM3D-WWM, simply SHYFEM hereafter) with Lagrangian trajectory and weathering modules (named FEMOIL).

Such tools are used to operationally provide the hydrodynamics at coastal scale and the fundamental variables controlling the oil-spill fate and to forecast its dispersion in the Bonifacio coastal area. It can reproduce and predict the space-time variability of the water circulation and sea conditions in the coastal area with a spatial resolution less than 0.1 km and a number of vertical levels such to properly resolve the surface layer, whose dynamics are the main mechanism of transport and dispersion of hydrocarbons.

The coastal system provides two daily forecasts, to maximum 3 days each, of the barotropic current velocities and waves in the Bonifacio Strait area (including the La Maddalena Archipelago) at steps of 6 hours. Numerical simulations are carried out over a computational domain (009°6.899' - 009°37.06' long. E and 41°6.2' - 41°25.683' lat. N) by means of a finite element staggered grid that is characterized by different spatial resolutions ranging from 10 m, in the La Maddalena Archipelago, to few kilometres in the off-shore areas.

The GUI

The GUI retrieves predictions generated by an ocean forecasting system, the coastal part of which is the BOOM.

The GUI is a user-friendly graphical interface used to create scenarios, to start numerical model simulations and analyse the obtained results. Usually the use of simulations generated by complex numerical models requires the interaction with computer codes and high-level processes. For this reason it is limited to users able to move through blocks of commands in machine codes. The GUI is designed to put a user, even with modest computer skills, able to easily interact with the numerical model.

In order to make operational the forecasting system it was therefore necessary to let it accessible to all users, even if not expert programmers, and facilitate the process of selection of scenarios, launch of the simulation and analysis of the generated results. The GUI has been designated through the "object-oriented" approach and implemented in Java language. These choices have two advantages: the first is the modularity of the system (easier to upgrade or modify the interface with new functions) and independence from the operating system (Windows, Unix, Linux, Sun, etc.) on the PC where the interface is installed.

The GUI can be used in two ways: in forecast mode that allows the user to obtain the scenario on the characteristics and position of the slick from the point and time of discharge to the following 72 hours; in backward mode allows to trace back to the estimate of the start date and of the possible area of an oil spill in the sea, known type and amount of oil pollution identified and the affected area.

Use of GUI

In the morning of the 10th of January 2011 in the industrial harbour of Porto Torres, north-western Sardinia,the oil spilled out from a ship tanker and then continuously for the following 18 hours. About 50 m3 of heavy crude oil were released into the sea then moving north-eastward along the coast for about 5 days. After several days the BOOM system was run in foreward mode in order to verify the dispersion of the oil along 100 km of coast as the slick found during a survey by an airplane at the entrance of the Bonifacio Strait one week after the accident.

The GUI reproduced fairly well the probable trajectory followed by the spilled oil during the whole period then revealing that the origin of the spilled oil was from Porto Torres as the slick found at the entrance of the Bonifacio Strait. Samples of floating oil were successively collected and chemically analyzed certifying its origin.

The coast guard in La Maddalena has included the results of some simulations in its local anti-pollution plan for 2009, submitted to the Italian Ministry for Environment. The local anti-pollution plan was completely revised, as it did not previously consider in detail the local circulation inside the La Maddalena archipelago.

References

Cucco A., Sinerchia M., Ribotti A., Olita A., Fazioli L., Sorgente B., Perilli A., Borghini M., Schroeder K. Sorgente R., *A high resolution real time forecasting system for predicting the fate of oil spills in the Strait of Bonifacio (western Mediterranean),* (2012), Marine Pollution Bulletin, 64, 6, 1186–1200

Cucco A., A. Ribotti, A. Olita, L. Fazioli, B. Sorgente, M. Sinerchia, A. Satta, A. Perilli, M. Borghini, K. Schroeder, and R. Sorgente, *Oil spills prediction in the Bonifacio strait area, western Mediterranean, (2012),* Ocean Science Special Issue, *"The MyOcean project: scientific advances for operational ocean monitoring and forecasting"*, Eds. P. Brasseur, M. Bell, J. A. Johannessen, P.-Y. Le Traon, and A. Schiller, 8, 4, 443-454

Olita A., A. Cucco, S. Simeone, A. Ribotti, L. Fazioli, B. Sorgente, R. Sorgente, *Oil spill hazard and risk assessment for the shorelines of a Mediterranean coastal archipelago,* Ocean and Coastal Management, (2012), 57, 44-52

Sorgente B., R. Sorgente, A. Olita, L. Fazioli, A. Cucco, A. Perilli, M. Sinerchia, A. Ribotti, *Effects of protection rules and measures in an important international strait area: the Bonifacio Strait,* Journal of Operational Oceanography, (2012), 5, 1, 35-44

Publishing and Citing Ocean Data

Lisa Raymond, WHOI, lraymond@whoi.edu (USA)
Cyndy Chandler, WHOI, cchandler@whoi.edu
Adam Leadbetter, BODC, alead@bodc.ac.uk
Ed Urban, SCOR, ed.urban@scor-int.org
Peter Pissierssens, IOC Project Office for IODE, p.pissierssens@unesco.org
Linda Pikula, NOAA, linda.pikula@noaa.gov

Data collected in the ocean sciences, whether generated from research or operational observations, are not always deposited in national or international repositories or data centers in a format that makes them retrievable and reusable. Often, there are insufficient incentives for data submission, resulting in low submission rates and even when submitted, a bare minimum of metadata. The problem is, in part, cultural. Publications of data do not carry the same weight in deciding promotion and tenure as papers that include analyses of the data. Promotion criteria do not take into account the innovation and complexity of data acquisition in challenging environments.

The advent of publisher and funding agency mandates to make data accessible is now the "stick" being used to force data accessibility. The "carrot" is the ability to cite the data accurately and to make re-use easier, making the openness advantageous to researchers. Scientists are now becoming aware that Digital Object Identifiers (DOIs) offer the means to easily cite their datasets and gain citation metrics.

The Use Cases

The Marine Biological Laboratory Woods Hole Oceanographic Institution (MBLWHOI) Library, the Scientific Committee on Oceanic Research (SCOR) and the International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission have developed and executed pilot projects related to two use cases: (1) data held by data centers are packaged and served in formats that can be cited and (2) data related to traditional journal articles are assigned persistent identifiers and stored in institutional repositories.

IODE has a history of fostering the establishment of standards and this collaboration is building a "community" of librarians, data managers and scientists to address the data publication paradigm.

The goal of the use cases has been to identify best practices for tracking data provenance and clearly attributing credit to data creators/providers so that researchers will make their data accessible. The assignment of persistent identifiers, specifically DOIs, enables accurate data citation.

Data Associated with Published Articles

The MBLWHOI Library has developed a workflow and metadata guidelines to deposit datasets in the Institutional Repository (IR), the Woods Hole Open Access Server (WHOAS). DOIs are assigned and registered with CrossRef. Libraries have been assigning DOIs to text documents for years and DOIs are now the de facto standard for datasets.

Elsevier Publishing sought collaboration with the MBLWHOI WHOI Library in 2011. The Library's ability to assign DOIs allowed a system to be set up that enables article records in ScienceDirect to display a banner that links to datasets deposited in WHOAS that are associated with Elsevier articles. The system works for DOIs assigned to datasets either before or after publication.

The Library is also collaborating with Biological and Chemical Oceanography Data Management Office (BCO-DMO) at the Woods Hole Oceanographic Institution. Tools and procedures have been developed to automate the ingestion of deposits from BCO-DMO. A Dublin Core compliant metadata record is deposited with a copy of each dataset into the WHOAS. The system also incorporates functionality for BCO-DMO to request a DOI from the Library. This partnership allows the Library to work with a trusted data repository to ensure high-quality data, while the data repository utilizes library services and is assured that a permanent archived copy of the data is associated with the persistent DOI.

Data Held by Data Centers

The Published Data Library (PDL) was recently implemented by the British Oceanographic Data Centre. It provides snapshots of specially chosen datasets that are archived using rigorous version management. The publication process exposes a fixed copy of an object and then manages that copy in such a way that it may be located and referred to over an indefinite period of time. Using metadata standards adopted across NERC's Environmental Data Centres, the repository assigns DOIs obtained from the British Library/DataCite to appropriate datasets.

Results

The project team has compiled a Cookbook (available online as IOC Manuals and Guides No. 64; http://www.iode.org/mg64) to enable other data centers and libraries to institute these services. The current version will be updated as more experience is gained. There will be a workshop with attendees who wish to start a data repository. Feedback from this class, and others who use the document, will help determine the weaknesses and strengths of the Cookbook. Further versions will be produced as appropriate. The intention is to share the Cookbook with the oceanographic community. It has already been distributed at an international meeting and to students at an e-Repository class at IODE. Response so far has been positive.

The British Oceanographic Data Centre and the MBLWHOI Library are actively assigning DOIs to appropriate datasets enabling accurate citation and preservation. This ensures accessibility to the broader community and provides an incentive to the original researcher to make the data available. Repositories that assign DOIs to data objects are providing services not yet provided by most large national or subject data repositories. These repositories are not meant to replace traditional data centers, but rather complement the data center by providing formal data publication services. Further, this project demonstrates that data publication that enables accurate citation of data sets is manageable for organizations of any size.

Establishment of an Operational Data Management System for Korea Ocean Prediction System

Sung-Dae Kim, Korea Institute of Ocean Science and Technology, sdkim@kiost.ac (Republic of Korea)
Hyuk-Min Park, Korea Institute of Ocean Science and Technology, hyukmin@kiost.ac
Jin-Hee Lee, Korea Institute of Ocean Science and Technology, jhlee88@kiost.ac
Sang-Ho Baek, Korea Institute of Ocean Science and Technology, prideshaq@kiost.ac
Soo-Young Park, Korea Institute of Ocean Science and Technology, sypark@kiost.ac

An operational data management system has been developed during establishment of Korea Operational Oceanography System (KOOS) cuncurrently. The system is composed of the data archiving system and the data service system and main target users of it were ocean researchers. The data archiving system was set up and started to be operated in 2010. Several physical oceanographic data around Korean peninsula were collected periodically and shared with ocean modellers for data assimilation and model validation. The data service system for the ocean experts was established to distribute observation data and model output files using OPeNDAP (Open-source Project for a Network Data Access Protocol) software and LAS (Live Access Server) software from 2011 to 2012. The KOOS OPeNDAP provides a way for ocean researchers to access oceanographic data anywhere on the internet and extract sub set from the original data set. The KOOS LAS provides on-the-fly visualization and analysis of multi-dimensional scientific data for web users.

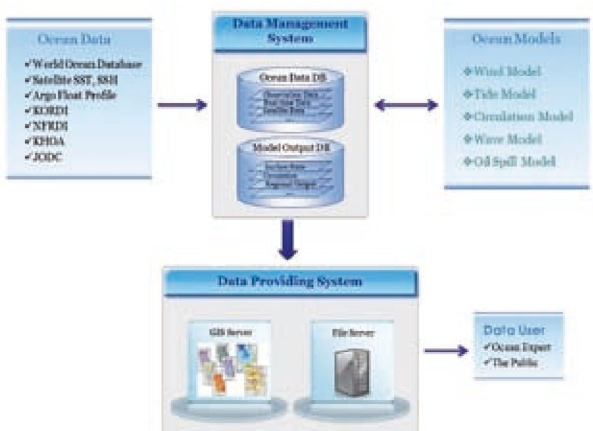


Fig. 1 - Structure of the operational data management systems for KOOS.

Data Archiving System

According to the request of ocean model researchers, several physical oceanographic data were collected and distributed through the KOOS data services. We started to collect data in the Yellow Sea, the East China Sea and the East Sea (Japan Sea) to provide them to the ocean researchers who are developing ocean prediction systems and data assimilation systems.

Almost temperature and salinity (TS) data available in the public domain and TS data produced by the Korean organizations were gathered periodically. The northern limit and the southern limit of the data collection area are 20°N and 55°N, and the western limit and the eastern limit are 110°E and 150°E, respectively. To archive TS profile data, ARGO data and World Ocean Database (WOD) data have being collected every week since 2010. The quality control software for TS profile data was developed using C language with QC criteria suggested by the ARGO program and the GTSPP (Global Temperature-Salinity Profile Program). It was applied to the collected TS profile data and QC flags were added to the profile data when the data were not meet QC criteria. The ARGO data were collected from ARGO GDAC USA (Global Data Assembly Center USA) and GDAC France in the format of netCDF and the profile data in the north west Pacific area were extracted from the original files. By the end of February 2013, the total number of vertical profile data collected from the ARGO GDAC was 44,607 deployed by 7 organizations of Korea, Japan, China, Australia and Canada. WOD data set was published in 1998 (WOD98), 2001 (WOD01), 2005 (WOD05), and 2009 (WOD09). Newly collected data are updated every 3 months and distributed through the WOD web site. Many historical TS profile data observed by CTD, OSD, BT, and PFL were retrieved from World Ocean Database 2009 and newly updated data are checked and collected every month. The total number of gathered data from WOD was 1,584,267.

To archive SST (Sea Surface Temperature) data and SSH (Sea Surface Height) data obtained by Satellite, AMSR-E SST, NCDC (National Climate Data Center) SST and AVISO SSH data were collected every week. We started to archive global 1/4° gridded AMSR-E SST data of level 3 on June 17, 2010, but data service was stopped on October 3, 2011 due to malfunction of satellite antenna. Since October 4, 2011, NCDC SST data set produced by optimum interpolation using AVHRR data have been collected from FTP service of NCDC. AVISO global 1/4° gridded SSH data is started to be collected from May 27, 2011. After we collect satellite data, we extract regional SST data and SSH data around Korean peninsula from the original global data sets and save them in the format of netCDF. We also produce the image files using extracted file and provide them through FTP.

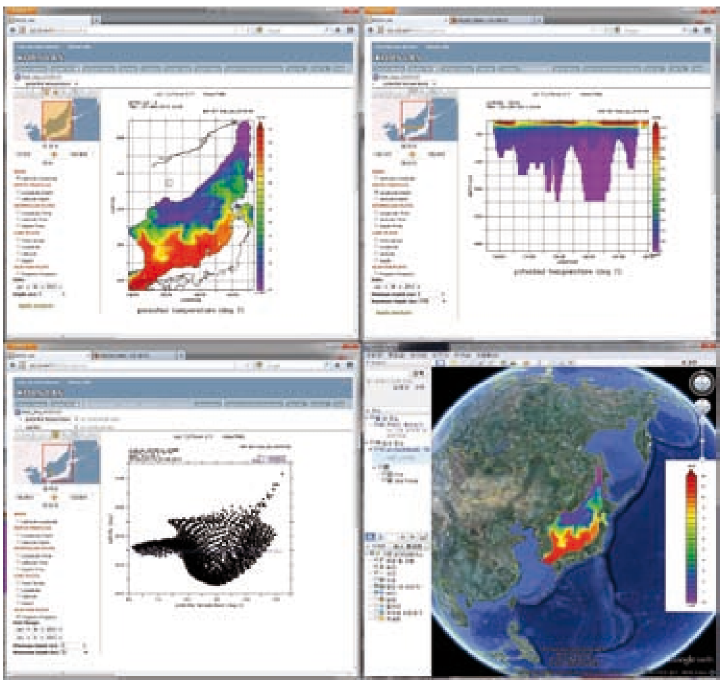


Fig. 2 - KOOS Live Access Server.

Data Service System for the Ocean Experts

The data service system for the ocean experts was established to distribute observation data and model output files. It consists of 3 file data services using OPeNDAP, LAS and FTP service. The OPeNDAP permits remote connection to the data file through various data manipulating software like Ferret, Matlab, IDL, etc. KOOS OPeNDAP service was set up to provide netCDF output files of KOOS and started its service in 2011. The Live Access Server is a highly configurable web server designed to provide access to geo-referenced scientific data. LAS enables the web user to visualize ocean data with on-the-fly graphics, to request custom subsets of variables and to compare variables. We set up KOOS LAS to provide various graphic images of several model output data produced by Korea ocean prediction system.

It offers unified access interface for multiple types of data and powerful graphic functions for section maps, depth profiles, hovmoller plots, line plots, time-series plots, and scatter plots. It also can produce animation images and KML files for the Google Earth. KOOS LAS supports the modelers and data users on their work to analyze the model output data.

Acknowledgements. This study was supported by the project titled “Development of Korea Operational Oceanographic System (KOOS) (PM57040)”.

A web-based GIS for oil spill prediction

Elena Zhuk, Cyprus Oceanography Centre, alenixx@gmail.com (Cyprus)
George Zodiatis, Cyprus Oceanography Centre, gzodiac@ucy.ac.cy
Stavros Stylianou, Cyprus Oceanography Centre, stylians@ucy.ac.cy
Robin Lardner, Cyprus Oceanography Centre, lardner@cytanet.com.cy

The Cyprus Oceanography Centre (OC-UCY) has implemented a web-based geographic information system for the existing oil spill forecasting system MEDSLIK, which has been developed at the OC-UCY. The online-MEDSLIK system allows authenticated users to use the main MEDSLIK functionality via an interactive and user friendly web application without the need of installing any software locally. The application offers three main functionalities: the input interface, simulation and visualization of the results. The web-based application consists of a number of different technologies, such as FLEX, PHP, and modestmaps API for mapping and visualization. Oil spill prediction results from the MEDSLIK module are presented remotely to the users through a rich web application in a geographical context.

Fig. 1 - Input Spill Prediction interface.

The oil spill prediction at present cover the regions of the Levantine Basin using the high resolution CYCOFOS forecasts and the Mediterranean using the MyOcean MFS regional forecasting data. The online MEDSLIK presently allows the predictions for up to two oil spills simultaneously or trajectories, including hindcasts (backward predictions). Through the output interface users have the ability to use basic map controls, view information about the amount of oil on surface, evaporated, dispersed, beached, size of spill and size of area, maximum and minimum viscosity. This information can be also viewed as time-series for the entire period of each simulation. Figs. 1 and 2 show the input interface in the Levantine Basin. Users have access to previous simulation as their applications saved separately in their own repositories. Users can register and access the online-MEDSLIK through the web site of the Cyprus Oceanography Centre www. oceanography.ucy.ac.cy. The online MEDSLIK is an important tool to assist the response agencies, locally and sub-regionally.

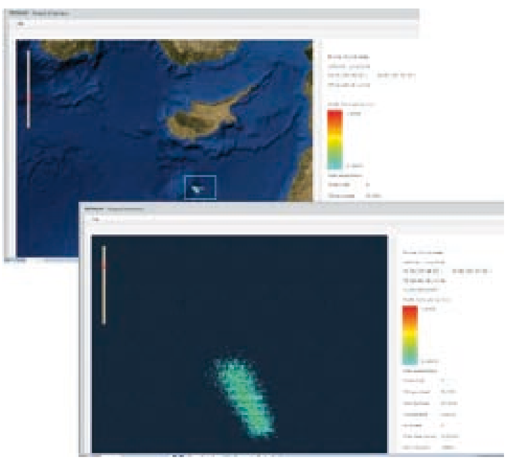


Fig. 2 - Output for Spill Prediction.

Enabling connectivity between data sources within the VECTORS project

Gerrit Hendriksen, Deltares, gerrit.hendriksen@deltares.nl (The Netherlands)
Sharon Tatman, Deltares, sharon.tatman@deltares.nl (The Netherlands)
Willem Stolte, Deltares, willem.stolte@deltares.nl (The Netherlands)
Fedor Baart, Deltares, fedor.baart@deltares.nl (The Netherlands)

Introduction

The importance of efficient data management in projects is more and more acknowledged especially in projects with high volumes of data and a high number of interaction between partners. However, the development of methods and tools for data management is costly and time consuming. Therefore Deltares introduced the OpenEarth approach in data management, which is a project superseding set of methodologies for sharing data, information and tools to enable smooth flows of information and knowledge using open source software. Intellectual properties rights on data, tools and knowledge can be protected in secure environments, but open distribution and sharing is promoted.

OpenEarth and VECTORS

Vectors of Change in Oceans and Seas Marine Life, Impact on Economic Sectors (VECTORS) is a European Commission Seventh Framework Programme (FP7) project. VECTORS is a multidisciplinary large-scale integrated European Project supported within the Ocean of Tomorrow call of the European Commission Seventh Framework Programme, which aims to improve our understanding of how environmental and man-made factors are impacting marine ecosystems now and how they will do so in the future. The multidisciplinary character of the project claims a high degree of data and information exchange between partners. Rather than sending data bilaterally between partners, a OpenEarth based central data management facility is used for unambiguous data exchange for common use.

OpenEarth data management in VECTORS aims at facilitating central access to data, model results and derived information in a dynamic and direct way which can not be facilitated by a normal web page. Data to be shared can either be stored and made accessible at partner institutes, or at the OpenEarth central data management system which is offered by Deltares.

Online access to data is generally done via netCDF formatted data in case of structured data such as model outcomes and raster data, and Relational Database Management Systems (RDBMS) for vector and unstructured data. Other types of online access of data can be arranged if needed. These systems can deal with measured data and dervied products, as well as modelled data and derived products. Model-generated data generally have the advantage of full spatio-temporal coverage, they are gap-free.

For demonstration purposes a Web Processing Service is integrated with Google Earth to visualize timeseries of modelled data from an OPeNDAP server (file format is netCDF) and measured data from a RDBMS. OPeNDAP and RDBMS enable direct model and data access via the internet without downloading complete files. Data within files can be selected based on temporal and geographical selection criteria. In this specific example high-resolution model results and monitoring data are aggregated on ICES squares and combined in time series plots.

Benefits

All partners within the project can benefit from this approach. Each institute can store its own model output or observations and is able to combine it without having the data. Dissemination is easily done by Google Earth, the low profile GIS viewer with time, space and object characteristics.

Fig. 1 - ICES oceanographic database observations as point features in RDBMS.

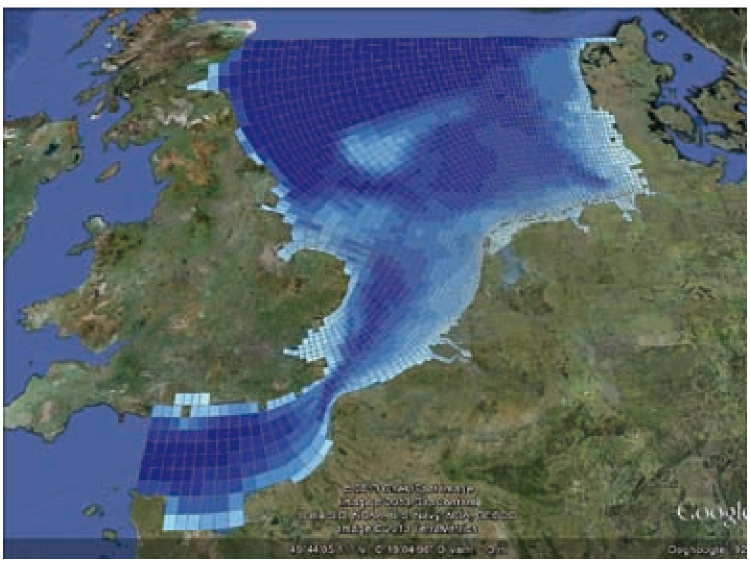


Fig. 2 - High-resolution output of Delwaq Model stored as netCDF.

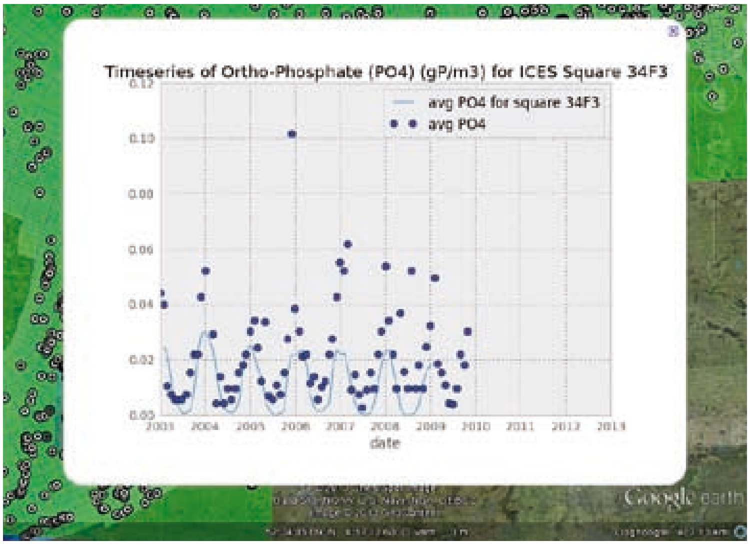


Fig. 3 - Aggregation on ICES rectangle and generation of timeseries plot.

Fig. 4 - Below some examples of the Coastal Viewer.

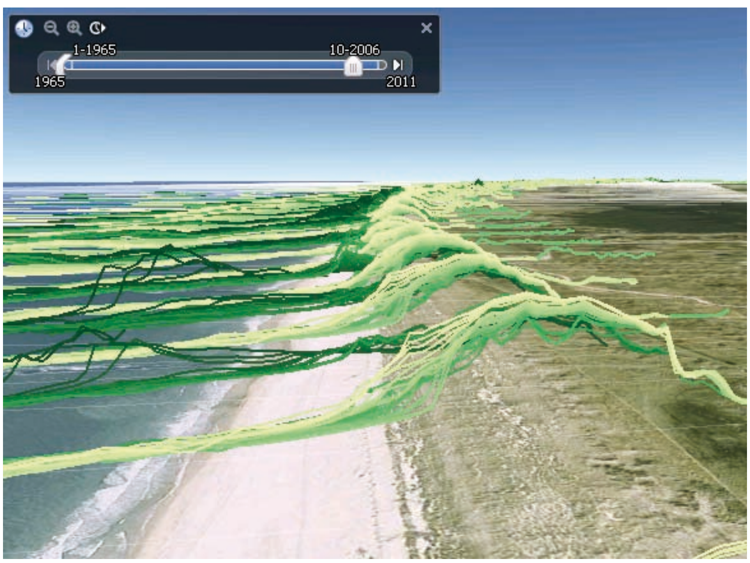


Fig. 5 - Development of Dutch coastal profiles.

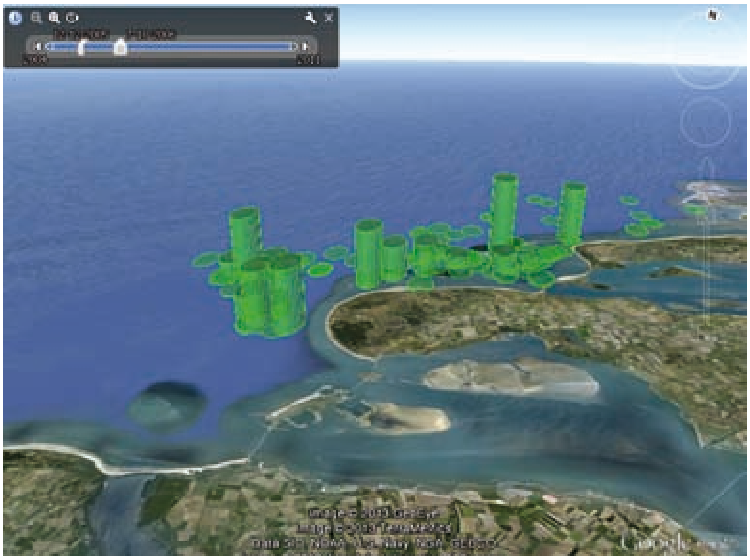


Fig. 6 - Observations of common scoters over time in front of the Dutch coast.

Fig. 7 - Animation of waterflows in the Westerscheld Eastuary

By utilizing Google Earth in the Content Management System (CMS) of Lizard it is possible to benefit from the ease and structure of a CMS and 3D Google Earth options combined in one portal. An example of such a portal is the Coastal Viewer (http://test.kustviewer.lizard.net).

References

lizard: http://lizard.org; VECTORS: http://www.marine-vectors.eu; OpenEarth: http://openearth.eu

"Up-Grade BS-Scene" and "CaspInfo" projects as examples of implementing standards provided by SeaDataNet and supported by INSPIRE

Ali Alyautdinov, Faculty of Geography of Lomonosov Moscow State University, alik@geogr.msu.ru (Russian Federation)

Ilya Kalinkin, Faculty of Geography of Lomonosov Moscow State University, ikl@geogr.msu.su

Vladimir Semin, Faculty of Geography of Lomonosov Moscow State University, vnsemin@mail.ru

SeaDataNet is a well known European research project of scientific cooperation. The main aim of this project is managing and standardization of the large and diverse data sets collected by the oceanographic fleets and the automatic observation systems from different countries and various scientific thematic topics. The SeaDataNet infrastructure networks and enhances the currently existing infrastructures, which are the national oceanographic data centers of 35 countries, active in data collection. Implementing SeaDataNet standards allows scientific organizations to publish their own data as to obtain necessary data from other oceanographic centers of the Europe.

INSPIRE (Infrastructure for Spatial Information in Europe) Directive is known as initiative of European Commission. INSPIRE Directive sets out the basic rules for the introduction of infrastructure for spatial information in Europe in order to carry out activities related to the environment, as well as events and activities that could have a direct or indirect impact on the environment.

The basic rules are:

- Metadata
- Network Services
- Data Sharing
- Interoperability of Spatial Datasets and Services
- Monitoring and Reporting

Implementing INSPIRE Directive gives extra possibilities for using and sharing data, thus creating up new informational space. In general INSPIRE Directive provides discovery, access and viewing services for different types of data using web technology.

The UP-GRADE BS-SCENE project is an FP7 EU funded project running from 2009-2011 that is building and extending the existing research infrastructure (developed under FP6 project BlackSeaScene 1) with an additional 19 marine environmental institutes/organizations from the 6 Black Sea countries. Within this project technical group from different countries developed special web-applications which was based on SeaDataNet standards for disseminating various information of Black Sea Region and supporting metadata directories for datasets - EDMED (European Directory Marine and Environmental Datasets), projects – EDMERP (European

Directory of Marine and Environmental Research Projects), organizations – EDMO (European Directory of Marine Organizations). The project has led to interesting results in the form of user interfaces that enable the users to retrieve much research information from the Black Sea. The user interfaces use map applications to support the search and these map interfaces use the latest OGC services. Also, interoperability services such as WMS (Web Map Service) and WFS (Web Feature Service), which are supported by INSPIRE Directive, were implemented in web-application. All results are accessible via the website http://www.blackseascene.net.

The CaspInfo project (Caspian Sea Environmental and Industrial Data and Information Service) is similar research project. Main aim of CaspInfo project is to create of the information model connecting a science, the legislation, the industry and structure business. One of the important elements of using such information model is presence, availability and reliability of the information on environment, including the sea, industrial activity, and the regulating certificates operating within the limits of the modern legislation of the different countries. The basic idea of project CaspInfo is close to SeaDataNet research project - compatibility and unification of the international and European standards of metadata, providing of access to the distributed bases of sea and ecological data. Data are provided by the professional national centers of gathering of the information – components of the all-European network on maintenance of on-line access to databases. The application creation methodology is based on use of the OpenGIS (Open Geospatial Consortium) standards. Basis of the concept of open systems is simplification computing systems at the expense of the international and national standardization of hard-ware and program interfaces.

The UP-GRADE BS-SCENE project and CaspInfo project are the examples of implementing SeaDataNet standards supported by INSPIRE Directive. These projects present the real way to create and to use new information model based on open standards. Such information model is opened for all users and gives real extra possibilities for exchanging scientific information and improving regional cooperation between institutes and organizations for environmental problems and protection.

The Application of Marine Data Management in Ports using Ocean Database and GIS

Mike Osborne, OceanWise Ltd, mike.osborne@oceanwise.eu, UK

Abstract

Marine environmental data and information is important to many real world situations, where ideally it is combined with operational data and managed holistically as a central resource. The data and information may then be accessed across a shared or distributed network and utilised in desktop and web based applications as part of a comprehensive 'Maritime Information Infrastructure'. Existing systems are not replaced but engineered to work in unison. An example of where this approach is being implemented and is already realizing tangible benefits is in the ports sector, where users of the 'infrastructure'' include Vessel Traffic Services (VTS), maritime operations, hydrographic surveyors, dredging operators and port engineers.

Based on projects delivered to five major UK Ports, this paper demonstrates how open standards and information systems, including GIS, are combined with off-the-shelf data management, publishing and workflow tools, and used to share and exchange data with multiple stakeholders. Hydrographic survey and other environmental monitoring data, some of it in real-time, is combined with data on aids to navigation, moorings, berths, and other infrastructure and operational data to improve efficiency and reduce costs and risk in a hard pressed commercial environment.

Keywords: Ports; Environmental Monitoring; Asset Management; Ocean Database; GIS

Ports and Harbours

Ports and Harbours are multi-faceted businesses, encompassing land and marine operations, asset management, engineering, health, safety and environment protection, security, plus numerous other activities (Fig. 1). Access to appropriate and 'fit for purpose' data and information is essential to the smooth running of these activities and hence the port. Historically, this data and information is the subject of a particular software application, rendering it inaccessible, or tied to a proprietary format. This has resulted in data 'silos' and multiple copies of the same data, which is both inefficient and prone to access and transcription errors. Furthermore, the exchange of data with contractors, regulators and other stakeholders is often a difficult and slow process.

Fig. 1 - Ports rely on a many types of data.

Data Policy and Management System

As a consequence, ports provide an ideal target for the development of an enterprise level spatial data infrastructure (SDI) to break down barriers and encourage greater efficiency and consistency through data sharing. Based on our experience of working with five major UK ports, the process commences by analyzing existing practices and systems using standard business process modelling techniques. Key data management principles are applied to form a data policy and strategy that clearly defines how existing systems can be best utilized and current workflows improved. A data management system is implemented through the application of appropriate software tools, training and mentoring. Marine data is complex and comes in a variety of different formats. Ocean Database (ODB), installed on a suitable RDBMS such as Microsoft SQL Server, is suitably structured to store and importantly load many simple and complex marine environmental and operational data types. Associated parameters used in sensor calibration, quality control and metadata creation are incorporated, facilitating onward data exchange and use. As a result, ODB provides an essential core component of the port data management system and wider information infrastructure.

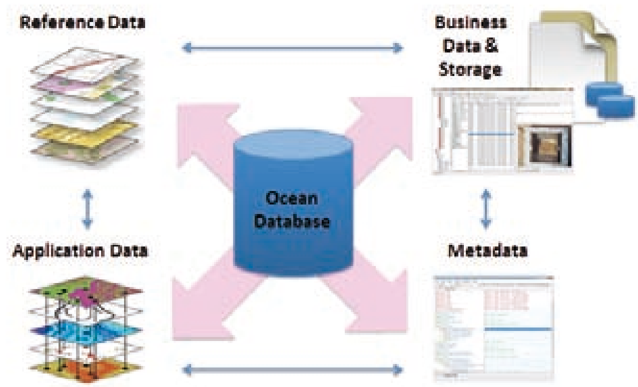


Fig. 2 - Example Port Data Management System.

Much of the data within a port is of a spatial nature i.e. it either represents a geographic feature (building, gate, wreck, navigational aid, sounding etc.) or can be attributed to a geographic feature (occupier, inspection date, sample value, tide measurement etc). This data is stored natively in ODB or retained in its native format and linked to metadata stored in ODB. Location is used to visualise, analyse and portray this data in a Geographical Information System (GIS). In addition, it can also be used to associate seemingly disparate datasets (and hence activities) and provide operational, application or business data with a spatial context. This is achieved by using location as the common factor and unique identifiers to provide a link between these and other datasets using open standards and services (Fig. 2).

Maritime Information Infrastructure

The application of data management principles and the SDI can be extended to other activities and technologies that are traditionally set apart to create a comprehensive 'Maritime Information Infrastructure' (Fig. 3). Data which is acquired for one activity can be used to support another activity, thus minimising data acquisition, storage and management overhead, reducing the potential for error and maximising re-use.

For example, environmental monitoring data can be presented in real-time at different locations and stored and accessed to support port development and regulatory compliance. Hydrographic survey data can be processed as input to paper chart production and dredging planning but also used in the creation of bathymetric Electronic Navigational Charts (bENCs) for use on Portable

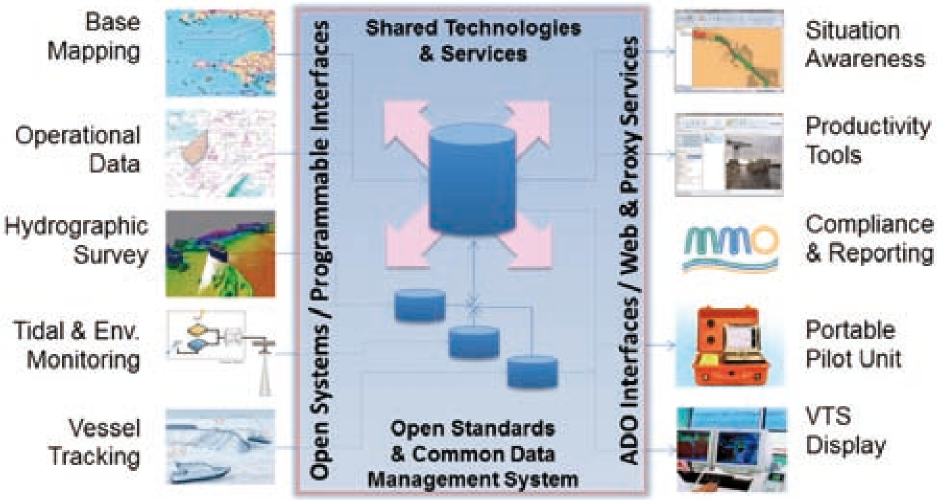


Fig. 3 - Maritime Information Infrastructure Components.

Pilot Units. The aim is not to replace existing systems but to create interoperability between these systems internally and agree industry protocols and open formats to facilitate data exchange with contractors. Advantages include the elimination of exchange sets internally and the automatic data sharing with stakeholders. Initially, this is likely to involve standard reports and templates, with web services being the ultimate goal.

ThemaMap: from data to maps

Jacques Madelaine, GREYC - CNRS UMR 6072, Université de Caen Basse-Normandie, Jacques.Madelaine@unicaen.fr (France)

Gilles Domalain, IRD - UMR 212 EME (Ecosystèmes Marins Exploités), Gilles.Domalain@ird.fr (France)

Wilbert Marin Soto, IMARPE (Instituto del Mar del Perú),Wmarin@gob.pe (Perú)

Marine research needs to manipulate different types of data coming from various sources and format: point's coordinates (fishing sites, GPS tracks...), polygon (administrative divisions, statistical areas), raster grid or geo-referenced image. Their processing generally needs the use of complex GIS software and to have specialized skills. ThemaMap aims to simplify the management and the processing of these data. It is an open-source application developed in cooperation between computer scientists and thematic researchers. It eases the production of thematic maps in order to disseminate the information as atlases. As it can easily rebuild maps from updated data, it may also be used to create monitoring panels combining different information sources.

ThemaMap organizes the information in layers as in standard GIS. Each layer has its own data source.[1] The data is presented (to the user) as a table (just like a RDBMS table or a spreadsheet). Various data transformations may be applied with the so called "adapters". One or several different visualizations of these data may then be configured and displayed on the map. Data sets are often multidimensional, and their direct visual representations are fuzzy. In order to get clear and readable visualization, ThemaMap's adapters allow to reduce the number of dimensions using aggregation under criteria of the modalities.

The visual representations are configured totally independently from the data source. Thus, any modification of the current data table will seamlessly update the current visualizations. The first visualization type is the geometric features rendering using geographical projections.[2] ThemaMap offers the classical range of symbolic data representations: disc, regular polygon, icons, pie, histogram, label. It can also easily represent flow and routes. Line width, dash pattern, line and fill colors are parameterizable. The symbols may all be resized (width, height, surface area) according to the value of a given variable (data table column). Furthermore, values of another variable may be used to colorize symbols; theirs colors depending on a classification method.[3]

ThemaMap accepts as primary data source the main geographical data standards: Esri shapefiles, MapInfo MIF/MID, Open Geospatial Consortium Geographical Markup Language or any DBMS that understands a SQL query (eg: Postgis). It accepts also NetCDF grid data. It can also load simple CSV file and it automatically interprets any couple of columns named latxxx/lonxxx as a geolocalization. A tool is also available to simply edit and modify data values or geometries.

New views of the data table may be computed with the so-called adapters which. They can operate both on geometries and data variables. They can be stacked on top of the existing data source. Any adapter can of course be removed to recover the data in the state it was before its was applied. A first example of adapters are filters that select a subset of the rows. Mono-valued filters select rows with a given value of a given column; multi-valued filters select rows where the value is member of a set; boolean expression based filters keeps only the rows for which the expression evaluates to true. Another simple adapter can be used to add a data column whose value is given with an expression (e.g. “density = population / st_area(region)”).

This add column adapter is slightly different from spreadsheet in the way that the expression is the same for all the cells of the new column. Expression may involve (just like as in a spreadsheet or in SQL): arithmetic, comparative or boolean operators, classical mathematical functions (trigonometric, logarithm, conversion (rad2deg, deg2km)), string manipulations (trim, concat, match), date/time manipulations.

More specific functions dealing with the geometries are compliant with the OpenGIS Simple Features Implementation. One can test for intersection (st_intersects), compute distance and area (st_length, st_area), and construct some new geometries (st_envelope, st_convex hull, st_intersection, st_union, makepoint, makedisc).[4] The last category of function is the statistical functions group: average, median, min, stdev, ... These functions give the possibility to specify groups defined by a general expression.

Two other adapters were developed in order to aggregate data according to specific criteria. The first one acts only on the alphanumeric data using an aggregate function such as sum or average, and may “pivot” the values of a column by creating one new column per existing modality of a variable . A second acts on geometries and is able to compute spatial aggregations such as convex hull, union or intersection of geometries sharing the same criteria. These tools allow to reduce or modify easily the number of dimensions of the data set.

It is also possible to aggregate spatial data on a regular rectangular grid using the same aggregate functions as above while keeping or deleting dimensions of the data set. The dimensions of the grid may be changed interactively. It is furthermore possible to smooth the result using kernels in case the aggregate function is a sum (different kernel functions are available). The construction of contours can be used to go back to vectorized data.

If the multidimensionality of the data is still important and forbid any clear visualizations, one can browse data using animations. Successive modalities of given variables may be selected using a mouse click or automatically thanks to a programmable timer. Each time the value change the map is redrawn seamlessly thanks to the independence of the visualizations configuration. For example, it is thus easy to synchronize different layers (e.g. catches and effort) for successive years. Furthermore, cascade control allows a variable modality to cycle over all its possible values for each change of another value. For instance, maps of catches may be displayed for each different gears and for each month of each year.

All the steps of the map’s construction may be saved and restored in a properties file. This file contains all the directives for the layers configuration, for the various features and symbols visualizations as well as for the data adapters.

This is convenient to reconstruct later a map or a single layer with updated data. It is moreover possible to store the data with the properties in a zip archive file in order to share the work.

Several types of exports are available. One can export the data in a GML file or simply in a CSV file. For its part, the maps is exportable as a jpeg or png raster image, or as SVG vector graphics image. The different frames of an animation may be exported as a collection of individual images or as an animation consisting of SVG embedded in a HTML page that can be replayed using a Web browser.

ThemaMap is used for building an atlas on small scale fisheries in Peru. The IMARPE institute collects data about individual fishing trips recording routes, catches, gears, … First visualizations of fishing areas may be simply made using discs of size proportional to the catches. Data may then be aggregate on a grid. This allows the presentation of the evolution of catches over the year, or of the repartition of the fish species using bar charts or graphs. Application of a kernel smoothing process lets areas of strong density to appear more clearly. GPS tracks of boats can be directly mapped. Speed, velocity, acceleration are calculated. These data are used to analyze the boat activity. Once again, aggregations may be used to produce more synthetic results making a monitoring dashboard.

References

[1] A special mechanism exists to share, if necessary, this source with another layer.

[2] ThemaMap is based on OpenMap (http://openmap.bbn.com) capabilities for the features rendering.

[3] Breaks may be set manually or automatically with methods such as quantiles, fixed intervals or nested means.

[4] The possibility to compute logical expressions using geometrical functions may also be used to create a new coverage from two or more existing layers

ThemaMap may be downloaded from https://themamap.greyc.fr.

General Bathymetric Chart of the Oceans (GEBCO) - Global Bathymetric Data Sets

Pauline Weatherall, British Oceanographic Data Centre (BODC), paw@bodc.ac.uk (UK)
Roy Lowry, British Oceanographic Data Centre (BODC), rkl@bodc.ac.uk (UK)

The General Bathymetric Chart of the Oceans (GEBCO) consists of an international group of experts who work on the development of a range of bathymetric data sets and data products with the aim of providing the most authorative bathymetric data sets for the world's oceans.

GEBCO operates under the joint auspices of the International Hydrographic Organization (IHO) and the Intergovernmental Oceanographic Commission (IOC) of UNESCO.

On behalf of the GEBCO Community, GEBCO's bathymetric data sets are maintained and made available by the British Oceanographic Data Centre.

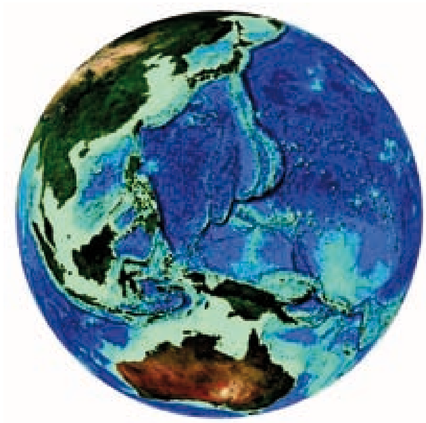


Fig. 1 - Image produced from the GEBCO global grid.

Information about recent developments and new products is given below. Further details about all of our products can be found at: *http://www.gebco.net/data_and_products/*

GEBCO's global gridded bathymetric data set

GEBCO's latest bathymetric product is a global 30 arc-second interval terrain model. The bathymetric portion of the grid was originally developed from a database of ship track soundings with interpolation between soundings guided by satellite-derived gravity data. Where they improve on this 'base grid', data sets developed by other methods have been included.

GEBCO aims to continually update its global grid file, working with regional experts to build an authoratitive global bathymetric model.

The grid was originally released in January 2009, with updated versions made available in 2010. A new release of the GEBCO grid is scheduled for publication in early summer 2013. This release will include a number of new bathymetric compilations, for example, data from the International Bathymetric Chart of the Arctic Ocean (IBCAO); the International Bathymetric Chart of the Southern Ocean (IBCSO); waters around Australia and data from the Lamont-Doherty Earth Observatory's Global Multi-Resolution Topography Synthesis.

The GEBCO grid is accompanied by a Source Identifier (SID) Grid. This indicates which cells in the GEBCO grid are based on soundings or existing grids and which have been interpolated.

GEBCO's grids are available to download from the internet in netCDF form:
https://www.bodc.ac.uk/data/online_delivery/gebco/

Free software is available for viewing and accessing data from the grids in ASCII as well as netCDF: *https://www.gebco.net/data_and_products/grid_display_software/*

Gazetteer of Undersea Feature Names

The GEBCO Sub-Committe on Undersea Feature Names (SCUFN) maintains and makes available a gazetteer giving the name, geographic location and extent of features on the sea floor.

The data set is currently accessible in the form of a spreadsheet. Work is in progress to make the data set available in GIS-friendly formats such as shapefile and KML.

The gazetteer can be downloaded from GEBCO's web site along with further information about SCUFN and the work they do in the governance of naming features on the sea floor:

http://www.gebco.net/data_and_products/undersea_feature_names/

IHO-IOC GEBCO Cook Book

In order to assist and encourage further participation in bathymetric grid development work, GEBCO has created a technical reference manual, the IHO-IOC GEBCO Cook Book, first released in October 2012.

It includes information on a wide range of topics, for example: gathering data; data cleaning; gridding examples and software overviews. Input has been provided by a number of individuals and organisations, all of whom are experts in their respective fields.

The Cook Book is made available as an Adobe PDF document and can be accessed via GEBCO's web site :

http://www.gebco.net/data_and_products/gebco_cook_book/

GEBCO world map

The GEBCO world map shows the bathymetry of the world's ocean floor in the form of a shaded relief colour map. Originally published in 2007, the second release of the map, based on the GEBCO_08 Grid, was made available in April 2013.

It can be downloaded from GEBCO's web site in the form of an Adobe PDF file :

http://www.gebco.net/data_and_products/gebco_world_map/

Web services

The GEBCO grid is made available as a Web Map Service (WMS), a means of accessing geo-referenced map images over the internet. These maps can be viewed in a web browser or Geographic Information System (GIS) and incorporated in a web application.

The WMS can be accessed from GEBCO's web site :

http://www.gebco.net/data_and_products/gebco_web_services/web_map_service/

Semantic interoperability on board: ongoing developments for EARS event logging

Karien De Cauwer, RBINS-MUMM, Karien.DeCauwer@mumm.ac.be (Belgium)
Yvan Stojanov, RBINS-MUMM, yvan.stojanov@mumm.ac.be (Belgium)
Jordi Sorribas, UTM-CSIC, sorribas@utm.csic.es (Spain)
Paolo Diviacco, OGS, pdiviacco@ogs.trieste.it (Italy)
Marie-Paule Corre, IFREMER, marie.paule.corre@ifremer.fr (France)
Alessandro Busato, OGS, abusato@ogs.trieste.it (Italy)
Serge Scory, MUMM, serge.scory@mumm.ac.be (Belgium)
Jean-Marc Sinquin, IFREMER, Jean.Marc.Sinquin@ifremer.fr (France)
Marc Nokin, IFREMER, Marc.Nokin@ifremer.fr (France)

Within the frame of Eurofleets, the development of common software tools included the development of a reporting tool. EARS, the Eurofleets Automatic Reporting System consists of an automatic data acquisition part, a manual event module and reporting functionalities. The manual event module aims in the first place to have accurately registered metadata which is only possible when registered as close as possible to the time of measurement or sampling. Besides complete and correct metadata, interoperable information is of high importance for further use, like reporting, browsing and dissemination, over the different cruises and vessels. Existing vocabularies should be considered as input and vice versa. Especially the SeaDataNet data dictionaries served by the Natural Environment Research Council (NERC) Vocabulary Server are highly relevant. By connecting to these, the standardisation of meta-information starts a step sooner: at sea instead of at the data centre.

When also looking at the user's needs, more issues arise. Difficult sampling conditions and time constraints are important parameters to consider when developing software for use on board. User friendliness is thus of utmost importance: fast entry, multilingualism, addition of new terms and relations, use outside, nested tools … To centralize all meta information in one tool, a clear way for users to enter and retrieve any additional characteristic or parameter accompanying the sample, observation or measurement according to the specific sampling or measurement procedure has to be foreseen.

A thorough brainstorming on the characteristics of an event has resulted in the development of event concept. The characteristics and user feedback of existing event loggers like CASINO+ of IFREMER and Ours, the Onboard Underway Registration of Samples, at MUMM have been taken into count. The different composing concepts as well as the relations between them have been identified:

- **Subject:** the domain in which the event takes place, e.g. a seismic system
- **Tool:** the device that produces the event, e.g. a Niskin bottle
- **Category:** the kind of event taking place, e.g. sampling

- **Action:** the actual task performed, e.g. close bottle
- **Comment:** a free text field that is left to the operator to enter more information
- **Actor:** the person performing the action
- **Action_property:** any additional characteristic or parameter accompanying a given action for which the user needs to enter a value onboard f.e. volume of water centrifuged.

Commonality between these concepts and existing vocabularies has been identified. In particular, the SeaDatanet L05 SeaDataNet device categories can be referred to for ‘Subject’, while the L221 Seavox device catalogue provides a good list of measurement devices for the ‘Tool’. Having events built over the combination of parts, allow high levels of flexibility but at the same time introduces the risk to get far from the actual acquisition practices. In order to avoid errors and at the same time increase user friendliness, it is very important to introduce a mechanism that can constrain choices upon knowledge and existing practices. The relations between the instances of different concepts are thus important.

At the start of Eurofleets 1, ‘Action’ terms have been defined for different tools and disciplines but it was obvious that all aspects of all disciplines could not be covered since the start. A flexible approach was chosen to prevent users being blocked during the survey. EARS Version 1 is built in such a way that relationships between terms are being created by scientific experts creating their personal and discipline-specific configuration tree. New terms can be easily added. After that, an event is generated by one double-click or drag-and-drop. In fact, an ontological instance is being created during the logging of information.

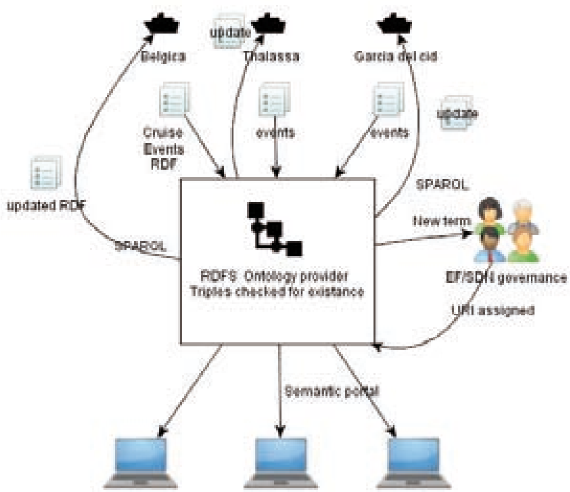


Fig. 1.

The ontology implementation for the event logger will enable the reference to existing controlled vocabularies or taxonomies as well as the usage and sharing of relations between terms as for example the possible actions with a given tool. A leap forward is ongoing within Eurofleets 2 through the full integration of the ontology. The so-far logged events are being gathered in an extended event ontology in which terms and relations of existing vocabularies are sourced from their originating namespace. As following step, the user interface to enter manual events will be adapted to be based on RDF files (SKOS, OWL) and SPARQL queries.

In order to manage the terms, a governance scheme will be implemented in close collaboration with SeaDataNet. New terms are to be expected as input for the NERC Vocabulary Server catalogues. Any additional list, relations between terms and commonly used local names will be served by a Eurofleets ontological server and SPARQL endpoint for use within the tool and for browsing linked event information. Contacts and concrete actions are defined to harmonise and establish a common governance beyond the EU within the ODIP project.

The EMECO Datatool: an online assessment and reporting system for co-production of environmental assessments of shared marine waters

Jo Foden, Centre for Environment, Fisheries and Aquaculture Science (Cefas), jo.foden@cefas.co.uk (UK)
David Mills, Centre for Environment, Fisheries and Aquaculture Science (Cefas), david.mills@cefas.co.uk
Kathryn Keeble, Blue Lobster IT Ltd kathryn@bluelobster.co.uk
Simon Keeble, Blue Lobster IT Ltd, simon@bluelobster.co.uk

We present the outcomes of an international collaborative exercise assessing the environmental condition of the central southern North Sea, which used a web-based Datatool.

The condition of the marine environment is regularly assessed for a number of drivers, at a range of scales. For example, national assessments such as UK's Charting Progress 2, European Directives such as the Marine Strategy Framework Directive, international agreements such as

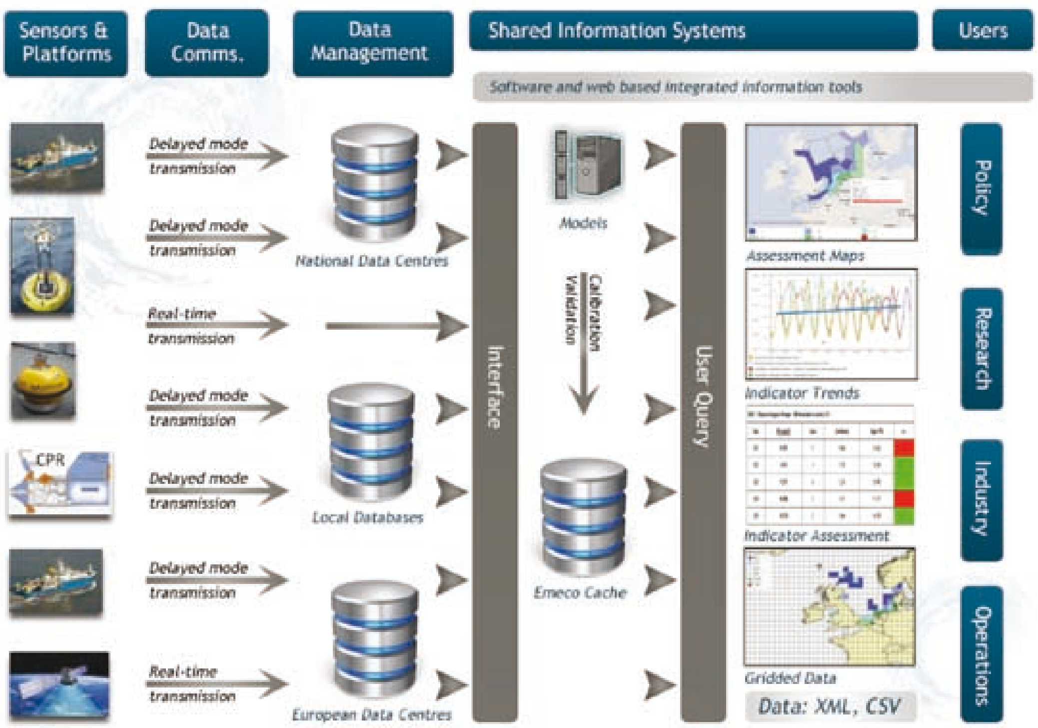


Fig. 1 - Schematic of the European Marine Ecosystem Observatory (EMECO) Datatool.

OSPAR, to global scale assessments such as the Global International Waters Assessment. Robust assessments are founded on a sound evidence base, relying on quality assured data from multiple sources.

Bringing together the required evidence can be a long and complicated multi-stage process. Geo-referenced data for multiple parameters need to be assembled in a common format for the purposes of analysis and assessment against predefined indicators or targets. It is particularly challenging to accomplish this for water bodies shared with European countries and to reach an outcome to an assessment agreed by all parties.

As part of the European Marine Ecosystem Observatory (EMECO) initiative, a web-based Datatool has been developed to automate and streamline the process of collating and standardising data from many different sources (Fig. 1). The Datatool uses open source technology and accepts data in a large number of common data formats, including NetCDF, XML, CSV, Access, ASCII, KML and TXT. The online querying function enabled us to select and interrogate data for several parameters in our collaborative exercise – nutrients, chlorophyll and dissolved oxygen. These parameters are commonly used for all marine assessments. We demonstrate how the Datatool was used to integrate multi-platform, multi-parameter and multi-national data, and present the results of queries that produced bespoke assessment products; maps, time-series plots and the raw data (CSV and XML formats). Providing the underlying data alongside assessment products gives transparency to the assessment process. For any assessment that may be legally challenged it is important to estimate the level of confidence and we show how this is carried out in the Datatool. Furthermore the Datatool has report-compilation functionality, which was used to collaboratively produce an agreed final assessment of the environmental condition of the offshore waters in the southern North Sea.

Kiel Data Management Infrastructure for Marine Sciences

Pina Springer, GEOMAR Helmholtz Centre for Ocean Research Kiel, pspringer@geomar.de (Germany)
Hela Mehrtens, GEOMAR Helmholtz Centre for Ocean Research Kiel, hmehrtens@geomar.de
Dirk Fleischer, GEOMAR Helmholtz Centre for Ocean Research Kiel, dfleischer@geomar.de
Carsten Schirnick, GEOMAR Helmholtz Centre for Ocean Research Kiel, cschirnick@geomar.de
Andreas Czerniak, GEOMAR Helmholtz Centre for Ocean Research Kiel, aczerniak@geomar.de

The Kiel Data Management Team (KDMT) is a joint group of large-scale projects at GEOMAR and Kiel University: the Cluster of Excellence ‘The Future Ocean’, the Collaborative Research Center ‘SFB 754’, the EU FP7 project ‘ECO2’, the Helmholtz Initiative and Networking Fund Project ‘MaNIDA’ and the BMBF funded projects ‘BIOACID II’ and ‘SOPRAN III’. Its aim is to provide ONE central place for description, storage and archiving of research data for all marine sciences in Kiel, independent of project status but with specific access restrictions for each project.

This system facilitates preparation of data for paper publication, data exchange within the project and data publication.

Services within the Virtual Research Environment (VRE)

The Kiel Data Management Infrastructure (KDMI) provides a sustainable support platform for marine sciences in Kiel. The KDMI offers integrated data management solutions for metadata description, exchange and citeable publication of the research data. The main goal is to provide private working platforms to working groups and projects including the possibility of sharing information on research activities with external partners and to enhance cooperation and outreach. Therefore the web based data management portal (https://portal.geomar.de/) offers private and public web pages, wikis, blogs and separate document exchange platforms. Information and data files about marine expeditions, numerical models, experiments and projects are merged in one central place, the Ocean Science Information System (OSIS). Different quality flags and access restrictions can be assigned to uploaded data files. As soon as files get the status ‘final’ they will be submitted to World Data Centers (e.g. PANGAEA) for publication. The linkage between data and print publications is also an essential objective and achieved by linking online data and records in the institutional repository OceanRep to expeditions or numerical models stored in OSIS and and other ‘Open Access’ data repositories.

A geographical search of sampling locations as well as visual representation of cruises and expeditions in virtual globe applications, eg. Google Earth, are easily on hand by means of autogenerated KML files.

Each sampling location is described in detail by coordinates, depth/altitude, sampling technique and principal investigator. Additional information about available data files, publications and other related links to OSIS and OceanRep is displayed as well.

Usage and Cooperations

Currently about 400 active registered users collaborate in more than fourty projects and working groups within this virtual research environment. Close cooperation with project leaders and managers ensures and enhances data flow between scientists as well as to 'Open Access' data repositories.

The KDMT participates in overarching data integration projects like the Marine Network for Integrated Data Access (MaNIDA) to achieve harmonization of data access in marine sciences, and the PubFlow project in order to develop new generic solutions for workflows within publication processes.

Near real-time quality data from ships to land

Pekka Alenius, Finnish Meteorological Institute, Pekka.alenius@fmi.fi (Finland)
Kimmo Tikka, Finnish Meteorological Institute, kimmo.tikka@fmi.fi
Riikka Hietala, Finnish Meteorological Institute, riikka.hietala@fmi.fi

Near real-time data flow from instruments to databases and Internet is todays routine. A research ship is the most versatile instrument that has great possibilities to quality control the data right on the spot. This requires that the ship's databases are up-to-date. Here we present an operational system that is used in Finnish Meteorological Institute and Finnish Environment Institute. It is a two-branch system where the station profile data part serves mainly research ship Aranda and the "flow-through" data part serves mainly data from Ferry Boxes installed on merchant ships on routine routes (the Alg@Line). The two branches share common metadata, vocabularies and automatic data flows to land based databases.

The research ship system consists essentially of Laboratory Information Management System (LIMS), scientific databases and cruise planning and reporting system. The LIMS and scientific databases make possible to quality control the data on-board. However, in practise the quality control is done in laboratory after the cruise, too. Thus the data sent from the ship during the cruise is flagged as raw-data. It is anyhow available to scientists of the institute in the same databases as the 'old' data. One part of research ship system's metadata is the XML-formatted cruise plan. The plan is updated by the system and by scientific crew on-board and it is send to land with the data. Thus it is possible to follow the plan even on land during the cruise. The Ferry Box system has automatic quality control routines because there is no scientific personnel on-board during the cruise.

Fig. 1 - Research ship Aranda.

Both of the system branches have their own main databases that are specifically constructed for the fundamental data type: either for station profile data or for 'flow through'-data from continuously moving platform. The system generates automatically data files to be sent from the ship to land. On land, automatic routines read the received files and upload the data to databases. The system sends only new data that has not yet been sent. There are some data types like navigation data and meteorological data that are sampled and stored on-board research ship with so high frequency that it is non-economical to send the complete data to land during the cruise. In these cases the system sends coarser time resolution data to land using e.g. 10 minutes time steps. The whole datasets are transferred to the institute after the cruise.

Preliminary versions of the system have been used for two decades and the present version for several years. Because of the communication limitations and costs the data has up to now been sent basically once a day. New advances have made possible to send the data more frequently. Part of the received data is directly put to the Internet (www.itameriportaali.fi/en_GB/) to be available for the public and most of the data is in the databases available for internal use of the local scientists already during the cruise. When the proper quality control of the data has been done in the laboratory after the cruise, the data is made available through e.g. SeaDataNet.

On-line ice resistance survey for the ships supporting winter navigation in ice channels of the Baltic Sea

Tarmo Kõuts, Marine Systems Institute at Tallinn University of Technology, tarmo.kouts@msi.ttu.ee (Estonia)
Madis-Jaak Lilover, Marine Systems Institute at Tallinn University of Technology, madis-jaak.lilover@msi.ttu.ee (Estonia)
Kaimo Vahter, Marine Systems Institute at Tallinn University of Technology, kaimo.vahter@msi.ttu.ee (Estonia)

The winter navigation occurs mostly in ice channels, which icebreakers brake into the fast or drift ice. Every ship proceeding in such an ice channel experiences certain resistance, which could vary quite considerably both in space and time. This study is aimed to describe the nature of ship resistance in ice channels in relation to forcing functions. In order to build up basic knowledge and understanding of ice resistance, *in situ* measurements directly in the areas of winter navigation are essential.

The measurement technology is based on the understanding that in different ice conditions (smooth or ridged ice, ice compression, etc.) the interaction of ice with ship hull results in ship hull vibration of different intensity which then is registered and taken as a measure of ship resistance in ice. The icebreaker of the Estonain Maritime Administration EVA-316 was instrumented and vibrations of the ship hull are recorded by means of 3D acceleration sensor which is tightly fixed to the ship hull.

Accelerations in three dimensions x-y-z are measured in 4Hz regime, processed, maximums per minute found and recorded on SD memory card together with GPS position and time in one minute interval. In order to have online control over the ship hull vibration, the acceleration data and ship position are transferred in real time to FTP server of the Marine Systems Institute, using GSM/GPRS protocol. Next, after a preliminary analysis of data, the ice resistance index is calculated. The comparison of the acquired ship hull vibration data with satellite ice images showed that the obtained data well distinguish the open water and ice conditions. Different severity of ice conditions could also be estimated.

An attempt to relate the ship hull vibration data and ice resistance index to the relevant forcing parameters like wind speed and direction was successfully accomplished. The analysis of 2012 March data of the Gulf of Riga cruises of IB EVA-316 revealed that in case of similar wind speed and ice properties, the case with the wind direction perpendicular to the ship course had 15% higher mean ship hull vibrations than the case of ship course parallel to the wind direction. We can conclude from the latter that the wind direction and especially the angle between the wind direction and ship course, is an important factor determining the ship resistance in ice channel. The ship hull vibration data showed adequately also different ship manoeuvres in ice, which icebreaker performs during the assistance of merchant vessels.

The recorded data on the ship hull vibrations together with the ship speed and course form dataset which enables to assign a specific rank for the severity of ice conditions directly in locations of ship operations, where this information is most needed. In order to augment that system with forecasting skill, a fuzzy logic relational scheme was defined, applied and validated. The forecast of ship resistance in ice was designed and realised in pre-operational mode for the fairway into Pärnu Port, Gulf of Riga, the Baltic Sea. The study concludes that the ship hull vibration measurements applied for the detection of ship resistance in ice channel is a useful tool for on-line monitoring of ice conditions along the fairways of winter navigation. The collected data can be used for validation of models of ice dynamics and forecast of ice conditions, as well as of ice compression in particular. The shipborne measurements of ice resistance serve also goals of statistical analysis of winter navigation in certain sea areas and help optimizing the future winter navigation.

Application of the Data-Interpolating Variational Analysis (DIVA) to sea-level anomaly measurements in the Mediterranean Sea

Charles Troupin, IMEDEA (UIB-CSIC), ctroupin@imedea.uib-csic.es (Spain)
Alexander Barth, GHER, University of Liège, a.barth@ulg.ac.be
Jean-Marie Beckers, GHER, University of Liège, JM.Beckers@ulg.ac.be
Ananda Pascual, IMEDEA (UIB-CSIC), ananda.pascual@imedea.uib-csic.es

In ocean sciences, numerous techniques are available for the spatial interpolation of in situ data. These techniques mainly differ in the mathematical formulation and the numerical efficiency. Among them, DIVA, which is based on the minimization of a cost function using a finite-element technique (Fig. 1). The cost function penalizes the departure from observations, the smoothness or regularity of the gridded field and can also include physical constraints. The technique is particularly adapted for the creation of climatologies, which required a large to several regional seas or part of the ocean to generate hydrographic climatologies.

Sea-level anomalies (SLA) can be deduced from satellite-borne altimeters. The measurements are characterized by a high spatial resolution along the satellite tracks, but often a large distance between neighbour tracks. This implies the use of simultaneous altimetry missions for the construction of gridded maps. An along-track long wave-length error (correlated noise, e.g. due to orbit, residual tidal correction or inverse barometer errors) also affects the measurement and has to be taken into account in the interpolation.

In this work we present the application and adaptation of Diva to the analysis of SLA in the Mediterranean Sea and the production of weekly maps of SLA in this region.

Determination of the parameters

The two main parameters that determines an analysis with DIVA are the correlation length (L) and the signal-to-noise ratio (SNR). Because of the particular spatial distribution of the measurements, the tools implemented in Diva for the analysis parameter determination tend to underestimate L and overestimate SNR, leading to noisy analysis (the observation constraint dominates the regularity constraint). Some adaptations of the tools are necessary to solve this issue.

Numerical cost

Because of the large number of observations to be processed (in comparison with in situ measurements on a similar period), the interpolation method employed is expected to be numerically efficient. Improvements in the implementation of Diva further improved the numerical performance of the method, especially thanks to the use of a parallel solver for the matrix inversion. The performance of finite-element mesh generator was also enhanced, so that interpolation of a data set of more than 1 million data points on a 100-by-100 grid can be performed in a few minutes on a personal laptop.

Analysis and error field

The analysis and error fields obtained over the Mediterranean Sea are compared with the available gridded products from AVISO. Different ways to compute the error field are compared. The impact of the use of multiple missions to prepare the gridded fields is also examined.

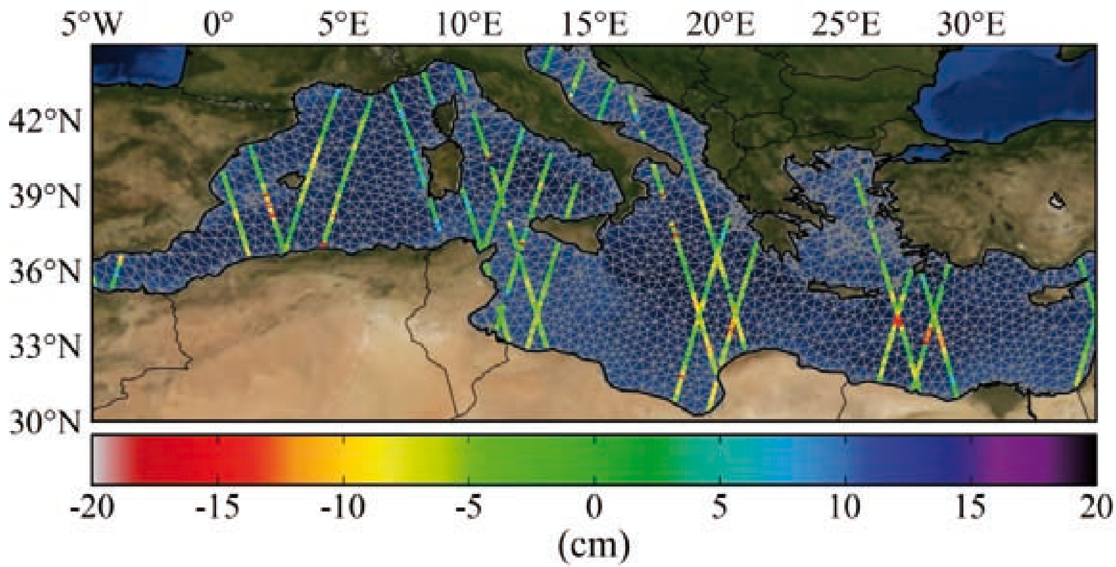


Fig. 1 - Area of interest, finite-element mesh and sea-level anomalies measurements (Envisat) for the period 6-13 May, 2009. [Data from AVISO].

A system for managing metadata using XML format

Elena Partescano, OGS-Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, epartescano@ogs.trieste.it (Italy)

Alberto Brosich, OGS - Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, abrosich@ogs.trieste.it (Italy)

Alessandra Giorgetti, OGS - Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, agiorgetti@ogs.trieste.it (Italy)

By definition, metadata, is information associated with data, answering questions as where, how, when and by whom the data was acquired. Taking part in European projects such as the EU-SeaDataNet (the Pan-European infrastructure for ocean and marine data management) made it necessary to use XML (Extensible Markup Language) as a standard file format for sharing metadata.

At present, the Italian National Oceanographic Data Centre (OGS/NODC) has all its data and metadata contained in an Oracle relational database, and some metadata are managed using XML documents following standard schemata (from SeaDataNet Project).

This paper shows the system OGS/NODC is currently using to manage oceanographic XML metadata, stored as is, into a relational database and a possible future development.

The participation to EU-SeaDataNetProject has grown the need to create and share metadata in XML format. Tools for managing it (as Mikado Software) are available, therefore the next step was linking metadata XML files to data loaded into relational database and building a web service to allow the connection between the managing tool (Mikado) and the database.

We evaluated and decided that the best solution was to load into the Oracle database the whole XML metadata files, using a specific data type. Using XML and XQuery functions it is possible to store, extract and manage different kinds of information that might be exchanged at the European level. Furthermore, with a RESTful (REpresentational State Transfer) Web Service we have a simple and standard interface for quickly and easily creating, modifying and deleting records containing XML documents inside the database. Finally, through the use of that RESTful Web Service it is possible to decouple the applications from the database, so that through the use of software that manages HTTP URLs, such as the Mikado (SeaDataNet project), the XML documents can be inserted, updated and deleted inside the database without the need for a direct connection to it.

Future development

Another way to manage XML files could to be to use a native XML database (Fig. 1): the documents are stored in a database designed especially for storing XML, supporting XPath and XQuery, to retrieve it. A native XML database doesn't use a relational model, used in common RDBMS databases.

In a native XML database (as for example: eXist, BaseX, Berkeley DB XML), the entire XML file can be stored in a single place. In this way, if the XML file or a part of it needs to be retrieved, only one searching index and only one reading index is required to recover the information, whereas a RDBMS database needs several indexes for searching and reading to recover the data. The main difference is that their inner model is based on XML and not something else.

The major advantage using a native XML database is that a database schema is not required to store textual or binary data and documents (as for the relational databases). XML databases allow to manages complex data relationships that are not easily managed in relational rows and columns. The database structure come from the XML schema, and can be easily adapted if necessary.

For these reasons, a native XML database fits perfectly with data-intensive uses such as archiving metadata.

In the future, we would prefer to use a native XML database to store the XML metadata file (Fig. 1), but even if it seems the more logical choice, several problems could arise, as for example how to join data and XML metadata (relational database and native XML database).

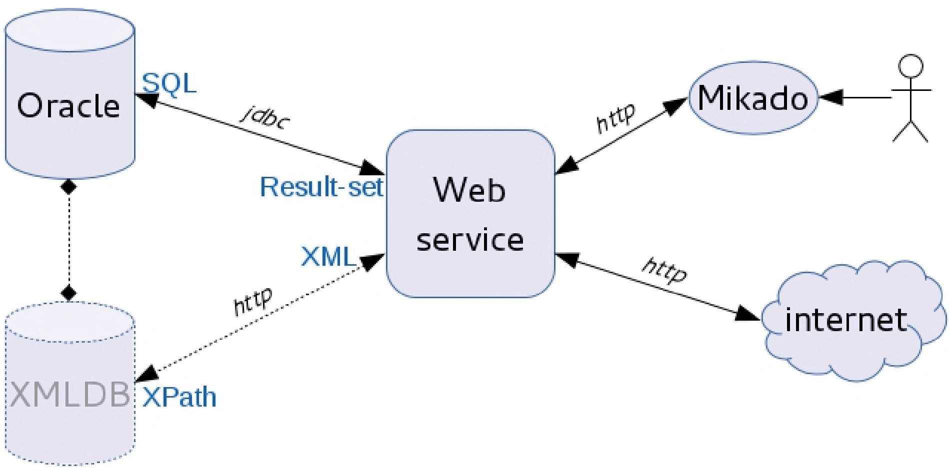


Fig. 1 - Work-flow (Dashed lines represent the future developments while continuous lines describe the present situation).

Conclusion

By using a database, relational as Oracle (currently) or Native XML database (possible future development), a RESTful Web Service and a software that manages HTTP URLs (as Mikado), we have the advantage that every single element is independent from the others. The independence between elements guarantees us more versatility. This has the benefit that changes in the database don't affect the operation of the system; at the same time, there is no obligation to use the Mikado software for managing the XML data and there is complete autonomy from the Web Service used including the implementing technology.

Browsing and Selection of CTD/XBT Data through Open-Source GIS Clients

Anna Vetrano, CNR ISMAR La Spezia, anna.vetrano@sp.ismar.cnr.it (Italy)
Alessio Bechini, University of Pisa, Dept. of Information Engineering, a.bechini@ing.unipi.it (Italy)
Catia Chiappini, CNR ISMAR La Spezia, catia.chiappini@sp.ismar.cnr.it (Italy)
Katrin Schroeder, CNR ISMAR Venezia, katrin.schroeder@ismar.cnr.it (Italy)

ABSTRACT

The management and storage of oceanographic data collected during dedicated in-situ surveys poses significant challenges, especially when their volume and variety are relevant[1]. This study is aimed at providing a convenient way to access distributed data regarding hydrological vertical profiles, focusing on measurements from instruments like CTD and XBT. The principal idea is to make a concerted use of an open source GIS client (namely, Quantum GIS) and relational, ISO-19125-compliant data repositories[2]. The adoption of a proper data/metadata model allows an effective access and browsing of the repositories' contents[3], and an ad-hoc plugin let the user export datasets made of elements selected from multiple sources, on a selected area, through the user-friendly GUI of the GIS client.

INTRODUCTION AND RATIONALE

In the last decade the European oceanographic community has intensively worked to create important networks and/or projects dedicated to the definition of standards for data and metadata, to proficiently exchange scientific data and equipment (SeaDataNet, Eurofleets, ecc.). A first step towards this ambitious goal could be represented by an effective organization of in-situ observational data[4] in local repositories, which may allow a simpler, integrated fruition for research purposes[5]. According to this vision, here we report an experience in browsing and exporting CTD and XBT data (Fig. 1) through the popular open-source QGIS client[6]. For example, this way to

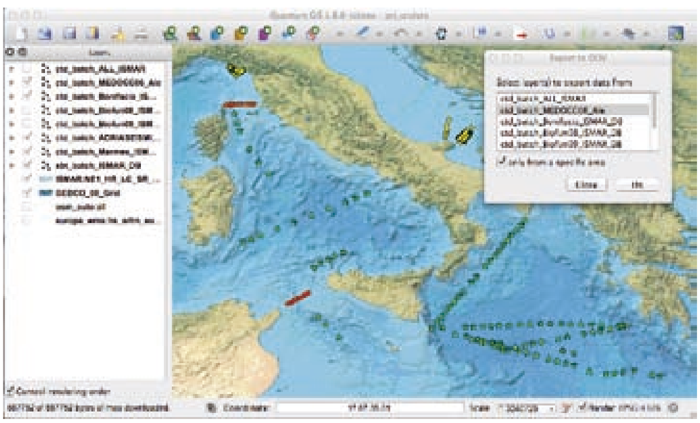


Fig. 1 - Browsing of CTD/XBT batches and corresponding exportation in QGIS.

operate can significantly help an ocean circulation modeler who asks for hydrological data in a particular basin, or in a channel connecting two sub-basins, for validation of a certain model itself[7]: The operator needs to retrieve data in a particular area, possibly across several years, maybe from larger datasets in different local repositories.

The overall approach relies on a flexible data-metadata organization within relational DBMS repositories, and a specific QGIS plugin to export the selected datasets.

DATA/METADATA ORGANIZATION

Data and metadata have been organized in a layered fashion, as depicted in Fig. 2. The two upper levels group up general and batch-specific features, respectively. Below, individual measures are kept according to their collection batch. Spatial attributes have been dealt with using PostGIS[8], in accordance to ISO-19125. Whenever information present on different tables has to be exposed to external GIS access, proper (joined) view can be defined.

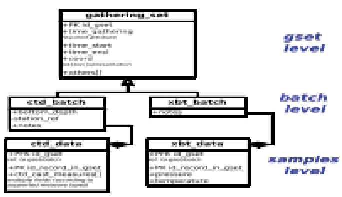


Fig. 2 - UML representation of the reference data/metadata model.

SYSTEM DESCRIPTION

In a GIS client, geo-referenced data are managed and presented in layers. In our case, layers can be defined accessing and selecting data from PostGIS sources; different layers can refer different repositories (Fig. 3). A specific collection of layers is represented by a “project”. The data user can interact with the QGIS GUI to specify what data to access/show, and can work on the map to precisely select the elements to be exported for later investigations, e.g. using ODV . In

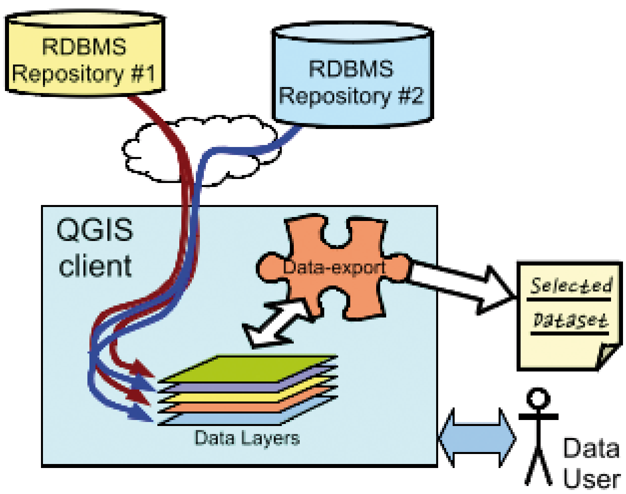


Fig. 3 - Schematic view of the interactions between the system components.

this setting, the crucial role played by the data exportation plug-in (Fig. 3), developed on purpose, can be performed referring to the standard data/metadata model. The described approach can be used for any kind of observational data, but in practice it has turned to be particularly useful with data that, like CTD and XBT, refer to single gathering points and span both extended areas and time intervals.

References

[1] K.S. Baker, C.L. Chandler, "Enabling long-term Oceanographic Research: Changing Data Practices, Information Management Strategies and Informatics," Deep Sea Research Part II: Topical Studies in Oceanography 55 (18–19) (2008) 2132–2142, DOI: 10.1016/j.dsr2.2008.05.009.

[2] Open Geospatial Consortium Website. URL: http://www.opengeospatial.org

[3] A. Bechini, and A. Vetrano, "Management and Storage of In-situ Oceanographic Data: An ECM-based Approach," Information Systems, 28(3), May 2013, Elsevier, pp. 351-368 - DOI: 10.1016/j.is.2012.10.004

[4] U. Send, "In-situ Observations: Platforms and Techniques," in: E.P. Chassignet, J. Verron (Eds.), Ocean Weather Forecasting, Springer, 2006, pp. 191–206, DOI: 10.1007/1-4020-4028-8_7.

[5] QuantumGIS Open-source Geographic Information System. URL: http://qgis.org

[6] A. Vetrano, E. Napolitano, R. Iacono, K. Schroeder, G.P. Gasparini, "Tyrrhenian Sea Circulation and Water Mass Fluxes in Spring 2004: Observations and Model Results," Journal of Geophysical Research 115 (C06023) (2010) 1–19, DOI: 10.1029/2009JC005680.

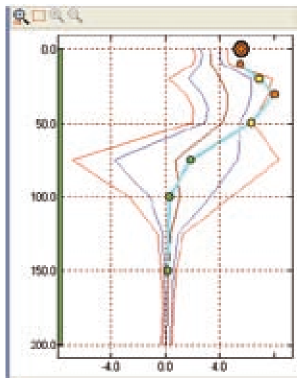
[7] R. Obe, L. Hsu, PostGIS in Action, Manning Publications, Co., Stamford, CT, USA, 2011

[8] R. Schlitzer, "Interactive Analysis and Visualization of Geoscience Data with Ocean Data View," Computers & Geosciences 28 (10) (2002) 1211–1218, DOI: 10.1016/S0098-3004(02)00040-7.

On-line quality control service of MHI NASU

Eugeny Godin, MHI NASU, godin.eugeny@nodc.org.ua (Ukraine)
Alexey Khaliulin, MHI NASU, khaliulin.alexey@nodc.org.ua
Andrey Ingerov, MHI NASU, ingerov.andrey@ nodc.org.ua
Elena Zhurk, MHI NASU, zhuk.elena@ nodc.org.ua

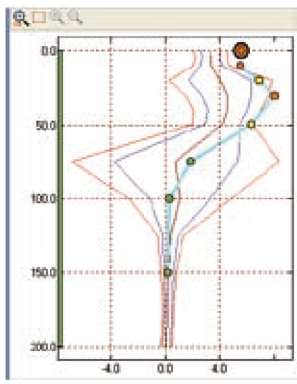
A test version of on-line quality control service was developed in MHI NASU. We took into consideration the approaches to data quality control of such data products as WOD; MEDAR/MEDATLAS II; NATO TU BLACK SEA databases also we paid attention to recommended by IOC UNESCO procedures for validation of oceanographic data. Using experience of international collaboration in the field of data collection and quality check we have developed the test version of on-line quality control service providing preliminary(automatic) data quality check procedures. QC procedure includes uploading of ODV file, metadata and data quality control and returns file with QC flags corresponding to SeaDataNet vocabulary L 201.



A

Metadata control provides such QC tests as
- location check
- date and chronology check,
- ship velocity check,
- sea depth check,
- last sounding value check.

Data quality control tests are
- sounding value check,climatic check, if possible (Fig. 1 a)
- range check (if climatic check is not available)
- density inversion check for hydrological data (Fig. 1 b)



B

Fig. 1 - Example (principles) of climatic (a) and density inversion (b) QC tests.

We estimated climatic characteristics for different squares of the Black Sea for each month for temperature, salinity, density, oxygen and hydrogen sulfide. For other parameters checking we use available climatic informations as well as regional and global ranges

Addressing uncertainty through data quality issues in environmental assessments to support improved decision making

Maria Pantazi, Hellenic Centre for Marine Research, mpantazi@ath.hcmr.gr (Greece)
Vassiliki Vassilopoulou, Hellenic Centre for Marine Research, celia@hcmr.gr
Yiannis Issaris, Hellenic Centre for Marine Research, issaris@hcmr.gr
Athina Kokkali, Hellenic Centre for Marine Research, athkokkali@gmail.com
Maria Salomidi, Hellenic Centre for Marine Research, msal@hcmr.gr
Alexandros Frantzis, Pelagos Cetacean Research Institute, afrantzis@otenet.gr
Jacob Fric, Hellenic Ornithological Society, jacobfric@ornithologiki.gr
Aliki Panou, Archipelagos - environment and development, aliki.panou@yahoo.gr

Acknowledgement and management of data quality assessment in environmental science for policy has become a key factor during the last decades.

In the framework of the MESMA (Monitoring and Evaluation of Spatially Managed Areas) FP7 project, an evaluative account of different qualitative aspects of the data (Funtowicz & Ravetz, 1990) describing the natural and socioeconomic components that occur in the region of the Greek case study (Inner Ionian Archipelago and the adjacent gulfs) has been attempted. This process can shed light on the uncertainty inherent in the data and should be considered particularly in cases that will be used in decision-making.

Data quality assessment

A qualitative assessment of semi-quantitative criteria based on a pedigree matrix that describes those aspects of data quality influencing the reliability of the overall result, has been developed as suggested by Pedersen Weidema & Wesnaes (1996). In our case, the specific pedigree matrix has been modified in order to help specify the reliability, completeness, temporal correlation, geographical correlation and data collection process quality of the ecosystem components under study (Fig. 1).

Each characteristic is divided into five quality levels with a score between 1 (high quality) and 5 (low quality). The “reliability” indicator relates to the sources, the acquisition methods and verification procedures used to obtain the data. The “completeness” indicator relates to the statistical properties of the data.

The “temporal correlation” indicator represents the time correlation between the year of study and the year the data were obtained. The “geographical correlation” indicator illustrates the geographical correlation between the defined area and the location of origin of the data used. Finally, the “data collection process quality” indicator refers to the collection process of the data.

Indicator Score	1	2	3	4	5
Reliability	Measured Data	Verified data partly based on assumptions	Non-verified data partly based on qualified estimates	Qualified estimate (e.g. by scientific expert)	Non-qualified estimate
Completeness	Representative data from all sites relevant for the study area considered over an adequate period to even out normal fluctuations	Representative data from >50% of the sites relevant for the study area considered over an adequate period to even out normal fluctuations	Representative data from only some sites (<50%) relevant for the study area considered OR >50% of sites but from shorter periods	Representative data from only one site relevant for the study area considered OR some sites but from shorter periods	Representa-tiveness unknown or data from a small number of sites AND/ OR from shorter periods
Temporal correlation	Less than 3 years of difference to year of study	Less than 6 years difference	Less than 10 years difference	Less than 15 years difference	Age of data unknown or more than 15 years of difference
Geographical correlation	Data from area under study	Average data from larger area in which the area under study is included	Data from area with similar environmental conditions	Data from area with slightly similar environmental conditions	Data from unknown area or area with very different environmental conditions
Data collection process quality	Data from targeted research conducted by the team involved in the case study	Data from targeted research conducted by other teams not involved in the case study	Data from targeted research conducted with different methodologies	Data from common research conducted with a standard methodology	Data from common research conducted with different methodologies

Fig. 1 - Modified pedigree matrix used for the assessment of the semi-quantitative indicators in our case study.

Results - Conclusions

The results of the data quality assessment indicated that the sampling effort was not evenly distributed throughout the study area; thus for many ecosystem components the accuracy of assessments could be hindered by lack of completeness in the data collection process. In terms of data reliability, specifically for small-scale fisheries, since, due to the low data availability, expert judgment was applied in the analysis process, further sources of bias were introduced. However, it should be noted that the geographical and temporal correlation of the overall ecosystem components exhibited high quality assurance (Issaris et al., 2012).

Finally, in an attempt to visualize the qualitative information gained from the implementation of the pedigree matrix, radar diagrams were produced (Fig. 2), presenting each indicator's score for every ecosystem component. Hence, certain ecosystem and socioeconomic components (coralligenous and deep corals, fan mussel, gold coral and small-scale fisheries) exhibited low quality assurance (high uncertainty) mainly due to the lack of completeness or the fragmented data collection process.

On the other hand, other ecosystem and socioeconomic components (loggerhead turtle nesting beaches, trawlers, purse seiners, industry, and tourism (blue flags and diving centres)) having higher data quality exhibited high quality assurance and therefore low uncertainty. The above should be also taken into account while designing appropriate monitoring programmes aiming to fill the identified gaps.

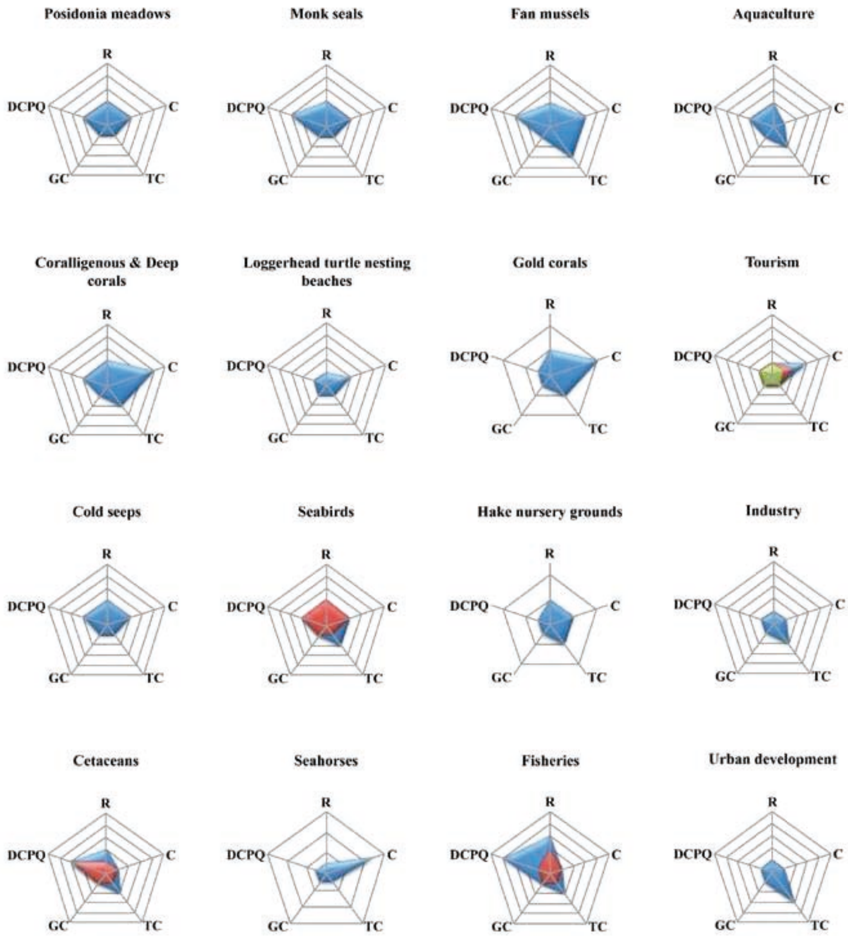


Fig. 2 - Radar diagrams for the ecosystem components, showing the scores for each indicator (R: Reliability, C: Completeness, TC: Temporal correlation, GC: Geographical correlation and DCPQ: Data collection process quality). For the ecosystem component: Tourism, the blue pentagon corresponds to marinas, the red pentagon to diving centres and the green to blue flags. For the ecosystem component: Seabirds, the red pentagon corresponds to the species C. diomedea and the blue to the species P. aristotelis desmarestii. For the ecosystem component: Cetaceans, the red pentagon corresponds to the species P. catodon, Z. cavirostris, T. truncatus, & S. coeruleoalba and the blue to the species D. delphis. For the ecosystem component: Fisheries, the blue pentagon corresponds to the coastal fleet and the red corresponds to the purse seiners & trawlers fleet.

Data validation in italian wave metric datasets

Marco Picone, ISPRA, marco.picone@isprambiente.it (Italy)
Gabriele Nardone, ISPRA, gabriele.nardone@isprambiente.it
Francesco Lagona, University of Roma Tre, francesco.lagona@uniroma3.it

Data from marine buoy networks are often exploited to validate remote sensing data or forecasting models.

The Italian data buoy network (RON, Fig. 1) acquire lots of oceanographic parameters. It consists of 15 oceanographic buoys, deployed along Italian coasts, on a seabed 100 meters deep. It collects directional sea wave data, meteorological data, sea surface temperature and in some cases quality parameters.

Fig. 1 - The Italian data buoy network.

Since 1989, a network has been in place along Italian coasts, but this was only able to measure directional sea waves and sea surface temperature. The new Italian Data Buoy Network implement the following measurements:

- Directional sea waves;
- High-accuracy Sea Surface Temperature (SST) and salinity;
- Meteorological parameters like wind speed and direction, air temperature and barometric pressure.

Marine dataset are disseminated of anomalous values, due to measurement errors, unmoorings, transmission errors, missing informations. Data quality control, or data validation, is essential whenever data are used by any individual or group other than the originators of the data. It consists in a sequence of operations that aspire to remove all incorrect measurements.

Standard validation procedure, such as threshold tests and spikes and outliers detection as descibed in the main international guidelines, can remove only measurements that appear to be clearly out of a resonable range and is often unable to process different time series that are strictly correlated, such as wave height, wave direction and wave period.

Validation can be however problematic in the presence of missing values in the network dataset. In order to account all these issues, a flexible cluster model can be used to validate date and simultaneously detect relevant regimes, such as sea regimes, wind regimes, atmospheric regimes, in italian seas. Traditionally, techniques of marine data clustering are based on distance-based methods. These methods have unknown statistical properties, precluding the possibility of formal inference on clustering results, such as the evaluation of the shape of the distribution of each cluster, the evaluation of the size of the confidence intervals and also the evaluation of any validation test on the data.

Latent-class models allow to cluster multivariate correlated data, such as wind and wave data, by defining the joint distribution of the data as a mixture of different distributions. Each observation can be validate according to the class membership and the confident inteval associated to the mixture model.

Latent class model for validation porpose

Even if in the literature only few works have been proposed in oceanographic topics, latent class (or mixture) approach is a very useful data mining method to obtain detailed information on the physical state of the ocean. It allows to cluster any kind of data, such as positive data (almost all oceanographic and meteorological measurement), circular data (tipically directional data, such as wind and wave direction), univariate or multivariate data, longitudinal data or spatial data.

Formally, given a latent (unobserved) variables that define the probability of a measured record yi to belong to class *1, 2,..., m,* the multivariate distribution of yi is given by

$$f(y_i) = \sum_{k=1}^{m} \pi_k f_k(y_i | \beta_k)$$

where β_k is a set of parameters of the mixture model, π_k is the probability that the observation i belongs to cluster k and m in the number of cluster. Typically, a maximum likelihood approach is used to estimate parameters and evaluate the class membership, through the implementation of an EM algorithm that iteratively converge to a solution.

As a results, these model are applied to multivariate time series of correlated measurements, such as wind speed and direction, wave height and direction, wave period. This procedure allows to cluster in relevant sea regimes all the observed events and define a confident interval for each observation profile. If the measurement does not belong to a 95% confident interval, the measurement is indicate as potentially incorrect. These anomalous values can be deleted and replaced by imputed values that correspond to the most probable value given the mixture model, obtained by using Monte Carlo methods.

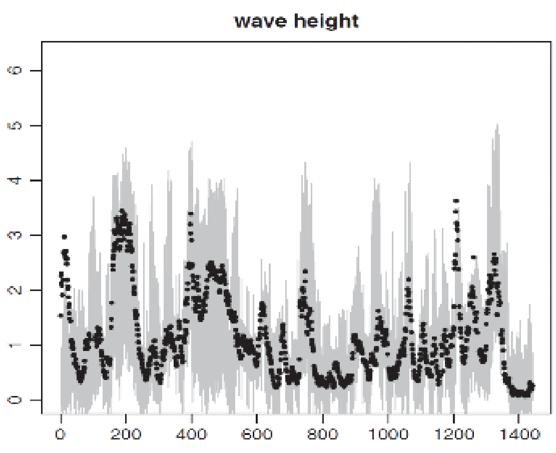


Fig. 2 - Confidence intervals for each observation.

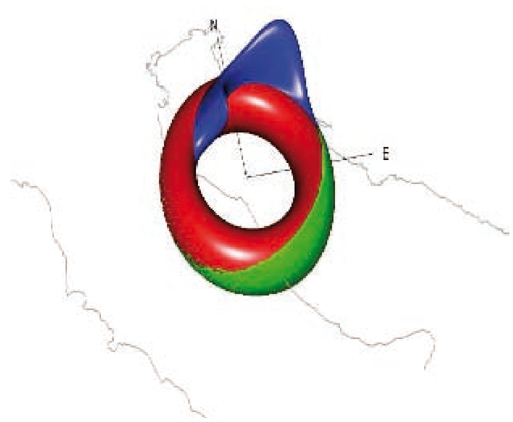


Fig. 3 - Joint distribution of wind-wave data.

R2R Science Eventlogger: Supporting Data Interoperability with Controlled Vocabularies

Cynthia CHANDLER, Woods Hole Oceanographic Institution, cchandler@whoi.edu (USA)
Andrew MAFFEI, Woods Hole Oceanographic Institution, amaffei@whoi.edu
Laura STOLP, Woods Hole Oceanographic Institution, amaffei@whoi.edu
Cynthia SELLERS, Woods Hole Oceanographic Institution, csellers@whoi.edu
Elizabeth COBURN, Woods Hole Oceanographic Institution, ecoburn@whoi.edu
Tim THIEL, Thiel Engineering, tthiel@thieleng.com
Robert Arko, Lamont-Doherty Earth Observatory, arko@ldeo.columbia.edu

Introduction

Methods used by researchers to record sampling events during a marine science research cruise to track and note when and where these occur are ubiquitous and vary widely. Handwritten notebooks, printed forms, watch-keeper logbooks, data-logging software, and customized software have all been employed. The quality of scientific results is affected by the consistency with which such events are recorded and integration of multi-cruise results is hampered because recording methods vary widely from cruise to cruise. The consistency and frequency with which events are entered is affected by the ease and speed with which a cruise participant can enter events during a cruise.

Following a year-long pilot project the US Rolling Deck to Repository (R2R) program developed an Eventlogger system for deployment on US research vessels. So far it has been deployed on 33 oceanographic research cruises. The initial principles followed and lessons learned during it's development and deployment are covered in this presentation. In addition we examine how post-cruise use of the Eventlogger data can support and enhance the discovery, analysis and integration of data held in oceanographic data repositories (e.g., the Biological and Chemical Oceanography Data Management Office (http://bco-dmo.org) funded by the US National Science Foundation).

Principles Followed and Lessons Learned

Principles followed and lessons learned during development and deployment of the R2R Eventlogger include the following: **(1)** work closely with and listen intently to scientists, repository owners, and other key stakeholders to understand how they will use the Eventlogger and match the software implementation accordingly; **(2)** use open-source software when available and appropriate; **(3)** auto-fill as much information as possible; **(4)** employ controlled vocabularies and data format standards to simplify steps required for event entry; **(5)** to the extent possible, map local controlled vocabulary terms to community terms, and ideally to globally unique identifiers for these terms in order to promote linked data; **(6)** plan to publish Eventlogger data in a standard way, to make data more discoverable and to enhance data access.

Eventlogger Development

Development of the R2R Eventlogger started during a year-long pilot project in which we examined the need for a digital event logging system, and designed several prototypes that incorporated the most useful functionality as articulated during the field testing. The current version of the R2R Eventlogger is based on the ELOG software package (http://midas.psi.ch/elog/) authored by Stefan Ritt. ELOG is an open-source package in the family of software known as 'weblogs'. Weblogs make it easy for people to use a Web browser client to enter information in a chronological fashion, in the form of short, time-stamped text messages ("entries") that are stored for later search, retrieval, display and export. The Eventlogger makes use of key data format standards and controlled vocabularies. The ELOG software has been enhanced by our team to provide a graphical user interface that enables the science party to customize a subset of information specific to their cruises. The custom configuration supports use of controlled vocabulary terms for cruise identifier, vessel name, start date, chief scientist name, cruise project name, cruise participant names, cruise participant affiliations, instrument (device) types, and instrument (device) actions.

An important design element of the Eventlogger system is the limited number of fields a user is required to complete when entering a new event. Pull-down lists of controlled vocabulary terms for instrument and person name and radio-button widgets for pertinent actions associated with these devices simplify event entry, improve accuracy and enable more efficient search. In addition, the Eventlogger is configured to monitor an NMEA stream on the shipboard network and insert latitude and longitude information automatically when the user presses the submit button for a new event.

Wherever possible, terms in local R2R controlled vocabulary lists are mapped to community-wide vocabulary terms. For example, SeaDataNet device category and type vocabularies (L05, L21 and L22) are used by the R2R program (Leadbetter, 2012). The R2R Eventlogger project has also begun to define and use a controlled term list for device actions (e.g., depart, return, deploy, recover, start, stop, startline, endline, release, etc.) that can be associated with each instrument. We have observed several significant advantages of using SeaDataNet community vocabularies; permanent identifiers assigned to each of the terms reduce ambiguity and an effective governance process enables term lists to be updated and extended. However, it is worth noting the importance of honoring 'local' device terms familiar to a research team, and including those terms in the Eventlogger configuration. During the pilot project, we observed the importance of familiar, local term use to the science team. The local terms, familiar to oceanographers, can be used during the cruise and then mapped to community-wide terms to enable resource discovery and use by a larger community in the future.

The Research Cruise Event Log: an Important Dataset

Using the R2R Eventlogger system, an event log created during a research cruise provides a chronological list of instrument/device related activities that occur during that cruise. These events can be exported as a comma-separated value file that includes event times in a standard time format and event-related fields populated with controlled vocabulary terms. Many of these terms are associated with unique identifiers that enable unambiguous interpretation.

After the cruise, the event log becomes a valuable record of all sampling activity during the cruise and can be made available as a data set. Including sampling event identifiers in discrete data sets reported by investigators, enables efficient and accurate integration of data sampled from the same sampling device (e.g., all measurements from analysis of Niskin bottle samples).

Although most work, to date, has been focused on ease-of-use for the science party and incorporation of controlled vocabularies into the application configuration process, we are now ready to investigate how event logs can provide links to much-needed temporal, spatial, and other contextual information for the data acquired during a cruise. A recent prototype effort of the R2R program has demonstrated, for example, how repositories with complementary data and metadata holdings for a growing number of oceanographic data types can be connected in an open, linked data network to provide oceanographic researchers with improved discovery and access capabilities (Arko et al. 2013).

References

R. Arko, C. Chandler, K.Stocks, S. Smith, PD Clark, A. Shepherd, C. Moore, S. Beaulieu. (2013) *"Rolling Deck to Repository (R2R): Collaborative Development of Linked Data for Oceanographic Research"*, Geophysical Research Abstracts, Vol. 15, EGU2013-9564, presented in Session *"ESSI1.2 Informatics in Oceanography and Ocean Science"* at EGU General Assembly 2013, Vienna, Austria.

Leadbetter, Adam (2012) *"NERC Vocabulary Server version 2.0"*. Revision 1.
http://www.bodc.ac.uk/products/web_services/vocab/documents/nvs2.0_documentation.pdf

SESSION 4

Services for Users and Education

- Historical evolution in data collection and management
- Tools for dissemination
- Test bed development for educational purposes

ORAL PRESENTATIONS

Surface Ocean CO2 Atlas (SOCAT) Version 2 – a showcase for transparent data management and international collaboration

Benjamin Pfeil, Geophysical Institute/University of Bergen, Bjerknes Center for Climate Research, PANGAEA, Benjamin.pfeil@gfi.uib.no (Norway)

Are Olsen, Geophysical Institute/University of Bergen, Bjerknes Center for Climate Research, are.olsen@gfi.uib.no (Norway)

Dorothee Bakker, University of East Anglia, D.Bakker@uea.ac.uk (UK)

Steven Hankin, Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration (NOAA/PMEL), Steven.C.Hankin@noaa.gov (USA)

Kevin O´Brien, University of Washington, Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration (NOAA/PMEL), kevin.m.o'brien@noaa.gov (USA)

Karl Smith, ISAO/University of Washington, National Oceanic and Atmospheric Administration (NOAA/PMEL), karl.smith@noaa.gov (USA)

Alex Kozyr, Carbon Dioxide Information Analysis Center (CDIAC), kozyra@ornl.gov (USA)

Christopher Sabine, National Oceanic and Atmospheric Administration (NOAA/PMEL), chris.sabine@noaa.gov (USA)

Maciej Telszewski, International Ocean Carbon Coordination Project (IOCCP), m.telszewski@ioccp.org (Poland)

Michael Diepenbroek, PANGAEA, MARUM / Universität Bremen, mdiepenbroek@pangaea.de (Germany)

The Surface Ocean CO2 Atlas (SOCAT) is a major global synthesis effort by the international marine carbon research community initiated by the International Ocean Carbon Coordination Project (IOCCP) in 2007. It aims to improve access to surface water carbon dioxide data measured on research vessels, voluntary observing ships and moored as well as drifting platforms by releasing regular updates of quality controlled and fully documented synthesis and gridded fCO2 products.

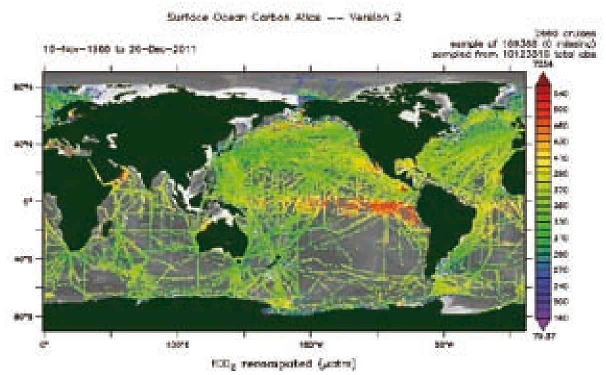


Fig. 1 - Distribution of fCO2 data using the LAS of SOCAT Version 2.

The second version of SOCAT will be published in June 2013 at the International Carbon Dioxide Conference in Beijing, China. It extends the first version SOCAT by 4 years and contains more than 10.1 million surface water fCO2 data from 2660 cruises between 1968 and 2011 making it the world's largest data collection of its kind.

Background

At the time when SOCAT was initiated in 2007 surface ocean CO2 data were archived in a wide range of formats and at numerous sites around the world, each with its own rules for access, and documentation of the data was frequently poor. This made it extremely time consuming if not impossible to generate comprehensive data synthesis products e.g. for large scale – long-term studies.

The international ocean carbon community under the lead of IOCCP decided to initiate SOCAT as a community driven effort to assemble, harmonize, document, quality control and re-compute surface water fugacity of carbon dioxide (fCO2) data into one open access database and to provide SOCAT gridded products of monthly fCO2 means on a 1° x 1° grid with no temporal or spatial interpolation. It took many years of hard and dedicated work by marine carbon scientists and data managers around the world to assemble and quality control the first two versions of SOCAT.

Data access and international collaboration

SOCAT is available through three access nodes demonstrating the international collaboration between the USA and Europe: PANGAEA (Data Publisher for Earth & Environmental Science), CDIAC (Carbon Dioxide Information Analysis Center) and the SOCAT Live Access Server (LAS) located at NOAA/PMEL (Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration).

The individual cruise files and the SOCAT synthesis product of the observational data are archived at PANGAEA and the SOCAT gridded products and concatenated files of the observational data of SOCAT are archived at CDIAC.

In addition are the SOCAT synthesis and gridded products accessible via the SOCAT LAS, a sophisticated, online data visualization and analysis tool.

The SOCAT website (www.socat.info) provides access to the nodes mentioned above.

In addition SOCAT is streamlined to the NODC (National Oceanographic Data Center), ICSU-WDS (International Council for Science World Data System), GEO Portal (Group on Earth Observation) and NASA GCMD (Global Change Master Directory).

Transparency

The methods and scripts for re-calculating CO2 within SOCAT are fully transparent, documented and open accessible. All archived SOCAT synthesis product and individual cruises are citable using DOIs (Digital Object Identifier). Cruises in SOCAT have been assigned a standardized cruise identifier called Expocode (consisting of the ICES/NODC platform code and the start day of the cruise using ISO 8601, YYYYMMDD) and every individual cruise was archived with detailed. Every data point in the SOCAT synthesis files has a link to its cruise of origin via the DOI-number making it possible to get instant and detailed access to input data and data treatment.

SOCAT is a perfect showcase for a community driven data synthesis, international collaboration, transparent procedures for data treatment and on how to publish global distributed data as a data product while giving credit to the data originator, scientists and data centers involved.

European Atlas of the Seas: Marine & Maritime Information for Science and Society

Vittorio Barale, Joint Research Centre, European Commission, vittorio.barale@jrc.ec.europa.eu (Italy)

The European Atlas of the Seas (MARATLAS) is a public, web-based, graphical information system (http://ec.europa.eu/maritimeatlas) aimed at practitioners, policy-makers, managers, teachers, students, and anyone interested to learn more about Europe's seas, their environment, as well as related human activities and policies in the European Union (EU). A preliminary version – a communication tool already foreseen by the EU Integrated Maritime Policy (IMP) in 2007 – is currently available in English, German and French, and presents maps with data obtained primarily from the European Commission (EC) and its Agencies.

The Atlas shall be further developed in the period 2013-2014, in order to provide access to more spatial and/or factual information on the European Seas, and to provide a suite of ready-made instruments for first-cut assessments of coastal and marine issues. In order to do so, MARATLAS will be hosted directly by the EC Joint Research Centre (JRC), to facilitate the development of new services and features, as well as the interaction with other information access tools currently available or planned.

The future Atlas will be compatible with, though separate from, the wider family of JRC information systems (e.g. the Global Marine Information System, GMIS, http://gmis.jrc.ec.europa.eu/), and with its spatial data product generation mechanisms, adding also data analysis and interpretation capabilities, for the combination of basic information into coupled ecological and socio-economic indicators. Further, it will be interfaced to other databases, including both the JRC in-house archives and relevant external sources, to ensure a reliable, up-to-date and authoritative window on all European Seas, using data and indicators taken from sources such as Eurostat and/or other Commission Services, or the European Environment Agency.

The resulting MARATLAS shall be equipped with map services and basic data analysis capability, such as product-to-product correlations, or time series visualisation, capable of providing a quick, but detailed and reliable, reference source of maps, facts and figures, to both the general public and a selected professional audience – including practitioners either holding policy-oriented or managerial positions, in both the public and private sector, or responsible for the planning and execution of projects and programs, or involved with educational institutions, non-governmental organisations and international bodies.

The ICES Data Type Guidelines and their place and role in the profusion of guidelines, manuals, standards, cookbooks and best practice

Lesley Rickards, BODC, ljr@bodc.ac.uk (UK)
Taco de Bruin, NIOZ, taco.de.bruin@nioz.nl
Neil Holdsworth, ICES, neilh@ices.dk
Hjalte Parner, ICES, hjalte@ices.dk
Sjur Ringheim Lid, IMR, Sjur.Ringheim.Lid@imr.no
Garry Dawson, UK Hydrographic Office, garry.dawson@ukho.gov.uk

Data held at the ICES Data Centre are used in various assessments for expert groups and regional sea conventions. In order to ensure comparable data with high quality, guidelines have been developed and adopted.

In particular, the ICES operational group, Data and Information (DIG), and its predecessors, have developed guidelines to assist those involved in the collection, processing, quality control and exchange of various types of (mainly) physical oceanographic data, for example, Moored Current Meter, Shipborne ADCP, Seasoar, Chlorophyll and Nutrient data. These guidelines have been adopted by the ICES Data Centre and are recommended to the ICES Community.

Each guideline addresses the data and metadata requirements of a specific data type. They cover three main areas:

- What the data collector should provide to the data centre (e.g. collection information, processing, etc)
- How the data centre handles data supplied (e.g. value added, quality control, etc)
- What the data centre can provide in terms of data, referral services and expertise back to the data collector

This presentation examines the place of the ICES Data Guidelines in amongst the many other guidelines, manuals, standards, cookbooks and best practice documents that exist today. When some of these guidelines were first developed (almost 30 years ago), there were few others available. Now there are many guidelines, standards, cookbooks and best practices documented and available from a variety of sources.

We have investigated current knowledge and usage of the present set of ICES Data Type guidelines and found some positives (they are known, used, referenced) and some negatives (not known, etc.). Before reviewing and updating the guidelines, it was agreed that the first priority was a communication strategy.

Thus the aim of this presentation is twofold: firstly to raise the profile of the ICES Data Type Guidelines, and secondly a 'call to action' to bring together the different bodies (including IODE/JCOMM, SeaDataNet, national activities) to make a more coordinated approach to linking the guidelines, and to map out the complex landscape.

There are a number of issues which the presentation will discuss:

- Each organisation, programme and project is trying to be "the" place to go – how can better signposting be provided so users can find what they need?
- Related to the above point – are good links in place both from the ICES Data Guidelines to more detailed manuals and to the ICES Data Guidelines from elsewhere?
- Are there overlaps or duplications between the ICES guidelines and other similar activities?
- Guidance is needed on which guidelines, manuals, etc., are best for which purpose.
- Provision of information describing the benefits to users.

Geoinformation Systems, Databases, and Climatic Atlases of the Sea of Azov, Black Sea, and the Caspian Sea

Gennady Matishov, Southern Scientific Centre of the Russian Academy of Sciences, matishov_ssc-ras@ssc-ras.ru (Russia)

Sergey Berdnikov, Southern Scientific Centre of the Russian Academy of Sciences, berdnikov@ssc-ras.ru

Valery Kulygin, Institute of Arid Zones of the Southern Scientific Centre of the Russian Academy of Sciences, kulygin@ssc-ras.ru

Natalia Yaitskaya, Institute of Arid Zones of the Southern Scientific Centre of the Russian Academy of Sciences, yaitskayan@gmail.com

The studies of long-term climatic trends in marine ecosystems, as a rule, are based on traditional oceanological approaches, from the one hand, and mathematic modeling, from the other. Both directions require available primary blocks of checked/controlled observations' data, united into megadatabases. The number of such databases in the contemporary informational space increases, all data are almost always available in operational regime. Research teams from various countries are involved in searches for historical primary data and their digitizing. At the same time, there are negative effects: the increase of databases' number increases the probability of obtaining the same blocks of data, that is the duplicates; approaches to data quality controls are not always agreed upon. The situation with already published maps and charts remains difficult. It is possible to find old bathymetric charts in a digital form but parameters' distribution charts frequently are not considered by the specialists as an information source.

An integrated approach is applied to the southern seas of Russia – the Sea of Azov, Black Sea, and the Caspian Sea – to rescue and preserve information (Climatic Atlas of the Sea of Azov 2006; Climatic Atlas of the Sea of Azov 2008). Within a new Atlas of Climatic Changes of Large Marine Ecosystems of the Southern Region of Russia (2013) the following has been developed and completed: the archives of primary oceanographic information for the period of 1884 to 2011 (167 505 sea stations for three seas, corresponding to the world standards of data quality control); cartographic database (346 thematic layers); instruments for the analysis of spatial-temporal information, which allow carrying out multi-dimensional analysis of spatial-temporal variability of oceanographic parameters of the seas. Geoinformational system is the basis for data storage.

Oceanographic database

Oceanographic database (ODB), apart from the main components, contains hydrological-hydrochemical and meteorological information, in total more than 18 different types of information and more than 100 parameters. When processing the primary information, considerable attention is paid to metainformation. The database of the Atlas envisages the storage of data on the observations' conditions and sampling (instruments, methods), as well as specialists involved in sampling and samples' processing. Each type of metainformation has got an appropriate code table (code table). The development of code tables was based on the code tables of the World Meteorological Organization (WMO) and the US National Oceanic and Atmospheric Administration (NOAA). Data quality control is an important and integrated part of the database development and development of data nets. Quality control procedure is an iterative multi-level process composed of: a) an automated stage of objective quality control of data; b) a stage of subjective analysis, carried out by a specialist. Quality control includes: the search and consideration of possible errors and uncertainties in data; quality flags for the analysed data at the level of station and the level of some values; data formats' control; check of spatial-temporal location; check of vertical structure of measurements; searches for duplicates; presence of added oceanographic characteristics in the code tables.

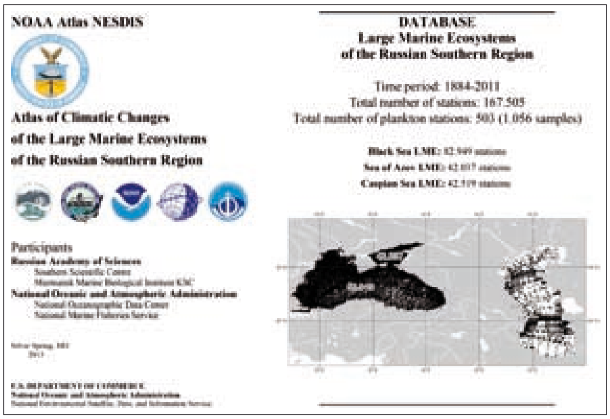


Fig. 1 - Atlas of Climatic Changes 2013.
Large Marine Ecosystems
of the Southern Russian Region.

Cartographic database

When developing a cartographic database in GIS, the synergy of two models of data, grounding the presentation and introduction of spatial and semantic information in GIS, georelational model, was used. To store information a base of geodata was chosen, containing: vector data (spatial classes and sets of data); raster data (sets of rasters and catalogues of rasters), represented by space images, topographic and thematic raster charts with spatial reference, results of spatial analysis in the way of raster surfaces; non-spatial tables, containing primary data and calculations' results, applying mathematic models.

GIS includes five thematic subsystems: 1) «Oceanographic Database», containing spatially registered/referred information from ODB; 2) «Bathymetry», containing both historical reviews of sea charts and most recent updates; 3) «Hydrology and Hydrochemistry» and 4) «Meteorology» developed for collecting, storing, and analysis of thematic cartographic information about the southern seas, published during the historical period, digitized distribution fields for hydrological and hydrochemical parameters. In total, more than 300 charts and schemes; 5) «Model evaluations», containing a set of subprogrammes to construct spatial distribution of parameters. Meteorological databases are separate modules of the considered GIS. They accumulate information from various open sources and data on reanalysis (the meteosites of European territory of Russia and six neighbouring countries). Despite the fact that the system has been developed with the help of ArcGIS software, information can be freely and easily converted into other formats and used together with other programmes.

POSTERS

Caribbean Marine Atlas

Ann-Marie Eversley, Environmental Protection Department, aeversley@epd.gov.bb (Barbados)

In June 2007, the potential benefits of an online atlas of coastal and marine data as a tool to improve regional environmental management were clearly demonstrated by the publication of the then prototype African Marine Atlas (AMA). This prompted the call to create a geographic data sharing platform for the Caribbean region modeled on the AMA. The development of the Caribbean Marine Atlas (CMA) project started with a stakeholder meeting in October 2007 in Bridgetown, Barbados. This was a regional information gathering workshop with 7 participating Caribbean countries (Barbados, Cuba, Grenada, Jamaica, St. Lucia, Trinidad and Tobago and Turk and Caicos). The workshop goals were: to inform the participating countries of the potential benefits of a CMA; to identify current national coastal zone management arrangements; to identify national and regional coastal and marine issues that could be the focus of the CMA; to identify the national resource requirements of the participating countries to enable full participation in a CMA Pilot Project; and to prepare a draft work plan of a CMA Pilot Project, for submission to, and approval by the respective national governments.

The coastal zone management issues identified from this workshop were found to be priorities across the region such as overexploitation, natural hazards, land based sources of marine pollution (LBS), and coral reef, seagrass, mangrove, and beach degradation. In terms of availability of national data, there were universal challenges to accessing data held by sister institutions. In addition, a survey data management policies indicated that there was little or no data quality control (geographic, range analysis), no established metadata schemes, and a lack of institutional resources such as personnel, training and equipment.

The workshop also identified several types of training necessary to increase the capacity of regional data managers and fulfill the requirements of the CMA project. These included ocean data management, data mining and web atlas creation. These workshops resulted in the advanced training of eleven regional data managers improved communication among Caribbean marine/coastal management agencies and the transfer of knowledge necessary to successfully maintain an online marine atlas. The creation of the CMA prototype atlas.caribbeanmarineatlas.org was another achievement.

The second phase of the CMA project received funding from the Government of Flanders in 2008 and featured the designation of a Regional Coordinator, Mr. Ramon Roach from Barbados. The second phase included the publishing of the CMA prototype online (November 2009), the addition of querying and download capability, expansion of member countries, improvements to the atlas interface, and additions to the functionally of the application.

The CMA is a multi-component platform (open-source and web-based) consisting of a web map server, a metadata catalogue and the online mapping application (Fig. 1). The interactive mapping application consists of a collection of GIS data layers (both vector and raster) representing various data types. The system also features advanced data control such as layer lists, feature selection, results tables, metadata viewing and links to websites, documents and auxiliary data. Some of these data controls are still in the developmental stages.

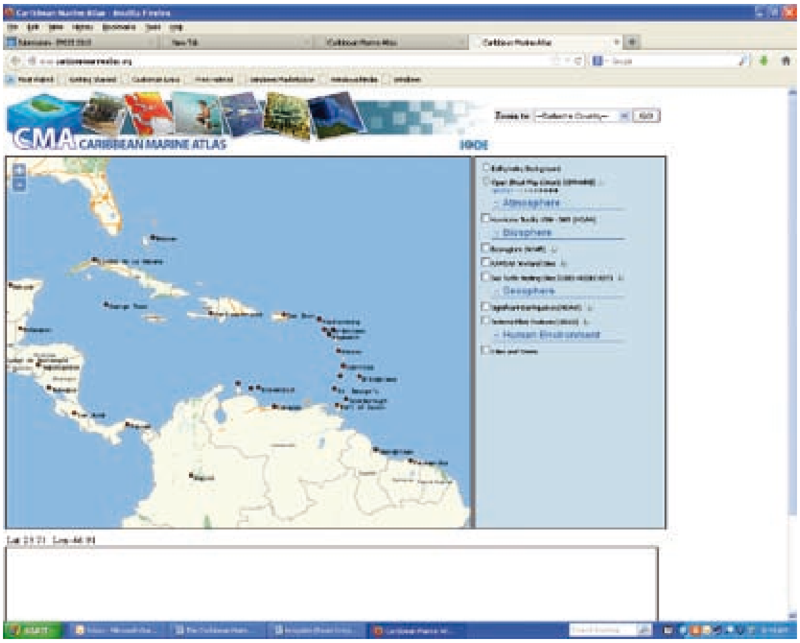


Fig. 1 - Caribbean Marine Atlas.

The thematic areas of the CMA are as follows:

• Geosphere • Natural hazards • Soil types • Land use	• Hydrosphere • Bathymetry • Physical oceanography • Chemical oceanography
• Biosphere • Habitats • Marine flora and fauna • Protected areas	• Atmosphere • Climate • Wather
• Human Environment • Settlements • Infrastructure • Tourism	

The design and implementation of national atlases is also part of the second phase of the CMA project. The national atlases are built on the same framework as the CMA, and share its functionality. The countries with functional prototype atlases are Barbados, Trinidad and Tobago, Jamaica, Cuba and Dominica, with Turks and Caicos expected to publish their atlas in late 2013.

Rescue of historical UK sea level charts and ledgers – creation of Open Educational Resources (OERs)

Lesley Rickards, British Oceanographic Data Centre, ljr@bodc.ac.uk (UK)
Elizabeth Bradshaw, British Oceanographic Data Centre, elizb@bodc.ac.uk
Jenny Andrew, British Oceanographic Data Centre, jedr@bodc.ac.uk

The British Oceanographic Data Centre (BODC) has, in collaboration with the University Of Liverpool School Of Engineering, produced Open Educational Resources (OERs) for historical sea level data from tide gauges. Recent Ocean Science graduates from the University have developed these resources using material we have digitised and/or scanned. These resources are for use in university Ocean Science courses, but are also of interest to the wider community including the public.

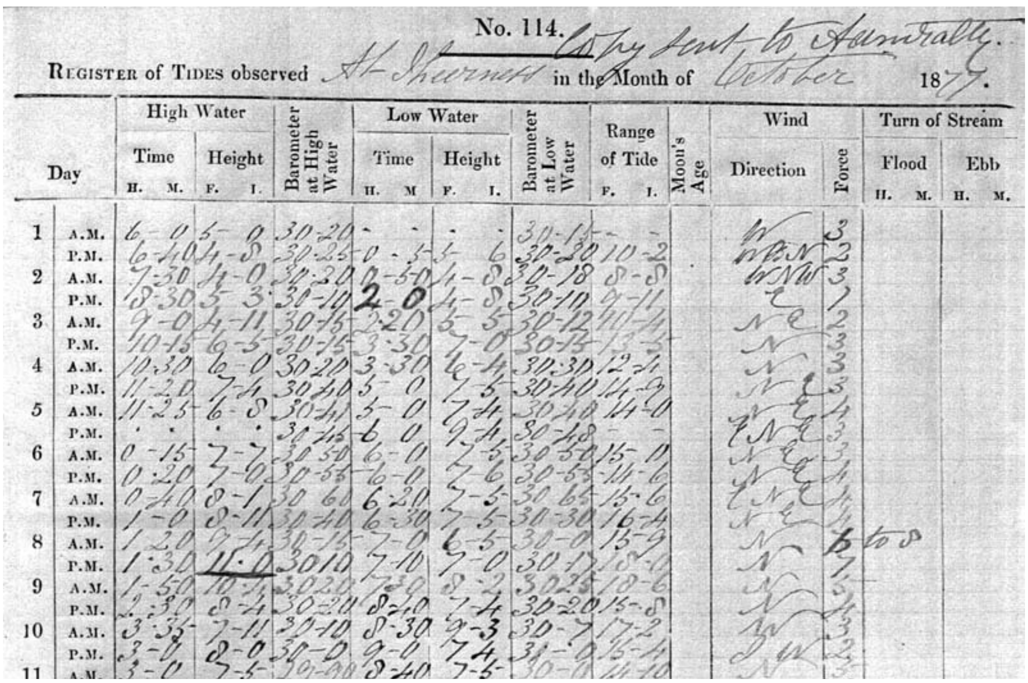


Fig. 1 - Scanned ledger for Sheerness, 1877. Underlined is the exceptionally high water on the 8th October, when there was a large flood on the East coast of the UK.

The Resources draw upon material digitised and scanned under the JISC Content programme 2011-2013, (Strand B: Mass Digitisation). BODC received funding for the "Rescue of historical UK sea level charts and ledgers" project', digitising ~160 site years of chart data from 22 sites around the UK coast, and scanning ledgers totalling ~500 years from 14 sites. The aims of the project were to increase the availability of long-term historical sea level time series, make the data electronically available to the wider community and raise awareness of the data's existence.

These Resources may answer specific questions, such as "what do we know about storm surges?" but they also help explain basic mathematical/physical principles.

The OERs will be Creative Commons 0 (CC0) to encourage distribution and reuse. They will also be deposited in JORUM, the JISC online repository for teaching and learning materials for the Higher Education/Further Education community. We will also investigate assigning paradata to the resources, to enable users to identify and evaluate useful material.

Another aim is to promote the use of BODC's data holdings, including the BODC Historical UK Tide Gauge Data delivery website:

- http://www.bodc.ac.uk/data/online_delivery/historical_uk_tide_gauge_data/

This tool allows users to search BODC's historical scanned charts and ledgers archive using a number of criteria, such as location, time period and parameters measured. After searching, users can examine the scanned images using a "magnifying glass" tool, with the option to login and download the high resolution original.

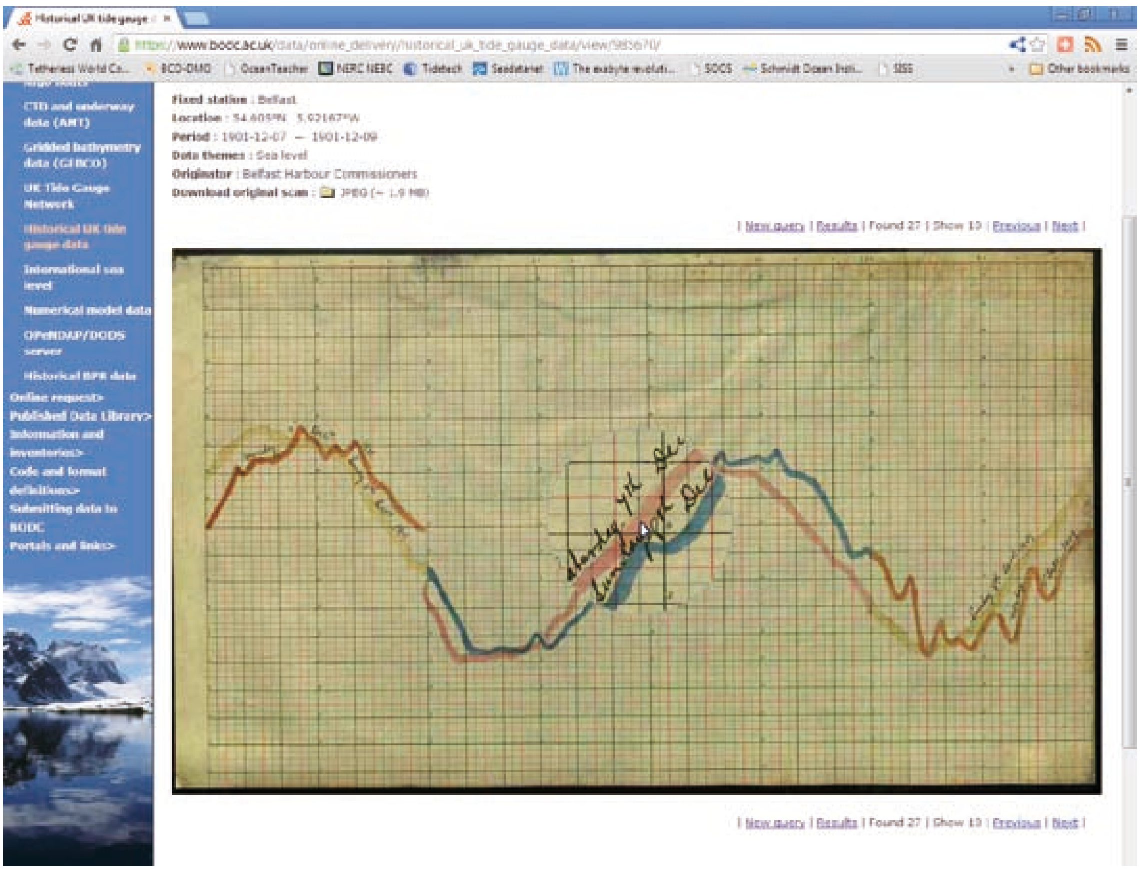


Fig. 2 - Screenshot of Historical UK Tide Gauge Data website, showing "magnifying glass" tool in operation.

Collaborative Knowledge Building

Giuseppe Manzella, ENEA, giuseppe.manzella@enea.it

and **OTTIMA** (Operational oceanography and Technologies for Maritime Security) Training Course:

Abbamondi Chiara, Andreotti Valeria, Angelica Carmelo, Buonocore Cira, Celentano Andrea, Ciancimino Sarah, Corrado Lorena, D'Angelo Alessandra, De Lauro Marinella, Di Luccio Diana, Ferrigno Federica, Frunzo Angela, Laurenza Tiziana, Leone Maddalena, Mirrione Alessandra, Pessini Federica, Riefolo Luigia, Roviello Valentina, Ruggieri Stefano, Scotto d'Antuono Annalina, Serrentino Fabrizio, Sgrosso Andrea, Sposato Marina, Tobia AngelaMaria, Tedesco Costanza

The OTTIMA training course is attended by post-graduate students from different disciplines. The aspects on Data Management and Information Systems are discussed by in the class as a collaborative process incorporating two main phases that constitute a cycle of (1) personal and (2) collaborative knowledge-building. Data management and information systems are presented in their general concepts and students elaborate them during ad hoc exercises, where exchange of personal thinking is occurring. The personal beliefs are successively articulated as a contribution to a collaborative knowledge-building process, that may lead through collaborative discourse toward the enriched self-understanding of a research community.

The cyclical character of this process allows increasingly complex questions to be posed on the basis of more and more sophisticated understanding. It explicitly considers the relationship of processes associated with individual minds to those considered to be socio-cultural. The collaborative knowledge-building starts from individual learning of several individuals. This happens when someone's personal belief is articulated in words and this public statement is taken up in a social setting and discussed from the multiple perspectives of several participants. The original statements are thereby articulated into a more refined and extensive discussion of the topic, subject to conflicting interpretations. The discussion consists of arguments providing rationale for different points of view. The interchange may gradually converge on a shared understanding resulting from a clarification of differences in interpretation and terminology.

The application has been done on these particular aspects:

- quality control
- data model
- metadata model
- needs of an information system

Each students has presented his own perspective and a synthesis has been proposed and discussed. Comparisons with models provided by European (e.g. SeaDataNet) and US projects (e.g. NODC, UNIDATA) has been performed.

The underlying the theory of learning defined by all students a social epistemology. Individuals generate personal beliefs from their own perspectives, but they do so on the basis of personal discipline based knowledge. Successively share language and external representations. Further, these beliefs become knowledge through communication, discussion, clarification: knowledge is a socially mediated product.

OceanTeacher goes global: The OceanTeacher Global Classroom

Claudia Delgado, IODE, claudia.delgado@iode.org (Belgium)
Peter Pissierssens, UNESCO, peter.pissierssens@unesco.org

One of the major objectives of the IODE Programme (http://www.iode.org) is to assist IOC Member States to acquire the necessary capacity to manage marine data and information, and thus become full and equitable partners in IODE. With the establishment of the IOC Project Office for IODE in April 2005 in Oostende (Belgium), IODE has been able, for the first time, to organize a sustained training programme for both data and information management, with OceanTeacher as IODE's capacity development platform and tool.

The OceanTeacher e-learning platform (http://www.oceanteacher.org) has two main components: the OceanTeacher Digital Library (a collection of knowledge and resources) and the OceanTeacher OpenCourseWare (a collection of course outlines and courseware). In addition the system includes video recordings of courses that are included in the OpenCourseWare pages (http://vimeo.com/iode). OceanTeacher has become a comprehensive web-based training system structured in a way to support classroom training (face-to-face), blended training (combining classroom and distance learning), online tutoring and online self-learning. When the Digital Library and the OpenCourseWare are used together for a training event this is called an OceanTeacher Classroom.

The objective of the OceanTeacher Academy is to establish a training environment that provides an annual teaching programme of courses related to oceanographic data and information management and related disciplines that will contribute to the sustainable management of oceans and coastal areas. The OceanTeacher Training Academy thus underpins all Ocean Data and Information Networks (ODINs) developed under IOC/IODE activities. The development of the OceanTeacher Academy started in 2005 with the establishment of the IOC Project Office for IODE in Oostende, Belgium.

From a training system that focused entirely on oceanographic data and information management, OceanTeacher has gradually developed into a multi-purpose training system focusing on several IOC (e.g. HAB, ICAM, MSP) as well as non-IOC (e.g. SeaDataNet, EUMETSAT) programmes. Ultimately OceanTeacher should be able to provide a multi-disciplinary training platform. Cooperation between SeaDataNet and IODE was established in 2006 within the framework of the first SeaDataNet project. Since then 5 training courses were organised at the IOC Project Office for IODE, involving 222 participants. The OceanTeacher e-learning platform is used for the SeaDataNet courses, and ensures that the training course contents, including training videos, are kept freely available in the long therm.

In seven years the OceanTeacher Academy has organized over 50 courses for over 1200 students from 120 countries and taught by 20 lecturers. One of the main conclusions is that we cannot provide training to all those who require it from just one location. In addition, the continuous increase in airfares makes the cost per student quite high. It was therefore decided to, as from 2012, further develop the distance learning capacities of OceanTeacher but, of more importance, to move towards a distributed architecture of regional training facilities. Such a distributed network of training facilities will allow a better focus on local and regional needs, as well addressing the language issue, which has been a recurrent comment from students.

In recent years video conferencing technology has developed remarkably, with increased quality of the communication provided. Videoconferencing systems now provide high quality, interactive visual solutions that improve all facets of education, giving teachers the tools to engage students with more interactive and innovative educational experiences. Ultimately, the technology pays for itself by reducing travel and accommodation cost and increasing the target audience. Video conferencing is therefore being used widely as a cost-effective alternative to expensive and environment unfriendly air travel. Especially large corporations with a global presence have been using this technology intensively. Top Universities have also joined the trend and provide online courses for students dispersed across the planet. Education without borders has become a reality. In developing countries the technology has been used far less due to the cost of necessary equipment and the lack of adequate Internet bandwidth. However these limiting factors are slowly being solved with web-based videoconference applications and the constant increase in available bandwidth.

During its 22nd Session in March 2013, the IODE Committee endorsed the OceanTeacher Global Classroom concept. In the meantime, the OceanTeacher Global-Classroom Pilot Project has been approved for funding; its implementation will start in late 2013, in order to fine-tune the methodology and technology. Discussions have already started with Member States in other regions - namely India, Kenya, Colombia and Spain - to establish similar centres. A successful experiment took place in March 2012, when 2 groups of students attended the same course: one group at the IODE Project Office in Oostende, Belgium, and another in INCOIS (International Training Centre for Operational Oceanography) in Hyderabad, India.

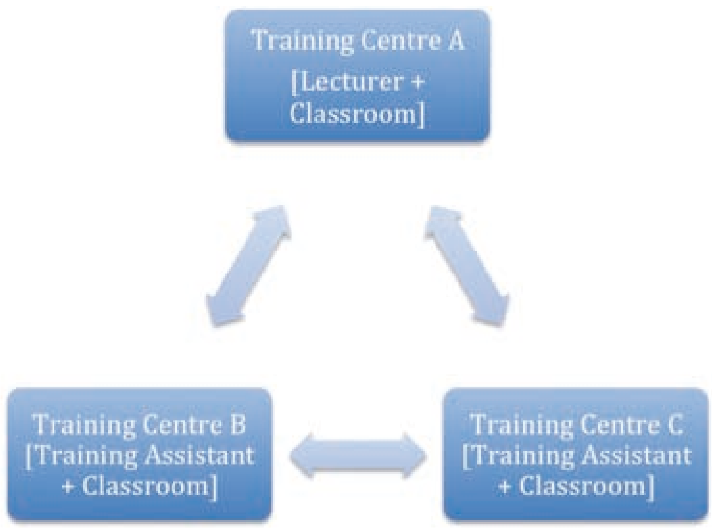


Fig. 1 - Schematic diagram of the OceanTeacher Global Classroom architecture.

The real innovation of the new OceanTeacher Global Classroom model is that we will blend traditional classroom-based training with distance learning, by using videostreaming technology to enable multi-site classrooms: a course taught in Oostende can be attended by students in Oostende but also in other Training Centres (e.g. Hyderabad or Mombasa). Similarly, a course taught in Mombasa can be attended also by students in Oostende and Hyderabad. Although the goal is to promote a more local/regional focus we should not forget the importance of bringing together students from different regions to foster south-south and north-south cooperation. By using advanced video conferencing technology we will be able to broadcast courses taught live in one region, to one or more other regional centres or invite individual lecturers to teach a class from their own university or even home. The system will also allow interaction between lecturer(s) and students and training assistant(s) in the other regional centres will be able to provide assistance with practical exercises. Additional benefits include a more efficient time-use for lecturers, less tiring travel for students and lecturers and of course, following from this, reduced carbon emissions. In additional benefit is that more attention can be given to local issues and hands-on support can be provided in local languages (while so far all courses were taught in English).

The expected results of this GlobalClassroom are to: 1) expand the learning experience from OceanTeacher's current 20 students per course to hundred(s); 2) combine focus on local/national/regional priorities while keeping a global perspective; 3) achieve true lecturer flexibility, i.e. make the best lecturers available anywhere in the world without the need for expensive, tiresome and time-consuming travel; 4) link classrooms in geographically dispersed locations; and 5) obtain an improved long-term impact of IODE-OTA capacity building efforts, including Continuous Professional Development (annual updating of capacity).

The OceanTeacher Global Classroom will therefore allow the building of a global course programme enabling lecturers to teach and students to attend courses from their usual place of study or work. Furthermore, using the same baseline curriculum in OceanTeacher we will be able to reach more students in more countries. In addition, through cooperation with local Universities accredited certificates will be issued.

Joint Nature Conservation Committee Marine Protected Areas Interactive Online Mapper

Dan Bayley, Joint Nature Conservation Committee, dan.bayley@jncc.gov.uk (UK)
Amy Ridgeway, Joint Nature Conservation Committee, amy.ridgeway@jncc.gov.uk (UK)

The Joint Nature Conservation Committee and the four country conservation agencies in the United Kingdom have an ongoing programme of work to identify Marine Protected Areas (MPAs) to contribute towards an ecologically coherent network. This JNCC interactive map (http://jncc.defra.gov.uk/page-5201) contains information on Marine Protected Areas in the UK. At the moment there are 107 Special Areas of Conservation (SACs) with marine components, 107 Special Protection Areas (SPAs) with marine components, one Marine Conservation Zone and two Marine Nature Reserves. Together these protect 8.4% of UK seas.

Dissemination of information on MPAs is a statutory requirement. This highlights the importance of communicating effectively to not only technical specialists and stakeholders but also the general public. The JNCC interactive web-based Geographical Information System (webGIS) is an important tool which allows us to reach a varied audience.

The Interactive Map collates and disseminates information on each UK Marine Protected Area. Like Google maps and Google Earth, you are able to zoom and pan to the area of interest. Users can then right-click on sites of interest to them to bring up more information about that specific MPA. Users are also able to search and zoom to habitat or species features of interest, to create bespoke pdf maps for printing (Fig. 1). Users can also download several of the datasets and import all of them via a Web Mapping Service (WMS) into their own desktop GIS systems. The Interactive Map is fully intergrated with the JNCC website allowing users to move seamlessly from the summary information provided in the interactive mapping environment, to more detailed relevant information in other sections of the JNCC website.

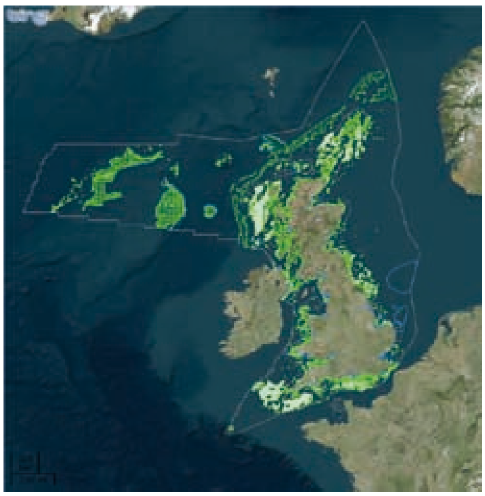


Fig. 1 - UK SACs, with areas of reef illustrated using the interactive map.

A future development will allow users to view a subset of sites by selecting the criteria they are interested in, for example, to view only sites that protect reef habitats.

This presentation will highlight the importance of communicating information on Marine Protected Areas and demonstrate the degree of functionality that enables the interactive mapper to be the ideal tool for distribution of the varied pieces of information on MPAs.

Back to the sea of the past to evaluate the status of today: Integration and valorisation of historical data

Ruth Lagring, Karien De Cauwer, Mia Devolder, Serge Scory

Royal Belgian Institute of Natural Sciences
Management Unit of the North Sea Mathematical Models (RBINS-MUMM),

ruth.lagring@mumm.ac.be (Belgium)

Historical marine datasets are an important scientific heritage. Not only do they allow us to give an idea on the status of the environment at the time of sampling, more importantly, long-term dataseries enable the assessment of global climate change providing insight in the evolution over the years. In the frame of the EU Marine Strategy Framework Directive, the 'Good Environmental Status', for which descriptors like eutrophication and contaminants were defined, has to be determined. To increase scientific knowledge of these descriptors, we have to go back to the data compiled in the past. However, long-term dataseries are not readily available and historical data, which are at risk of being lost, need to be recovered, integrated and quality checked before they can be compared with current data. In project PMPZ-DBII (2009-2012), data of 1970-1982 from the Belgian Continental Shelf have been recovered and integrated in the existing database and studied to assess usefullness for long-term studies.

In Belgium, modern oceanography with systematic measurement campaigns for assessing the quality of the marine environment, started in 1970 with the 'Projet Mer/Projekt Zee' or PMPZ (1970-1976). All compartments of the ecosystem were studied with the colaboration of over 200 scientists to explore the Belgian Continental Shelf (BCS). This initial program was followed by other research actions, programs and monitoring campaigns, resulting in a vast quantity of scientific data of the BCS since 1970. Techniques for digitization were not as advanced as today and data that were obtained via measurement campaigns aboard ships were not centralised in adequate databases. In the frame of PMPZ-DBII, the objective was to centralise and integrate the PMPZ-data with current data originating from monitoring and research programs, and to make them publically available. The IDOD database or "Integrated and Dynamical Oceanographic Data management system" was specifically developed to deal with marine environmental data and is as such suitable for these long-term datasets.

Original data and metadata from the years 1970 until 1982 could be recovered from technical reports, cruise reports and grey literature, digitized and imported. The reconstruction of the extensive historical research actions was very complex and many obstacles during the import process, which will be further described, needed to be overcome, like:

- A high number and high variety of data was dealt with which were scattered over the country and mostly in paper format.

- Final reports contain aggregated values and maps which are difficult to import in a database. Original values, which could be imported, were found in the technical reports.
- Available information was limited and metadata was often missing, making it very time consuming to recover.
- Information needed to be cross-linked to reconstruct the complete data and their metadata, like values, samples, campaigns, positions, dates, depth, sampling gears, analyses methods, data-originators, platforms, etc.

The IDOD database has been extended with data from over 120 campaigns and about 1600 sampling events, like water values for pigments (5864), nutrients (3298), heavy metals (2302), physical measurements (3295) and sediment values for heavy metals (1649). As a result, today the IDOD database has dataseries available for over 40 years, accessible via the BMDC website (http://www.mumm.ac.be/datacentre/).

Finally, as a first attempt to valorise the PMPZ-data, the evolution of heavy metals and pesticides in the sediment of the BCS from 1970 until today has been studied. During this assessment, a lot of limiting factors had to be dealt with and it was found that important meta information was lacking, which could be retrieved by consulting the data-originator. For example, the sediment fraction analyzed was not clearly described in the reports. After cross-checking data of different tables and testing the possible hypotheses, the fraction could be derived and corrected. However, it appeared to be smaller than the fraction used in the current monitoring practices (63 μm). Recalculated concentrations of lindane and mercury, appeared to be higher than the recent concentrations. It was concluded that a thorough study on normalization of these historical data in the BCS is needed to obtain a reliable comparison with current data. The project has learned us to cope with characteristics of historical datasets and that, in order to valorise and use them for long-term assessments, specific additional efforts are required to quality check and intercalibrate the data. The results of this study will be presented.

Nordic Microalgae – an information system for aquatic microalgae

Bengt Karlson, SMHI, bengt.karlson@smhi.se, (Sweden)
Örjan Bäck, SMHI, orjan.back@smhi.se
Lars Johan Hansson, SMHI, lars.hansson@smhi.se
Arnold Andreasson, SMHI, arnold.andreasson@smhi.se
Marie Johansen, SMHI, marie.johansen@smhi.se
Malin Mohlin, SMHI, malin.mohlin@smhi.se
Ann-Turi Skjevik, SMHI, ann-turi.skjevik@smhi.se
Andreas Loo, SMHI, andreas.loo@smhi.se
Lotta Fyrberg, SMHI, lotta.fyrberg@smhi.se

Phytoplankton and other microalgae are diverse and changes in systematics and taxonomy are not uncommon. Correct identification of organisms is essential when monitoring phytoplankton. The most common tool is light microscopy although molecular techniques as well as fluorescence and electron microscopy are also used.

The aim of the Nordic Microalgae information system is to present up to date information on microalgae and related heterotrophic groups at the web site www.nordicmicroalgae.org. The site is available to the public since mid-2011. A version adapted for small screen devices, e.g. smart phones, is available since April 2013 at http://m.nordicmicroalgae.org.

Fig. 1 - Dinophysis a phyto-plankton causing shellfish toxicity.
Source: www.nordicmicroalgae.org.

The system is based upon the concept of taxon sheets which contain information on a certain taxon, e.g. name, taxonomic position, harmfulness, size, biovolume information, images and links to other resources. Images and videos of organisms are contributed by users. Downloadable check lists of species and higher taxa are regularly quality controlled and updated.

A database with quantitative data from phytoplankton monitoring is part of the system. It is available at www.smhi.se as part of the Swedish National Oceanographic Data Centre. Modern database and web technology facilitates interactive production of distribution maps and graphs of time series of data.

The system is part of the Swedish Lifewatch project which contributes to EU Lifewatch.

A Climate Data Access and Visualization Tool for Marine Ecosystem Managers

Karsten Shein, NOAA National Climatic Data Center, Karsten.Shein@noaa.gov (USA)
John Stachniewicz, Global Science and Technology, Inc., John.Stachniewicz@noaa.gov
L. DeWayne Cecil, Global Science and Technology, Inc., Dewayne.Cecil@noaa.gov
Steve Ansari, NOAA National Climatic Data Center, Steve.Ansari@noaa.gov

1. Introduction

Marine ecosystems are under increasing pressure from climate change and variations. Climate drivers such as changing water temperature, water chemistry (e.g., pH, salinity), water level, or storminess result in adverse responses including mass mortality, loss of habitat, increased disease susceptibility, and trophic cascade feedbacks. However, although there exists a great quantity of climate data, and a variety of tools are available to aid in obtaining and displaying these data, marine ecosystem managers are faced with challenges in their attempts to obtain appropriate climate data and to transform those data into meaningful visualizations capable of supporting sound management decisions. The primary challenges are:

- The lack of uniformity in data set access and format;
- The requirement that data first be obtained before it may be visualized; and,
- The assumption that marine ecosystem managers are climate data experts …
  - They understand what data they require and where those data may be found.
  - They understand which of several similar data sets are most appropriate to a given management question.
  - They have the tools and training to access, visualize, and analyze those data.

In an effort to better address the needs of marine ecosystem managers and overcome these barriers, the U.S. National Oceanographic and Atmospheric Administration (NOAA) has created the Integrated Marine Protected Area Climate Tools (IMPACT) project, within which data access and visualization tools are being developed. The overarching goal of IMPACT was to ensure that data access, visualization, and understanding were based upon the actual management questions being posed, and therefore could optimize the resources needed to facilitate the identification, access, visualization, and use of appropriate data.

2. WCT-IMPACT

One such tool is the WCT-IMPACT software. NOAA's National Climatic Data Center (NCDC) has been working with the Gulf of the Farallones National Marine Sanctuary Ocean Climate Center to enhance and expand the functionality of NOAA's Weather and Climate Toolkit (WCT). The WCT (Fig. 1) is a freely available, Java-based tool (http://www.ncdc.noaa.gov/oa/wct/) designed to access and display a variety of NCDC's georeferenced climate data products (e.g., satellite, radar, reanalysis). However, the WCT requires the user to have already identified a data set of interest and gained access to it. This can limit its utility by users who are not knowledgeable about which data sets are relevant to their needs and where those data sets can be found.

The IMPACT extention to the WCT eliminates those requirements. Instead WCT-IMPACT couples a user query approach with a quasi-expert system that determines, subsets, retrieves, and loads the the appropriate data products for visualization and analysis by the user. At the core of WCT-IMPACT is a needs-assessment that was conducted through interactive and iterative dialog between NCDC programmers and GOFNMS managers and scientists. Relevant data products are identified based on the environmental variables that ecosystem managers have indicated are important to their ecosystems.

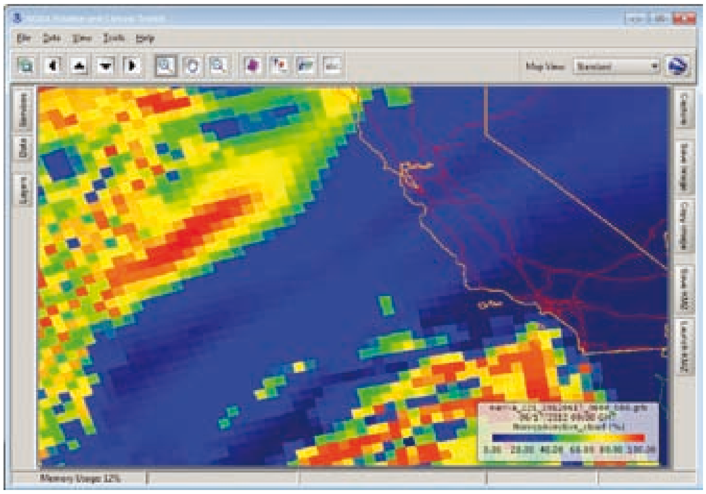


Fig. 1 - NOAA's Weather and Climate Toolkit image of NARR-A non-convective cloud cover off the US West Coast on 17 June 2012 at 9:00 GMT.

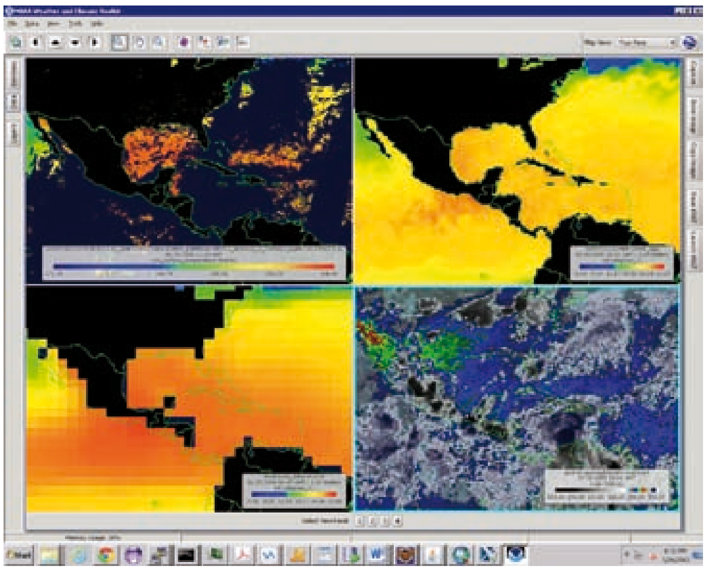


Fig. 2 - Four-frame data visualization window from the WCT-IMPACT software.

To address the data location and access issues, WCT-IMPACT contains an index file that connects environmental variables (e.g., cloud cover, sea surface temperature) with high-quality, high-resolution, and peer-reviewed Climate Data Records (CDRs). Additionally, NCDC servers routinely subset these datasets and provide climatological summary information over specific areas, reducing the time between access and visualization. Furthermore, data visualization includes the ability to examine conditions at a specific time, or animated over a time range. Multiple variables may be viewed synchronously on a multi-window display currently being developed (Fig.2). Furthermore, to address the erroneous assumption of marine managers as climate experts, WCT-IMPACT has incorporated contextual help features that explain the data and provide links to extended metadata. Climate tutorials are also being incorporated into the software. Finally, ecosystem managers asked for a tool that would perform and display the results of basic data analyses (e.g., sums, averages, probabilities).

This functionality is currently under development.

3. Summary

The WCT-IMPACT data access and visualization tool has been developed specifically to address the stated needs of marine ecosystem managers to combine data access and data visualization, and deliver relevant climate information to support management decisions, develop ecosystem-scale climate assessments, and produce visual aids in support of education and outreach efforts to the public, stakeholders, and policy-makers.

At every step of development, marine ecosystem managers have been included in the framing of requirements, so that not only is the tool developed to be of optimal use by them, but that the climate data managers who maintain the tool also become more aware of the needs of the marine ecosystem community. Features are included because the marine managers have asked for them, rather than because the data managers desire the tool to serve every possible need of every possible user. WCT-IMPACT is currently still in a prototype stage, but is expected to be released operationally within the upcoming months, via NCDC's website.

Australia's marine virtual laboratory

Roger Proctor, Integrated Marine Observing System, roger.proctor@utas.edu.au (Australia)
Peter Oke, CSIRO, peter.oke@csiro.au
Uwe Rosebrock, CSIRO, uwe.rosebrock@csiro.au

In all modelling studies of realistic scenarios a researcher has to go through a number of steps to set up a model in order to produce a model simulation of value. The steps are generally the same, independent of the modelling system chosen. These steps include determining the time and space scales and processes of the required simulation; obtaining data for the initial set up and for input during the simulation time; obtaining observation data for validation or use in data assimilation; implementing scripts to run the simulation(s); and running utilities or custom-built software to extract results. These steps are time consuming and resource hungry and have to be done every time irrespective of the simulation – the more complex the processes, the more effort is required to set up the simulation.

The Australian Marine Virtual Laboratory (MARVL) is a new development in modelling frameworks for researchers in Australia. MARVL, uses the TRIKE framework, a java-based control system developed by CSIRO to allow a non-specialist user configure and run a model, to automate many of the modelling preparation steps needed to bring the researcher faster to the stage of simulation and analysis. The tool is seen as enhancing the efficiency of researchers and marine managers, and is being considered as a educational aid in teaching.

In MARVL we are developing a web-based open source application which provides a number of model choices and provides search and discovery of relevant observations allowing researchers to:

1. efficiently configure a range of different community ocean and wave models for any region, for any historical time period, with model specifications of their choice, through a user-friendly web application,
2. access data sets to force a model and nest a model,
3. discover and assemble ocean observations from the Australian Ocean Data Network (AODN, http://portal.aodn.org.au/webportal/) in a format that is suitable for model evaluation or data assimilation, and
4. Run the assembled configuration in a cloud computing environment, or download the assembled configuration to run on any other system of the user's choice.

In its initial application MARVL has been applied to the estuarine / coastal region of the South East Tasmania (Derwent Estuary, Huon Estuary, Storm Bay), which has an established observing system, modelling system, and management regime. Fig. 1 shows a screen shot of the user-interface used to automatically configure the Derwent Estuary model. The window to the right shows details of the model configuration, including the start time, run length, grid size etc, and showing a map of the domain for which forcing fields are extracted to force the model. The Inner window (entitled "Create New Run") is used to select the model from a list of pre-defined models.

MARVL is now being applied in a number of studies around Australia ranging in scale from locally confined estuaries to the Tasman Sea between Australia and New Zealand.

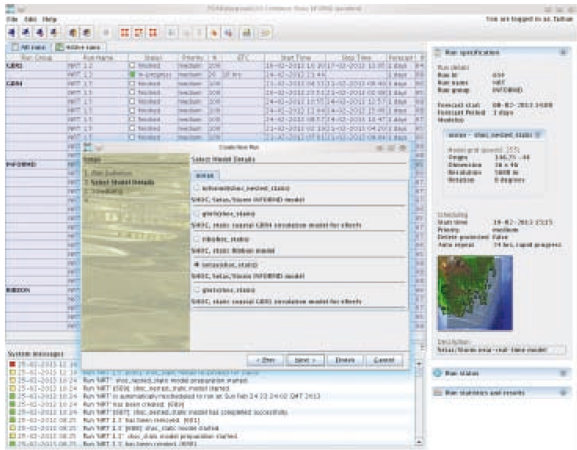


Fig. 1 - Screen shot of MARVL user-interface.

Web interface for the Black Sea database with GIS functionality

Elena Zhuk, Marine Hydrophysical Institute of National Academy of Science of Ukraine, zhuk.elena@nodc.org.ua(Ukraine)

Alexey Khaliulin, Marine Hydrophysical Institute, khaliulin@yahoo.com

Andrey Ingerov, Marine Hydrophysical Institute,ingerov@rambler.ru

Eugeny Godin, Marine Hydrophysical Institute, godin_ea@mail.ru

Marine Hydrophysical Institute keeps a lot of various data for the Black Sea region such as hydrophisical and hydrochemical data obtained during research cruises, current and temperature measurements from drifters, satellite images, operational oceanography data. The web GIS software was developed to keep and manipulate this data. This software is permanently improved and now it allows users to:

- import oceanography data from ODV files into the database. The quality control of data is made before loading new data into the database;
- select hydrochemical and hydrophisical data from the data base. A user can select any number of stations. The selection can be done visually using the following criteria: hydrochemical and hydrophisical parameters, date or seasonal period, space, cruise, research vessel, drifter information. The results are presented both on the map and plot for every selected parameter. Metadata and data are also shown. Visual selection of hydrophisical and hydrochemical data is shown in Fig. 1. In Fig. 2 you can see drifter's trajectory and temperature profile;

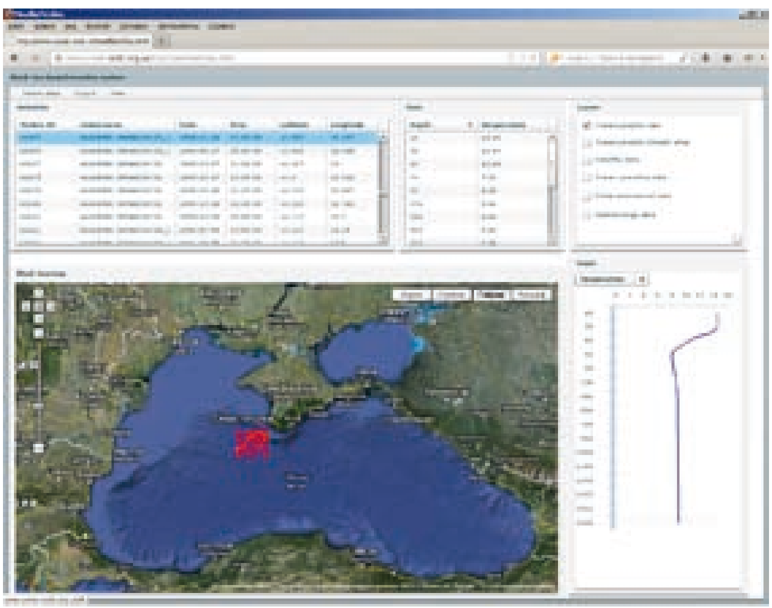


Fig. 1 - Oceanography data selection.

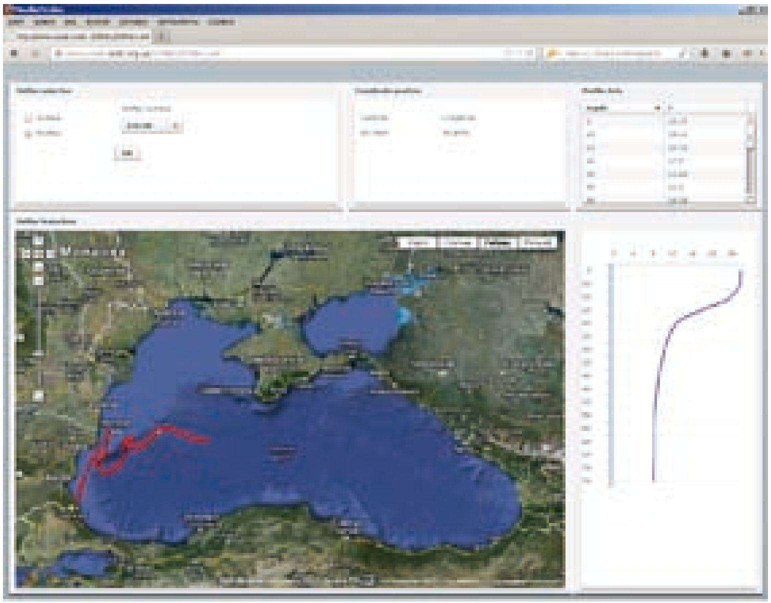


Fig. 2 - Drifter's data selection.

- export selected oceanography information to ODV and ODV SDN2 Spreadsheet files;
- make overlays of different information such as satellite images (MODIS AQUA Sea surface temperature, Chlorophyll concentration and Water leaving radiation), climatic maps (monthly, seasonal and annual) for hydrophisical and hydrochemical parameters, operational oceanography data, and meteorological information;
- interactive work with maps.

Celebrating 80 years of the Permanent Service for Mean Sea Level (PSMSL)

Kathleen Gordon, Permanent Service for Mean Sea Level (PSMSL), kmg@noc.ac.uk (UK)
Andrew Matthews, Permanent Service for Mean Sea Level (PSMSL)
Lesley Rickards, Permanent Service for Mean Sea Level (PSMSL), ljr@bodc.ac.uk

The PSMSL was established as a "Permanent Service" of the International Council for Science in 1958, but in practice was a continuation of the Mean Sea Level Committee which had been set up at the Lisbon International Union of Geodesy and Geophysics (IUGG) conference in 1933. Now in its 80th year, the PSMSL continues to be the internationally recognised databank for long term sea level change information from tide gauge records. Here, we present a review of the past, present and planned future activities of the PSMSL.

The PSMSL dataset consists of over 2100 mean sea level records from across the globe, the longest of which date back to the start of the 19th century. Where possible, all data in a series are provided to a common benchmark-controlled datum, thus providing a record suitable for use in time series analysis. The PSMSL dataset is freely available for all to use, and is accessible through the PSMSL website (www.psmsl.org). We review the state of the PSMSL dataset, describing the historical evolution of the catalogue and highlighting some of the geographic regions where data coverage remains poor.

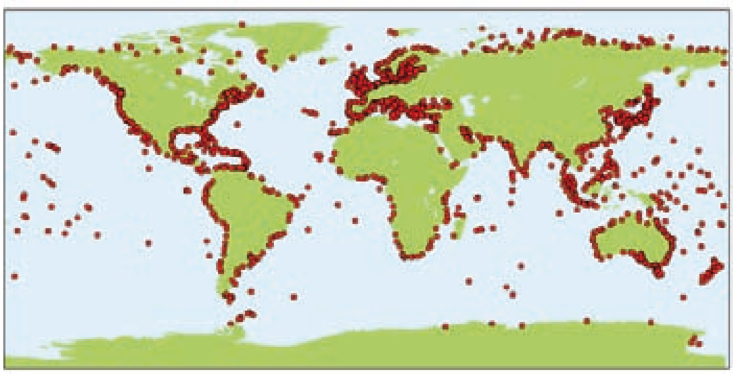


Fig. 1 - PSMSL Data Explorer.

The PSMSL also provides technical and scientific outreach information to coastal stakeholders and the general public. It provides training and supports to tide gauge operators across the globe. It plays an important role in the operation of the Intergovernmental Oceanographic Commission's Global Sea Level Observing System (GLOSS) and has participated in each of the Intergovernmental Panel on Climate Change assessment reports.

We also describe recent changes made to the PSMSL website, and describe some products added to allow users to explore the dataset interactively. We highlight some of the efforts being made to increase the quality and quantity of metadata available, and describe ongoing attempts to link our dataset with sources of higher frequency sea level data and other related data.

Tunisian Coastal Marine Atlas

Melika Belhassen, INSTM, belhassen.malika@instm.rnrt.tn (Tunisia)
Nouha Barraj, INSTM, nouha.ensi@gmail.com,
Soumaya Lahbib, INSTM, lahbib_soumaya@yahoo.fr
Saida Messaoudi, INSTM, messaoudi.saida@instm.rnrt.tn

Ocean Data and Information Network for Africa (ODINAFRICA) brings together more than 40 marine related institutions from twenty-five countries in Africa, including Tunisia since 2000. The network strives to address the challenges faced in ensuring that ocean and coastal data and information generated in national, regional and global programmes are readily available to a wide range of users in an easily understandable format.

One of the main outcomes of ODINAFRICA (2010-2013) is the Coastal Marine Atlas. The objective of the Atlas is to disseminate appropriate timely and relevant information about environmental variables along the Tunisian sea.

The Methodology used in this work consists to identify collect and organize different geospatial data with various temporal and spatial scales from online resources of public-domains. Data processed includes different topics:

– Hydrosphere (Chlorophyll, Sea Surface Temperature (SST), Sea Surface Salinity (SSS), etc…),
– Atmosphere (air-temperature, precipitation, wind speed and direction, etc…),
– Basemaps(Physical features, Bathymetry, etc…),
– Geosphere (Geology, Earthquakes, Tsunami, etc…),
– Human environment (Administrative and Maritime boundaries, Infrastructure, etc…).

Local datasets were also collected from several national data bases and includes mainly oceanographic and human environmental topics.

Data were converted into GIS format and implemented in a web application “Smart Atlas” developed by Coast and Marine Research Center, Ireland (CMRC).This application allowed to publish GIS data and to connect to metadata files generated under Geonetwork, The Atlas is online in the URL “www.africanmarineatlas.org”.

The Atlas has three different levels of lecture: regional, national and local. The regional level covers the Mediterranean and Red Sea, it goes from 39N, 41Eto12N, -6W and includes 208 layers. The national level covers the Tunisian coasts from 40N, 17E to 30N, 5Eandinvolves 205 layers. The local level covers the Golf of Gabes region and goes from 34N, 11E to 33N, 10E and includes 27 layers.

Tunisian coastal marine atlas is an ongoing activity and several national data bases are in the way to be processed. This product is a useful tool for decision makers and for marine and coastal management of Tunisian ecosystems.

SeaMobile apps – marine data and forecast services available on mobile devices

Daniel Hayes, OC-UCY, dhayes@ucy.ac.cy (Cyprus)
George Zodiatis, OC-UCY, gzodiac@ucy.ac.cy (Cyprus)
Dick Schaap, MARIS BV, dick@maris.nl (The Netherlands)
Vassos Vassiliou, OC-UCY vassiliouv@ucy.ac.cy (Cyprus)
Costas Vrioni, OC-UCY vriocom@gmail.com (Cyprus)
Claudia Coman, OC-UCY coman.claudia@ucy.ac.cy (Cyprus)
Andria Karaolia, WabbieWorks, andria@wabbieworks.com (Cyprus)
Iacovos Constantinou, WabbieWorks, iacovos@wabbieworks.com (Cyprus)

Recently a new project was initiated in Cyprus to develop the SeaMobile information service for giving tourist and recreational users access to customised marine data and forecasting services, adopting smart phones and mobile touch screen devices as distribution platforms. The project is co-funded by the Repubic of Cyprus and the European Regional Development Fund.

The SeaMobile project as a pilot has a focus on developing marine environmental information for tourist and recreational users, and later this can be expanded with dedicated services for businesses or public authorities. The project involves formulating and developing platform-optimised content and related services for users from the recreation sector. User requirements will be developed, and technical solutions for multiple mobile platforms will be tested. A number of information services will be defined and developed based upon the existing Cyprus Oceanoghraphy Center services. These are developed, maintained and operated by Cyprus Oceanoghraphy Center and their output, fine-tuned for the tourist and recreation sector, will be made available for distribution on mobile devices.

The application can be also used as an educational tool and for boosting public awareness of marine environmental issues. This might affect people’s attitude toward the environment and their actions. Furthermore, the information services can increase safety, because users will be warned of possible high waves or heavy currents and take these into account when planning their activities such as sailing, boat trips and others. Moreover, a dedicated promotion campaign will be designed and implemented together with the Cyprus Tourism Organization to reach out to the target community of users. This will add flavour to the services that are already offered to tourists and make Cyprus even more attractive as a holiday destination.

Authors Index

A

Abbamondi Chiara 293
Aburto-Oropeza O. 123
Adams Leoni 224
Alcorn Trevor 148
Alenius Pekka 259
Alexandrov Laura 179
Allen Terry 60
Allison Dicky 184, 213
Almeida Sara 163
Alory Gael 134
Alvera-Azcárate Aida 206
Alyautdinov Ali 243
Amaike Ken-ichi 127
Andreasson Arnold 301
Andreotti Valeria 293
Andrew Jenny 291
Angelica Carmelo 293
Ansari Steve 302
Appeltans Ward 203
Arko Robert A. 18, 38, 276
Arranz Lucia 87
Arroqui Juan 76
Azrilevich Paola A. 176

B

Baart Fedor 216, 239
Bäck Örjan 23, 301
Baek Sang-Ho 235
Baeye Matthias 209
Bakker Dorothee 281
Baldoni Ana 76
Balopoulou Stavroula 71
Baptista Rui 163
Barale Vittorio 283
Barraj Nouha 310
Barth Alexander 206, 263
Basic Josip 191
Bayley Dan 298
Bechini Alessio 267
Beckers Jean-Marie 206, 263
Behnken Andree 90
Belhassen Melika 310
Belokopytov Vladimir 136
Belov Sergey 103
Belova Kristina 103
Beltràn Joan Pau 125
Berdnikov Sergey 286
Blower Jon 44, 58
Bocharova Tatiana 146
Bohorquez Julio 144
Boldrini Enrico 32
Bon Antoine 44
Borghini Mireno 63
Borremans Catherine 22
Borst Kees 157
Bourassa Mark A. 38
Bradshaw Elizabeth 291
Brauer Peer C. 29
Breitbach Gisbert 141
Brosich Alberto 121, 265
Buga Luminita 179
Buonocore Cira 293
Busato Alessandro 108, 148, 253

C

Cabanes Cécile 143
Candela Leonardo 68, 105
Car Nicholas 36
Carbotte Suzanne 38
Carrara Paola 188
Carval Thierry 25
Cases Oscar 44
Castelli Donatella 105
Celentano Andrea 293
Cermelj Branko 165

Chalkiopoulos Antonis 25
Chandler Cynthia 18, 38, 184, 197, 213, 233, 276
Chassanite Aurore 96
Che-Bohnenstengel Anne 111
Chiappini Catia 267
Chunyaev Nikita 227
Ciancimino Sarah....................................293
Claus Simon 42, 200
Claverie Vincent..........................44, 58, 219
Coatanoan-Girou Christine 23, 25, 143
Coburn Elizabeth 276
Colella Simone 113
Coman Claudia 311
Constantinou Iacovos 311
Coro Gianpaolo.........................68, 105, 203
Corrado Lorena 293
Corre Marie-Paule...........................108, 253
Cotton David...60
Crosnier Laurence 219
Cucco Andrea..230
Czerniak Andreas....................................257

D

D'Angelo Alessandra 293
Dadic Vlado 82, 191
Dawson Garry 284
De Alfonso Marta.....................................25
De Bruin Taco 181, 284
De Cauwer Karien 108, 253, 299
De Dianous Rémi........................44, 58, 219
De Hauwere Nathalie 200
De Koster Ronald....................................181
De Lauro Marinella.................................293
De Leo Francesco 96
DeWayne Cecil L. 302
Delcroix Thierry......................................134
Delgado Claudia......................................295
Deneudt Klaas 42
Devine Laure...197
Devolder Mia 299
Dezeljin Damir..165
Di Luccio Diana......................................293
Diepenbroek Michael..............................281
Diggs Stephen 193
Diviacco Paolo 108, 148, 209, 253
Domalain Gilles 248
Dorandeu Joël 219

E

Efthimiou George....................................153
Ellenbroek Anton 203
Esquivel Jimenez 79
Eversley Ann-Marie 289

F

Fazioli Leopoldo 230
Fernandez-Guerra Antonio.........................84
Ferrigno Federica 293
Ferrini Vicki 38
Fichaut Michèle 15, 23
Finney Kim 34
Fleischer Dirk..257
Foden Jo 255
Foglini Federica 96
Forneris Vega 113
Frantzis Alexandros 271
Freeman J. Eric 40
Fric Jacob 271
Frunzo Angela...293
Fuentes Jose Eduardo..............................144
Furmanek Tomasz 154
Fyrberg Lotta..301

G

Gaillard Fabienne....................................134
Galanis George.......................................221
Galparsoro Ibon 182
Galvarino Charlton 184
Garcia Hernan 197
Gatti Joaquìn 76
Gertman Isaac 136
Gheorghe Flavian......................................92
Giorgetti Alessandra 121, 188, 265
Gies Tobias.......................................25, 161
Gioia Antonio...68
Glaves Helen 21, 117, 148, 180
Glöckner Frank Oliver 84

Glover David 184, 213
Godin Eugeny 270, 307
Golde Ilmer 163
Gòmara Sonia....125
Gordon Kathleen 309
Gorringe Patrick 118
Grande Valentina....96
Grandi Alessandro 47
Griffiths Charly 224
Griffiths Guy 44, 58
Groman Robert 184, 213
Grouazel Antoine 25, 143
Guerrero Raùl....76
Guibert Sarah 115
Guillotin David 25

H
Hall Per 90
Hällfors Guy....186
Hällfors Seija 186
Hammarklint Thomas 25
Hankin Steven....281
Hansson Lars Johan 301
Hartman Anders 25
Hasselbring Wilhelm....29
Harvey Julio....73
Hassen Esmael M.....174
Hayes Daniel 311
Heckendorff Stephan....208
Hendriksen Gerrit....157, 239
Hernàndez Alberto 87
Hernandez Francisco 26, 42, 200
Hietala Riikka 259
Hodge Jonathan....36
Holdsworth Neil 66, 284
Huber Robert....56, 90
Huttunen Maija 186

I
Ingerov Andrey 136, 270, 307
Ioakim Chryssanthi 153
Iona Athanasia....92
Iona Sissy 23, 71
Ishii Masayoshi 127
Issaris Yiannis 271
Italiano Angela....203
Ivankovic Damir 82, 191

J
Jaccard Pierre 25
Jensen Hans Mose....42
Johansen Marie 301
Jolibois Tony 44, 58

K
Kaitala Seppo 25
Kalinkin Ilya 243
Kalkavouras Konstantinos 92
Karagevrekis Pelopidas....71
Karlson Bengt 301
Karaolia Andria 311
Kassis Dimitrios 25
Keeble Kathryn 255
Keeble Simon....255
Keeley Robert 55
Kerr Yann 115
Khaliulin Alexey 136, 270, 307
Kideys Ahmet E. 138
Kim Sung-Dae 235
Kinkade Danie 184
Kizu Shoichi....127
Kokkali Athina....271
Kononets Mikhael....90
Konovalov Sergey....197
Koppe Roland 172
Kottmann Renzo 84
Kõuts Tarmo 261
Kozyr Alex 52, 197, 281
Krasemann Hajo 141
Krieger Magali....134
Krivenko Yevgeniya....136
Kulesza Przemysław 169
Kulygin Valery 286

L
Lagona Francesco 274
Lagring Ruth 299
Lahbib Soumaya 310

Larigauderie Carole 115
Lampitt Richard 56
Lardner Robin 238
Laurenza Tiziana 293
Lavrova Olga 146
Leadbetter Adam..........................18, 49, 233
Lear Dan..224
Ledang Anna Birgitta 25
Lee Jin-Hee 235
Lehfeldt Rainer 129
Leone Maddalena 293
Lilover Madis-Jaak 241
Lisi Francesca 96
Lisovsky Richard 96
Loo Andreas 301
Lopes Ana 163
Lopez-Sagastegui Catalina........................79
Lora Sebastian 125
Louarit Abdelaziz 211
Loubrieu Benoît 151, 209
Loubrieu Thomas 25, 44, 58, 219
Loupian Evgeny 146
Lowry Roy 18, 49, 251
Luecker Mathias......................................161
Lykiardopoulos Angelo 71, 92

M

Macario Ana 172
Madelaine Jacques 248
Maffei Andrew 38, 276
Malde Ketil 174
Mancini Sebastien 34
Manzella Giuseppe 25, 118, 293
March David 125
Marinova Veselka 25
Martinelli Massimo 222
Mateescu Razvan 179
Matishov Gennady 286
Matthews Andrew 309
Maudire Gilbert 15, 45
McCaan Mike..73
Mehrtens Hela 257
Melles Johannes 129
Melnikov Denis.......................................227
Menzel Michael 208
Meola Bruno 177
Messaoudi Saida 310
Michida Yutaka 127
Mihailov Emanuela 179
Mikhailov Nikolay 103
Miller Stephen 21, 38
Mills David 255
Minster Jean-Bernard 94
Mirrione Alessandra 293
Mitropoulos Dimitris 153
Mityagina Marina...................................146
Mohlin Malin 301
Mokrane Mustapha 94
Molinari Graciela 76
Moncoiffé Gwen 197
Morales David F......................................144
Moreno-Bàez Marcia 79, 123
Moroni Davide 222
Moussat Eric 151
Myroshnychenko Volodymyr..................138

N

Nakano Toshiya.......................................127
Nardelli B.Buongiorno 113
Nardone Gabriele 274
Nast Friedrich 111
Nativi Stefano 32
Neprokin Oleksandr 96
Nikolaidis Andreas..................................221
Nokin Marc 63, 108, 253
Notrastefano Giulio 25
Novellino Antonio...................................118

O

O'Brien Kevin...281
O'Grady Eoin..156
Obregòn-Noriega V. 123
Oggioni Alessandro.................................188
Oke Peter..305
Olita Antonio...230
Olive Joan 63
Olsen Are 281
Olsonen Riitta 186

Orbe Iratxe 182
Oropeza Aburto..79
Osborne Mike 245
Ostrovskii Alexander 167
Ozer Tal...136

P

Pagano Pasquale 68, 105, 203
Pagnani Maureen 56
Palazov Atanas 159
Pannocchi Andrea 87
Panou Aliki..271
Pantazi Maria 271
Papadopoulos Vassilis..............................153
Park Hyuk-Min 235
Park Soo-Young 235
Parner Hjalte 101, 284
Partescano Elena 188, 265
Pascual Ananda 263
Perivoliotis Leonidas 25
Pessini Federica 293
Petit de la Villéon Loic 134
Pfeil Benjamin 281
Picone Marco 274
Pieri Gabriele 222
Pikula Linda 233
Pinardi Nadia 47
Pinto Carlos..66
Pisano Andrea 113
Pissierssens Peter 203, 233, 295
Pop Ristov Aleksandar.................................84
Pouliquen Sylvie 25, 118
Proctor Roger 21, 34, 305

Q

Quemener Gael 209
Quimbert Erwann..211

R

Ramaekers Gemma 216
Rauch Shannon 184
Raymond Lisa 233
Reed Greg 81, 197
Ribera D’Alcalà Maurizio.........................188
Ribotti Alberto 230
Rickards Lesley 25, 60, 94, 118, 284, 291, 309
Ridgeway Amy...298
Riefolo Luigia 293
Ringheim Lid Sjur..............................25, 284
Roiha Petra..25
Romero Laia.......................................44, 58
Rosebrock Uwe..305
Roviello Valentina......................................293
Rozo Daniel M..144
Ruggieri Stefano 293

S

Sabine Christopher......................................281
Sagan Slawomir 131
Sagarminaga Yolanda................................182
Sagen Helge 23
Sala Joan 44
Salihoglu Baris..138
Salomidi Maria..271
Salvetti Ovidio 222
Sánchez Jose Antonio 182
Sánchez-Rodriguez A................................123
Santinelli Giorgio.......................................216
Santoleri Rosalia 113
Saramago Ana 163
Sarretta Alessandro 188
Sato Kanako 127
Satra Le Bris Catherine 211
Schaap Dick ...15, 21, 23, 32, 117, 118, 148, 311
Schirnick Carsten.......................................257
Schlitzer Reiner 23, 55, 197, 208
Schmitt Thierry 151
Schnetzer Julia 84
Schroeder Katrin 267
Scory Serge 23, 108, 253, 299
Scotto d’Antuono Annalina......................293
Sebastian Kristian 125
Seeley Becky 224
Sellers Cynthia 276
Semin Vladimir 243
Serrentino Fabrizio....................................293

Seta Akihiro 127
Sgrosso Andrea 293
Shein Karsten 302
Shimizu Yugo 127
Skjevik Ann-Turi 301
Silvoni Gabriela 176
Simoncelli Simona 23, 47
Sinerchia Matteo 230
Sinquin Jean-Marc 108, 209, 253
Sloyan Bernadette 193
Smith Karl 281
Smith Shawn 38, 40
Soetje Kai 25, 161
Solovyev Vladimir 167
Sorensen Kai 25
Sorgente Roberto 230
Sorribas Jordi 63, 108, 209, 253
Soto Wilbert Marin 248
Souza Dias Francisco 200
Spinu Alina 179
Sposato Marina 293
Springer Pina 257
Stachniewicz John 302
Stahl Henrik 90
Stefanov Asen 159
Stevenson Alan 180
Stocks Karen 38
Stojanov Yvan 108, 253
Stolp Laura 276
Stolte Willem 239
Stylianou Stavros 221, 238
Suga Toshio 127
Suzuki Toru 127, 197

T

Tampucci Marco 222
Tarot Stéphane 115
Tatman Sharon 239
Tedesco Costanza 293
Tel Elena 87
Telszewski Maciej 281
Tengberg Anders 90
Tezcan Devrim 23, 138
Thiel Tim 276
Thijsse Peter 181
Tikka Kimmo 186, 259
Tintore Joaquin 25, 125
Tobia Angela Maria 293
Tonani Marina 23, 47
Tréguer Mickael 44, 211
Trigo Santiago 76
Tronconi Cristina 113
Troupin Charles 206, 263
Turpin Victor 143
Tyberghein Lennert 26

U

Urban Ed 233
Uvarova Ivan 146
Uysal Zahit 138

V

Vahter Kaimo 261
van der Kolff Gerard 157
Vandepitte Leen 26, 42
Vanhoorne Bart 26, 42, 200
Vassiliou Vassos 311
Vassilopoulou Vassiliki 271
Vetrano Anna 63, 267
Viazilova Natalia 229
Villa Paolo 222
Viloria Amaia 87
Vinci Matteo 49, 121
Virkkunen Mikko 186
Volpe Gianluca 113
Von Schuckmann Karina 25
Vos Erik-Jan 216
Vrioni Costas 311
Vucic Ivan 82, 191
Vyazilov Evgeny 227

W

Waldmann Christoph 90
Waumans Filip 26
Weatherall Pauline 251
Webb Tom 203
Webster Chloë 177
Wedhe Henning 25

Wichorowski Marcin 131
Wiebe Peter 184, 213
Wójcik Michał 169
Woodruff Scott 40
Worley Steven J. 40

Y

Yaitskaya Natalia 286
Yilmaz Elif...138
Yoritaka Hiroyuki 127

Z

Zananiri Irene...153
Zatsepin Andrey 167
Zhuk Elena 238, 307
Zhurk Elena 270
Zimianitis Vaggelis 153
Zodiatis George 221, 238, 311