Bollettino
di Geofisica
teorica ed applicata

Vol. 57
SUPPLEMENT

IMDIS 2016
International Conference on Marine Data and Information Systems

11-13 October, 2016
Gdansk, Poland

Istituto Nazionale di Oceanografia
e di Geofisica Sperimentale

ISSN 0006-6729

Vol. 57 - SUPPLEMENT - pp. 1 - 252

BOLLETTINO DI GEOFISICA teorica ed applicata

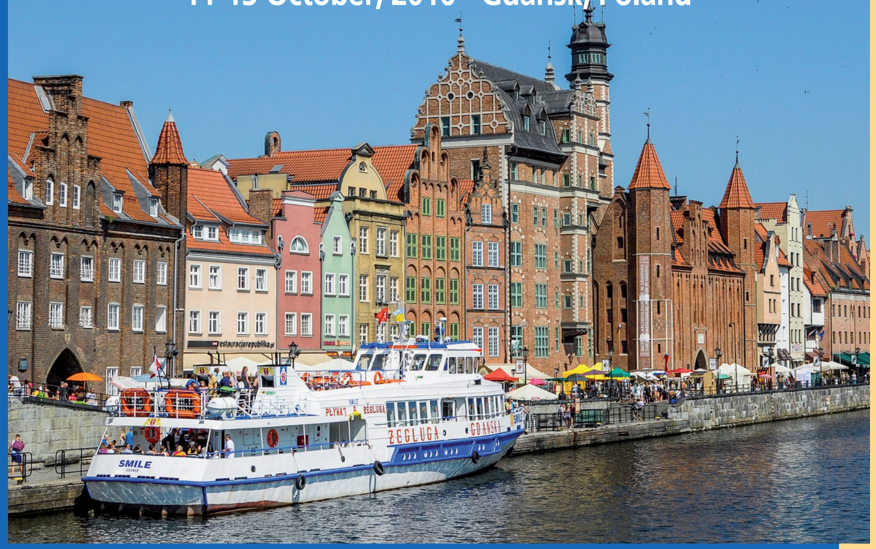
Vol. 57

Bollettino di Geofisica
teorica ed applicata

An International Journal of Earth Sciences

SUPPLEMENT

IMDIS 2016
International Conference on Marine Data and Information Systems
11-13 October, 2016 - Gdansk, Poland



Istituto Nazionale di Oceanografia
e di Geofisica Sperimentale

free colour printing

Bollettino di Geofisica
teorica ed applicata

An International Journal of Earth Sciences

Scope of the journal

The "Bollettino di Geofisica Teorica ed Applicata" is an international journal dedicated to the publication of original papers dealing with Deep Earth Geophysics, Near Surface Geophysics, Exploration Geophysics, Borehole Geophysics, Geodynamics and Seismotectonics, Seismology, Engineering Seismology, Geophysical Modelling, Geodesy, Remote Sensing, Seismic and Geodetic Networks, Oceanography, and their application in the fields of Energy, Natural Resources, Environment and Climate, Policies and Regulations, Risk and Security, Technological Development.
Issues of about 80 pages are published quarterly. Papers dealing with Solid Earth and Oceanography are grouped into two parts, A and B, respectively. The journal is published by OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale), Trieste, Italy.

The "Bollettino di Geofisica Teorica ed Applicata" is indexed in: Science Citation Index Expanded (also known as SciSearch®), SCOPUS, Current Contents® /Physical Chemical and Earth Sciences, Journal Citation Reports/Science Edition, Cabell's International.

The "Bollettino di Geofisica Teorica ed Applicata" is included in the product of EBSCO Publishing Inc.

SUBSCRIPTIONS 2016

	Italy	Other countries
Annual subscription rate	€ 65	€ 130
Back volumes	€ 75	€ 150
Separate issue	€ 22	€ 45
Back issue	€ 25	€ 50

All prices include post and packaging. For private persons (excluding institutes, laboratories, libraries, etc.) subscribing directly, the subscription rates are reduced by 20%.

Claims for missing numbers will be honoured only if received within 100 days of normal delivery date.

Subscription orders, changes in address and other business correspondence should be sent to the following address:

Bollettino di Geofisica Teorica ed Applicata
c/o Istituto Nazionale di Oceanografia
e di Geofisica Sperimentale
Borgo Grotta Gigante, 42/c
34010 Sgonico, Trieste, Italy
e-mail: bgta@ogs.trieste.it
http-server: www2.ogs.trieste.it/bgta

Publication B.A.M. exempt
Article 4 - 1° comma n. 6 - D.P.R. n. 627 of 6/10/87

INSTRUCTIONS TO AUTHORS

Types of contributions Scientific and technical papers, and review articles are the main contributions accepted. Moreover, the journal will also publish short notes, limited to a length of four pages including references; technical notes dedicated to the presentation of current research activity or significant original data. In the case of articles of strictly regional or local interest, the contribution should be limited to the essentials.

Submission of papers The manuscript should not have been published or submitted for publication elsewhere. Submission of a multi-authored manuscript implies the consent of all participating authors. Manuscripts shoud be submitted online at this web address: www.editorialmanager.com/bgta.

Reviewing process Before publication, each manuscript will undergo a reviewing process. The Editor-in-Chief will assign the manuscript to an Associate Editor who is directly in charge of the reviewing process of the manuscript and will send the manuscript for evaluation to reviewers. The final decision on publication will be made either by the Associate Editor or by the Editor-in-Chief. Editing and styling will be done at the discretion of the publisher.

Manuscript preparation All papers must be written in English and should not normally exceed 20 printed pages, including figures and tables. The manuscript should be written in maximum consistency with the style and layout of the journal.
The title of the paper should unambiguously reflect its contents. A suggestion for an abbreviated running title should be given.
The author name(s) and affiliation(s) should follow the title. Where an author has moved permanently or temporarily to a new address, this should be indicated as a footnote on the first page. An author to contact should be specified with full postal address, phone, fax, and e-mail.
Each paper must be provided with an Abstract of about 150-200 words, describing concisely the purpose and results of the paper. If the paper was presented at a meeting, mention should be made thereof in the acknowledgements.
The SI system should be used for all scientific and laboratory data: if it is necessary to quote other units, these should be added in parentheses. Abbreviations for units should follow the suggestions of the American Institute of Physics (AIP) Style Manual. Where abbreviations are likely to cause ambiguity or to not be readily understood by an international readership, units should be written in full. The full stop should not be included in abbreviations, e.g., m (not m.), ppm (not p. p. m.), "%" and "/" should be used in preference of "per cent" and "per".
Footnotes should be avoided, especially if they contain information which could equally well be included in the text.
The use of proprietary names should be avoided. Papers essentially of an advertising nature will not be accepted.

References should be cited at the appropriate point in the text with the author's name followed by the year of publication; thus, "according to Finetti (1982)" or "as shown in an earlier paper (Marussi, 1983)". When there are two or more papers by the same author published in the same year, the distinguishing letters, a, b, etc. should be added to the year. The name of two coauthors must be linked by "and"; for three or more, the first author's name must be followed by "et al.". References should be listed together at the end of the paper and arranged alphabetically without numbering. Only cited references should be included in the list. Examples of layout of references are given below.

Slejko D., Caporali A., Stirling M. and Barba S.; 2010: *Occurrence probability of moderate to large earthquakes in Italy based on new geophysical methods*. J. Seismol., **14**, 27-51, doi: 10.1007/s10950-009-9175-x.

Richter C.F.; 1978: *Elementary Seismology*. W. H. Freeman and Company, San Francisco, 768 pp.

Rikitake T.; 1975: *Dilatancy model and empirical formulas for an earthquake area*. In: Wyss M. (ed), Earthquake prediction and rock mechanics, Birkhuser Verlag, Basel und Stuttgart, pp. 141-147.

Garcia J. and Slejko D.; 2009: *Scenario earthquakes for tsunami generation in the Ionian Sea*. In: Slejko D. and Rebez A. (eds), Gruppo Nazionale di Geofisica della Terra Solida, 28° Convegno Nazionale, Riassunti estesi delle comunicazioni. Stella Arti Grafiche, Trieste, pp. 375-380.

Garcia J., Slejko D., Alvarez L., Peruzza L. and Rebez A.; 2003: *Probabilistic seismic hazard assessment for Cuba*. In: Slejko D. (ed), Gruppo Nazionale di Geofisica della Terra Solida, Atti del 21° Convegno nazionale, CD-Rom, Prospero, Trieste, file 04.20.

DISS Working Group; 2010: *Database of Individual Seismogenic Sources (DISS), Version 3.1.1: A compilation of potential sources for earthquakes larger than M 5.5 in Italy and surrounding areas*. <http://diss.rm.ingv.it/diss/>, doi: 10.6092/INGV.IT-DISS3.1.1, last access March 2015.

Montaldo V.; 2005: *Seismic hazard and uncertainties assessment in north-eastern Italy: comparing approaches with varied geological and seismological background*. PH.D. Thesis in Geological Sciences and Geo-Tecnologies for Environment and Territory, XVIII cycle, Università degli Studi di Milano - Bicocca, 146 pp.

Carulli G.B.; 2006: *Carta geologica del Friuli Venezia Giulia*. SELCA, Firenze.

Koulakov I.; 2011: *High-frequency P and S velocity anomalies in the upper mantle beneath Asia from inversion of worldwide traveltime data*. J. Geophys. Res., **116**, B04301, 22 pp., doi: 10.1029/2010JB007938.

Figures .jpg or .tiff files of each figure, with a 300 dpi resolution, should be provided upon request. Colour illustrations are accepted, and will be not charged to the authors.
Size of characters should be such that after reduction the smallest character will not be less than 1.5 mm.

Numbering and captioning Each illustration should be given a figure number in Arabic numerals. Each table should be given a table number in Arabic numerals.
Figures and tables must have a caption or title which is referred to in a list at the end of the manuscript. Captions should never be incorporated in the drawings. Scales should be indicated graphically and not numerically.
Equations should all be numbered sequentially with Arabic numbers, and refered to within the text by Eq. or Eqs. followed by the pertinent number in brackets.

Proofs The (contact) author will receive a set of proofs to be checked by typesetting editing. No new material may be inserted in the text at the time of proof reading unless accepted by the Editor-in-Chief. Proofs should be returned within a week.

Page charges and reprints There will be no page charges, including colour pictures. The .pdf file of the printed paper will be freely available at the BGTA web site. Paper reprints can be ordered on a reprint order form that will be sent with the page proofs and will be charged at current printing prices.

Copyright Publication of an article implies the transfer of the copyright from the author to the publisher, by signing the "Transfer of Copyright" agreement after manuscript is received. Permission to quote from this journal with acknowledgement of the source is granted.

Vol. 57 – SUPPLEMENT

Bollettino di Geofisica
teorica ed applicata

An International Journal of Earth Sciences

IMDIS 2016
International Conference on Marine Data and Information Systems

11-13 October, 2016 - Gdansk (Poland)

Organised by
IOPAN and IMGW
jointly with MEDIN and IFREMER, BODC, CSIC, MARIS and OGS,
as part of the SeaDataNet consortium.

Istituto Nazionale di Oceanografia
e di Geofisica Sperimentale

ISSN 0006-6729

TABLE OF CONTENTS

Session 1 - Data Services in ocean science 11

Oral presentations

L. Rickards, M. Vardigan, I. Dillo, F. Genova, H. L'Hours, J.-B. Minster, R. Edmunds and M. Mokrane
Developments in the certification of data centres, services and repositories through an RDA/WDS/DSA partnership 13

J. Buck, L. Darroch and A. Kokkinaki
Data standards; why are they important and what can they do for oceanography? 16

C. Wood
Development of Linked Data Services to Support Widespread Exposure of Data 18

P. Diviacco, A. Busato and P. Fox
A relaxed approach to handling semantics in data management 21

G. Coro, P. Pagano and U. Napolitano
Bridging environmental data providers and SeaDataNet DIVA service within a collaborative and distributed e-Infrastructure 23

J.-P. Babau, J. Tokotoko, O. Kande and A. Bilal
Noumea: a Model Driven Framework for NetCDF-CF Data Extraction and Analysis 26

J. Gourrion and T. Szekely
Improved statistical method for hydrographic climatic records quality control 29

J. Buck, T. Carval, T. Loubrieu and F. Merceur
A single DOI for Argo; a generic approach to making datasets that grow and evolve with time citable on legacy infrastructure 30

K. Finney, S. Mancini, R. Proctor and N. Atkins
Managing Australian Ocean Data Network (AODN) Vocabularies 32

E. Partescano, A. Brosich, A. Giorgetti, M. Lipizer and V. Cardin
(Near) real-time data publication for coastal and deep-sea observing system using OGC Sensor Web Enablement (SWE) standards 34

H. Parner
Making HELCOM Eutrophication Assessments Operational 37

A. Barth, S. Watelet, C. Troupin, A. Aida, G. Santinelli, G. Hendriksen, A. Giorgetti and J.-M. Beckers
OceanBrowser: on-line visualization of gridded ocean data and in situ observations 39

P. Thijsse and W. Som de Cerff
CLIPC: Tools and visualisation services for climate datasets to assess climate change 41

C. Troupin, J. P. Beltran, B. Frontera, S. Gómara, M. Gomila, A. Krietemeyer, C. Muñoz, M.À. Rújula, I. Serra and J. Tintoré
Data processing and visualization at the Balearic Islands Coastal Observing and Forecasting System (SOCIB) 43

V. Dadič, D. Ivankovič, S. Muslim, D. Jelavič and G. Beg Paklar
Oceanographic system to control the impact of construction works in the sea in a sensitive coastal area - Case Study 45

E. Zhuk, G. Zodiatis, S. Stylianou and R. Lardner
Web interface for the Oil Spill prediction software 48

Posters

G. Coro, G. Panichi and P. Pagano
A Web application to publish R scripts as-a-Service on a Cloud computing platform 51

J.-M. Sinquin, X. Lurton, C. Vrignaud, G. Mathieu and H. Bisquay
DORIS Software : Processing and management of Sound Velocity Profiles for echosounding applications 54

S.-D. Kim and H.-M. Park
Calculation of Regional QC ranges of temperature and salinity for Korean waters 55

E. Godin, A. Ingerov, E. Zhuk, E. Isaeva and I. Gertman
Desktop QC software test version Alexey Khaliulin, MHI RAS (Russia), khaliulin@yahoo.com... 56

E. Zhuk, G. Zodiatis, S. Stylianou, A. Karaolia and A. Nikolaidis
BYTHOS online visualization of ocean forecast 58

S. Mieruch, A. Wegener, R. Schlitzer and A. Wegener
WebODV: Providing Ocean Data View Services over the Internet 60

R. Thomas, A. Kokkinaki and R. Lowry
WebODV: Providing Ocean Data View Services over the Internet 62

D. Cano, E. Tel, C. González-Pola, R. Somavilla, A. Viloria, C. Rodríguez, M. Ruiz Villareal and A. Lavín
New developments on Biscay-AGL Observatory. From derived products to sensor networks 64

E. Bradshaw, L. Rickards, A. Matthews and S. Jevrejeva
The Global Sea Level Observing System (GLOSS) Data Rescue plan 66

G. González-Nuevo, E. Tel, A. Cabrero, P. Otero, M. Ruiz-Villareal and J.M. Cabanas
Subsurface Water Sampling Network of the Instituto Español de Oceanografía: from data acquisition to final users 68

F. Foglini, F. De Leo, J.M. Gómez, V. Grande, R. Leone and F. Marelli
A user-centric approach to validate the EVER-EST Virtual Research Environment infrastructure for the Earth Science 70

C. Borremans and F. Merceur
Marine Science Videos 73

M. Treguer, C. Satra Le Bris, E. Quimbert and J. Meillon
SEXTANT, a marine spatial data infrastructure: implementation of OGC protocols for the dissemination of marine data 75

E. Politi, R. Scarrott, M. Terra Homem, H. Caumont, N. Grosso, A. Mangin and N. Catarino
Unlocking the potential for coastal innovation growth using Earth Observation data and cloud infrastructure 77

M. Mokrane, S. Harrison and L. Rickards
The ICSU World Data System (WDS): An update on progress and current developments... 80

M. Osborne
Creating a Common Operating Picture and Improving Efficiency using Open Standards - Practical Examples from UK Ports 83

E. Viazilov
Decision support system (DSS) for disasters based on integrated database ± output of possible impacts and recommendations for decision makers........ 85

Session 2 - Marine information and data management 87

Oral presentations

H. Glaves and D. Schaap
Developing a common global framework for marine data management 89

M. Fichaut and D. Schaap
From SeaDataNet II to SeaDataCloud 91

S. Belov, N. Mikhailov and T. Spears
IODE Ocean Data Portal - platform to build national distributed data systems 94

S. Sá, N. Dwyer and M. Wichorowski
Towards the development of an overarching Marine Research Infrastructures information system 96

A. Lykiardopoulos, S. Balopoulou and S. Iona
SeaDataNet network monitoring services: Features and Statistics 98

È. Goďy and B. Saadatnejad
Arctic Data Centre, bridging scientific and operational data streams 100

A. Morvik, T. Hannant, A. Pobitzer and H. Sagen
Norwegian Marine Data Centre 102

S. Menegon, P. Penna, M. Bastianini, G. Stanghellini, F. Riminucci, T. Minuzzo and A. Sarretta
CNR-ISMAR in situ observations network: new approaches for an interactive, high performance, interoperable system 104

M. Wichorowski and K. Bachowiak-Samoyk
Aggregation and processing of data originating from heterogeneous sources for multidimensional analyses 106

K. O'Brien, S. Worely, B. Simons, B. Pfeil and E. Burger
Integrating Data and Information Across Observing Systems 108

P. Strömbergr, A. Andreasson and N. Nexelius
The end of reporting deadlines? Optimising dataflow by harvesting - a case study using open source REST-server technology from SMHI in corporation with ICES 110

T. Huang
Taking on Big Ocean Data Science 111

A. Leadbetter, D. Smyth, R. Fuller and E. O'Grady
Adding Big Data's Velocity to Oceanographic Data Systems 113

T. Loubrieu, J. Détoc, C. Borremans, A. Thorel and H. Azelmat
Sensor Nanny: support observation operators with equipment and data management services on the cloud 116

R. Proctor, P. Blain and D. Fruehauf
Enhancing User Access to Oceanographic Data Through Commercial Cloud Services 118

Posters

P. Gorringe and E. Buch
An Overview of the WMO Information System (WIS) 121

T. Vandenberghe, P. Diviacco, J.-M. Sinquin, J. Sorribas, R. Munoz, O. Garcia Sanchez, S. Scory and Y. Stojanov
EARS : Repositioning data management near data acquisition 123

A. Kobelev and E. Viazilov
Universal database of Unified System of Information on the World Ocean...................... 124

N. Mihaylov, E. Vjazilov, S. Belov and A. Voronsov
Approaches and implementation of marine data harmonization in Russian Federation.. 126

R. Lagring, M. Adam, A. Goffin, Y. Stojanov, L. Tyberghein and T. Vandenberghe
The implementation of a data management plan to uplift historical data: long-term change detection in the Belgian Continental Shelf.. 127

S. Jürgensmann, S. Feistel and S. Bock
IOWMETA - A modular metadata information system enabling the extraction of ISO-compliant metadata from heterogeneous sources 128

P. Otero, E. Tel and G. Gonzalez-Nuevo
IEO marine data discovery, representation and retrieval through metadata interoperability 130

C. F Postlethwaite, H. Williams, L. Rickards and G. Evans
MEDIN ± The Hub for UK Marine Data.. 132

E. Partescano, A. Crise, A. Giorgetti, L. Corgnati, C. Mantovani, V. Forneris, C. Tronconi and M. Drudi
A step toward data interoperability: NetCDF metadata comparative analysis in RITMARE Italian Project.. 133

V. Grande, A. Kruglov, F. Foglini, O. Leposhkin, R. Lisovskyi and O. Neprokin
The CoCoNet solution for management and access heterogeneous marine datasets and metadata 135

M. Arnold, C. Griffiths, M. Hall, E. Hughes and D. Lear
DASSH: where research and data meet... 137

B. Pfeil, M. Telszewski, K. O'Brien, T. Tanhua, A. Olsen, E. Burger and A. Kozyr
Initiating the Global Data Assembly Centres for Marine Biogeochemistry........................ 139

H. Glaves
ENVRIplus: creating a coherent, interdisciplinary and interoperable cluster of Environmental Research Infrastructures across Europe 141

T. de Bruin, L. Ponsoni, J. Welsh, J. Nauw, W. Sistermans, R. de Koster, J. Nieuwenhuis, T. Richter, M. Tacoma and R. Brink
Project 56: Data Management 143

R. Proctor, P. Bricher, J. Beja, A. Kozyr, A. Van de Putte, J. Cusick and S. Diggs
Raising the bar on data recovery and discovery for the Southern Ocean Observing System.. 145

P. Gorringe, V. Fernández and G. Nolan
The European Marine Observing Network and the development of an Integrated European Ocean Observing System. An EuroGOOS Perspective.. 146

Session 3 - Marine environmental data bases: infrastructures and data access systems 147

Oral presentations

L. Petit de la Villéon, S. Pouliquen, H. Wehde, J. Tintore, T. Carval, L. Sjur Ringheim, S. Tamm, S. Tarot, V. Marinova, L. Perivoliotis, M. de Alfonso Alonso-Muñoyerro, T. Hammarklint, F. Manzano Muñoz, C. Troupin, K. Balem and C. Guyot
Marine environmental data bases: infrastructures and data access systems Copernicus Marine Environment Monitoring Service In Situ TAC: an In situ operational data provision system for operational oceanography 149

G. Breitbach, W. Petersen, S. Reinke, P. Gorringe and A. Novellino
A common European database for underway data from FerryBoxes 151

F. Roquet, L. Boehme, M. Bester, H. Bornemann, S. Brasseur, J.-B. Charrassin, D. Costa, M. Fedak, C. Guinet, A. Hall, R. Harcourt, M. Hindell, K.M. Kovacs, C. Lydersen, C. McMahon, B. Picard, G. Reverdin and C. Vincent
When diving animals help us to observe the oceans: the MEOP data portal 152

L. Vandepitte, D. De Pooter, K. Deneudt, B. Vanhoorne, P. Provoost, W. Appeltans, N. Bailly, M. Eliezer, E. Fujioka, P. Goldstein, A. Giorgetti, M. Lewis, M. Lipizer, K. Mackay, G. Moncoiffe, S. Nikolopoulou, S. Rauch, A. Roubicek, C. Torres, A. Van de Putte, M. Vinci, N. Wambiji and F. Hernandez
Expanding the Ocean Biogeographic Information System (OBIS) beyond species occurrences, by including associated environmental data - experiences from the OBIS-ENV-DATA project 155

S. Lahbib, M. Dugenne, M. Libes, C. Sammari, M. Belhassen, M. Thyssen and G. Grégori
High resolution and automated flow cytometry data management 158

A. Nikolaidis, G. Zodiatis, S. Stylianou, G. Nikolaidis, E. Zhuk and E. Akylas
DINEOF daily cloud-free SST for the Eastern Mediterranean and Black Sea 160

A. Matthews, L. Rickards, E. Bradshaw, K. Gordon, A. Hibbert, S. Jevrejeva, S. Williams and P. Woodworth
The challenge of providing metadata for a 200 year long global mean sea level dataset 162

C. Coatanoan, S. Simoncelli, R. Schlitzer, M. Fichaut, D. Schaap and S. Iona
SeaDataNet II data products: the North Atlantic Ocean Region 164

S. Simoncelli, A. Grandi and S. Iona
New Mediterranean Sea climatologies 166

N. Holdsworth and C. Pinto
Development of the Impulsive Noise Register System 168

P. Gorringe, G. Manzella, D. Schaap, S. Pouliquen, L. Rickards and A. Novellino
EMODnet Physics: past, present and future 170

M. Vinci, A. Giorgetti, M. Lipizer and A. Brosich
EMODnet Chemistry: biogeochemical data management for long-term research and support to EU policies 171

D. Schaap
EMODNet Bathymetry - building and providing a high resolution digital bathymetry for European seas 174

S. Claus, L. Vandepitte, F. Waumans, B. Vitorino Pino and F. Hernandez,
Towards better data access: the development of a new data download tool for marine biological data ... 176

J. Populus, A. Palasov, N. Pinardi, E. Moussat and F. Blanc
EMODnet essential data needs and gaps: a comparative review of the Atlantic, Black Sea and Medsea Checkpoints ... 178

S. Scory, M. Fichaut, S. Iona, A. Novellino and D. Schaap
Streamlining the data ingestion process in the EMODnet context ... 181

Posters

G. Konshin and V. Seglinsh
Classification of geological data as a key to optimizing the preparation of CDI and ODV in the Geo-Seas project ... 183

V. Harscoat, S. Pouliquen and AtlantOS WP7 partners
Towards an integrated EU data system within AtlantOS ... 186

A. Kozyr and M. Krassovski
W.A.V.E.S ± Web Accessible Visualization and Extraction System for GLODAPv2 Database ... 188

C. Stegen Landa, S. Jones, B. Pfeil and T. Johannessen
The ICOS OTC Data Lifecycle Plan ... 189

L. Paglialonga, H. Mehrtens, P. Springer, C. Schirnick and C. Faber
Integrating data management services in marine sciences ... 192

P. Weatherall, K.M. Marks, M. Jakobsson and L. Rickards
Global Bathymetric Data Sets - General Bathymetric Chart of the Oceans (GEBCO) ... 195

T. Suzuki, M. Ishii, T. Kuragano, K. Sato, T. Suga, K.-I Amaike, Y. Karigome, S. Kizu, T. Nakano, Y. Shimizu, T. Yasuda, H. Yoritaka and Y. Michida
XBT Data Management and Quality Control in Japan (II) Improving Database by Historical XBT System ... 197

T. Vandenberghe, S. Degraer, T. Jauniaux, F. Kerckhof, B. Rumes, S. Scory, J. Haelters and E. Toussaint
MarineMammals.be: improving marine mammal stranding, by catch and observation data collection ... 199

M. Adam, R. Lagring, Y. Stojanov, F. Strobbe and T. Vandenberghe
Dealing with (historical) data and making it accessible: Data Inventory and Tracking System (DITS) applied in the scope of the "4 decades of Belgian marine monitoring" project (4DEMON) ... 201

G. Santinelli, G. Hendriksen and A. Barth
Analysis of ocean in situ observations and marine data products: Web-visualization and Services for EMODnet Chemistry ... 203

C. Pinto, I. Martinez, B. Schoute, N. Holdsworth and C. Morgado
The ugly duckling that turned into a beautiful swan (ICES Stock Assessment Database) .. 206

C. Borremans, V. Tosello, E. Pelleter and J. Detoc
ARCHIMEDE - A software for management of marine geological and biological samples and data ... 209

M. McCann, C. Sakamoto, J. Plant and K. Johnson
Biofloat: Python Software for Bio-Argo Float Data ... 211

D. Ivanković, V. Dadić and S. Muslim
Database based Operational Eddy Detection in High-Frequency Radar Surface Velocities ... 214

I. Baramidze, T. Gvarishvili, E. Mikashavidze, G. Makharadze, K. Bilashvili and M. Varshanidze
An analysis of biological data of the Georgian Black Sea coastal zone ... 216

T. Szekely and J. Gourrion
CORA4.3: A global delayed time mode validated in-situ dataset ... 217

E. Moussat, N. Pinardi, G. Manzella, F. Blanc and S. Simoncelli
EMODnet MedSea Checkpoint for Blue Economy efficiency and environment protection ... 218

Session 4 - Services for Users and Education ... 221

Oral presentations

X. Hoenner, S. Mancini, R. Proctor, J. Everett and K. Swadling
Bringing ocean observations to the classroom - student access to open data ... 223

A. Leadbetter, E. O'Grady and N. Burke
Ireland's Integrated Digital Ocean ... 224

V. Barale
A picture is worth a thousand words: the European Atlas of the Seas ... 227

P. Thijsse, M. Wernand and H. van der Woerd
The EyeOnWater concept: Marine observations via participative science ... 228

A. Drago, A. Zammit, R. Tarasova, A. Gauci, A. Galea, J. Azzopardi, G. Ciraolo and F. Capodici
KAPTAN - A smartphone application for mariners ... 231

P. Otero, G. Gonzalez-Nuevo, E. Tel, M. Ruiz-Villarreal and J. M. Cabanas
Insights from the development of a downstream web-service to visualize ocean and meteorological forecast data at Iberian Atlantic beaches ... 234

Posters

S. Kaitala, S. Lehtinen, S. Suikkanen, J. Ruohola and H. Kaartokallio
Development and use of HELCOM COMBINE phytoplankton dataset in SeaDataNet compliant format for the Baltic Sea ... 239

E. Zhuk, A. Khaliulin and G. Zodiatis
The Black Sea GIS developed on basis of free software ... 240

C. Troupin, B. Frontera, J. P. Beltran, A. Krietemeyer, K. Sebastian, S. Gómara, M. Gomila, R. Escudier, M. Juza, B. Mourre, A. Garau, T. Cañellas and J. Tintoré
Medclic: the Mediterranean in one click ... 242

P. Alenius, R. Hietala and K. Tikka
Evolution of physical oceanography observations in Finland ... 244

SESSION 1

Data services in ocean science

- Standards and quality-assurance issues
- Services and visualisation tools
- User oriented services

ORAL PRESENTATIONS

Developments in the certification of data centres, services and repositories through an RDA/WDS/DSA partnership

Lesley Rickards, National Oceanography Centre (United Kingdom), ljr@bodc.ac.uk
Mary Vardigan, ICPS University of Michigan (United States of America)
Ingrid Dillo, Data Archiving and Networked Services (Netherlands)
Françoise Genova, Centre de Donnees astronomiques de Strasbourg (France)
Hervé L'Hours, UK Data Archive (United Kingdom)
Jean-Bernard Minster, University of California, San Diego (United States of America)
Rorie Edmunds, WDS International Programme Office, NICT (Japan)
Mustapha Mokrane, WDS International Programme Office, NICT (Japan)

The International Council for Science World Data System (ICSU-WDS) is striving to build worldwide 'communities of excellence' for scientific data services by certifying Member Organizations - holders and providers of data or data products - from wide-ranging fields by using internationally recognized standards. WDS Members are the building blocks of a searchable common infrastructure, from which a data system that is both interoperable and distributed can be formed.

As part of the process of developing the WDS, certification criteria and a procedure for evaluating candidates for membership were developed by the WDS Scientific Committee to ensure the trustworthiness of WDS Members in terms of authenticity, integrity, confidentiality and availability of data and services. Certification is important because it promotes trust and confidence in the usability and persistence of shared data resources. It also helps repositories, data centres and services improve their practices and procedures. The WDS has 17 certification criteria grouped under the following headings:

- WDS general requirements and policies (Organization specific requirements)
- Organizational framework
- Management of data, products and services
- Technical infrastructure

As of February 2016, the World Data System has 95 Member Organizations in four different categories: 61 Regular Members; 10 Network Members; 6 Partner Members; 18 Associate Members, with reviews underway for several other applicants.

In addition to the WDS Certification, several certification standards and accreditation procedures for data repositories at differing levels of rigour have been developed worldwide. These include: Data Seal of Approval (DSA), Network of Expertise in long-term Storage and

Accessibility of Digital Resources in Germany (nestor) Seal/German Institute for Standardization standard 31644 and the Trustworthy Repositories Audit and Certification criteria/International Organization for Standardization (ISO) standard 16363. The DSA and the WDS are two examples of organizations that have set up core certification mechanisms; however the two standards have evolved and operated independently. The primary focus of DSA has been on digital repositories in the Humanities and Social Sciences, whereas the focus of WDS has been on Earth and Space Sciences for historical reasons. Nevertheless, both initiatives have fully multidisciplinary missions.

Fig. 1 - The different certification levels, each with increasing rigour.
DSA and ICSU-WDS represent core certification.

In 2014, under the umbrella of the Research Data Alliance (RDA)/WDS Interest Group on Certification of Digital Repositories, a Working Group (Repository Audit and Certification DSA-WDS Partnership WG) was established - consisting of representatives from the DSA and WDS communities and beyond - to explore and develop a DSA-WDS partnership with the objectives of realizing efficiencies, simplifying assessment options, stimulating more certifications and increasing impact on the community.

The WG has delivered a harmonized catalogue of criteria for core certification of repositories that is drawn from the DSA and WDS requirements, as well as a set of common procedures for repository assessment. This presentation will examine and discuss these, outlining the issues overcome in establishing the catalogue of Common Requirements. These are expected to replace the current certification catalogues of the Data Seal of Approval and the ICSU World Data System. They include an introduction on the benefits of certification, background information, all guidance text and a glossary. Both the new catalogue and (to some extent) the procedures have been validated through a testbed comprised of DSA and WDS community members undergoing renewal of their certifications. The testbed results have been published in short report; the insights resulting from this and subsequent improvements to the Common Requirements will be assessed.

By reconciling the fundamental and lightweight certification mechanisms of DSA and ICSU-WDS, the joint WG has taken a first step in simplifying the array of certification options, and shown the value to be gained from a certification procedure requiring relatively low investment of time and effort. Looking to the future, a proposal building on the WG outputs, Trust4Data, has been submitted to the EU. Its objective is to establish a Core Certification Service for a wide range of research data repositories and data centres. It carries the ambition to develop and create a standard information (object and process) model for Core Certification and deploy this in a service within a dedicated group of early adopters. A brief overview of the proposal and its outputs will be given.

Data standards; why are they important and what can they do for oceanography?

Justin Buck, National Oceanography Centre (United Kingdom), juck@bodc.ac.uk
Louise Darroch, National Oceanography Centre (United Kingdom), lorr@bodc.ac.uk
Alexandra Kokkinaki, National Oceanography Centre (United Kingdom), alexk@bodc.ac.uk

Abstract

As oceanographic data volumes and complexity grow with the development of new sensors and observing technologies it is becoming increasingly important to process and distribute data efficiently. To this end the British Oceanographic Data Centre work on the application of oceanographic data and software standardisation at levels from sensor to data delivery, is on-going across a range of European projects. This paper presents a solution which enables exposure of data as both Linked data and by applying the Open Geospatial Consortium Sensor Web Enablement standards and introduces semantic interoperability in the data output formats. The result of this work is to remove the barriers hindering the effective sharing of data across scientific domains and international boundaries.

Paper

As oceanographic data volumes and complexity grow with the development of new sensors and observing technologies, it is becoming increasingly important to process and distribute data efficiently. To this end the application of oceanographic data and software standardisation at levels from sensor to data delivery is on-going across a range of European projects. This will make the ingestion and data processing of oceanographic data more efficient and serve new users such as the producers of 'big data' data products and operational data assimilation/ingestion who require data to be unambiguously ingestible and served via APIs that enable machine to machine interaction. The two primary standards being implemented are the application of World Wide Web Consortium (W3C) Linked Data and Open Geospatial Consortium (OGC) Sensor Web Enablement (SWE) standards1,2.

The implementation of both standards at BODC is on-going via the EU Horizons 2020 BRIDGE, SenseOCEAN and AtlantOS projects and UK funded Celtic Seas project. The solution uses open source tools including ERDDAP, 52North SOS service and linked data software, which operate on top of an internal database holding sensor metadata.

Data modelling, which constitutes a step during Linked data publication, entails data modelling through the use of standardised domain and/or upper ontologies. Sensor metadata and sensor observation data are modelled through the fusion of a set of W3C ontologies namely: The Semantic Sensor Network ontology, the om-lite ontology, the GoodRelations ontology, the Time ontology, the GeoSPARQL ontology, the Provenance ontology and the SKOS ontology. All metadata is then semantically annotated using NERC Vocabulary Server vocabularies and concepts. The fusion of ontologies is shown in Fig. 1. This enables the storage of metadata sufficient for OGC SensorML, OGC Observations and Measurements, and Linked data based formats such as Semantic Sensor Network (SSN) and JSON-LD.

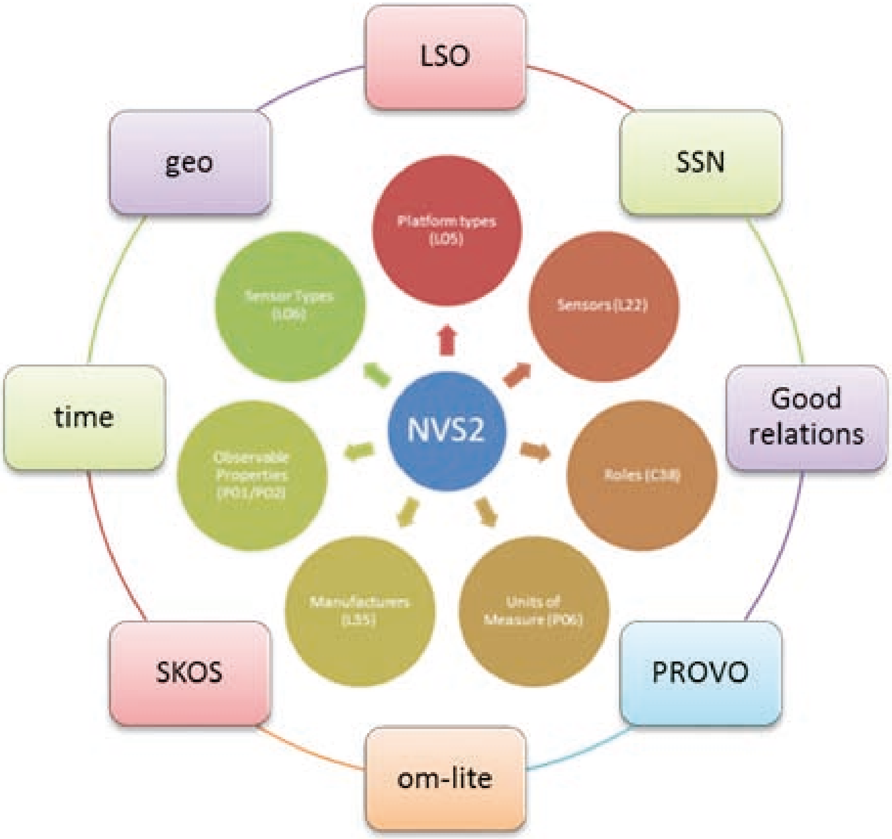


Fig. 1 - The ontological model used to stored metadata sufficient for production of both SSN and Sensor ML data descriptions.

In addition to exposing data to new users the application of such standards will make it possible to readily share data internationally via collaborations such as the Ocean Data Interoperability Platform (ODIP). World standards tend to be open and flexible to accommodate several scientific domains and applications, which can result in various incompatible implementations. To overcome this problem, international collaboration is essential, to enhance communication and specialise the standards across the domain. Collaboratively implementing data standards removes the barriers hindering the effective sharing of data across scientific domains and international boundaries.

References

World Wide Web Consortium. (2013). *Linked Data*. Available: http://www.w3.org/standards/semanticweb/data. Last accessed 8th October 2014.

Open Geospatial Consortium. (2014). *Sensor Web Enablement (SWE).* Available: http://www.opengeospatial.org/ogc/markets-technologies/swe. Last accessed 8th October 2014.

Development of Linked Data Services to Support Widespread Exposure of Data

Chris Wood, British Oceanographic Data Centre (United Kingdom), chwood@bodc.ac.uk

Historically, environmental data centres have tended to store all the metadata and data they control on internal databases. The lack of public interfaces to such databases means that the dissemination of the stored data is then a relatively labour intensive process, whereby a scientist or other interested user contacts the data centre requesting a particular piece of data, a data centre employee retrieves the relevant data, and emails it to the user. In the last few years, web technologies have allowed fairly complex websites to be developed which allows more of the metadata to be exposed and for increasing amounts of data and metadata to be downloadable directly by users. However, this approach still presents limitations to both the data centre and end users; data centres have to ensure that the development of such systems provides an intuitive interface for end-users to interact with – an unintuitive system means that end-users are likely to engage less with the interface and revert to contacting the data centre directly. The data centre also must ensure such a system is up to date with trends in interface design and searching ability to keep up with the experience that ever-increasingly technically-savvy web users expect. End-users, meanwhile, have to become familiar with the search interface for every service they wish to interact with.

This particular paradigm is being solved by some data centres, including the British Oceanographic Data Centre (BODC), by using an emerging World Wide Web Consortium (W3C) standard: the Resource Description Framework (RDF) Model. RDF is a data structure where the description of the data, as well as the relationships between both data and metadata is named. These descriptions and relationships are made up of three elements, and, as such are normally referred to as triples. Triples can be stored in specialist databases, known as Triplestores, and often allow the triples to be queried using the SPARQL query language. The relationships between data and its metadata is often described using Unique Resource Identifiers (URIs), and as such are ideal for exposing data on the web. Where a triplestore provides both a web-accessible front-end and the ability to query with SPARQL (commonly known as a SPARQL endpoint), it provides a convenient method for mass amounts of data to be exposed with little effort needed by a data centre. Triplestores become particularly useful when the data they hold is richly described using published ontologies (both very generic and subject-specific), and thus allows end-users to search the data using terms they are already familiar with.

Up until now, exposure of data via triplestores has been slow due to a chicken-and-egg style problem: data centres have been reluctant to spend time and effort implementing triple stores while there are few subject-specific ontologies that will allow their data to be searched effectively; meanwhile, ontology authors have not been creating subject-specific ontologies when there are too few triplestores that would use the ontology. This has led to a slow uptake of the technology by data-centres, with the number of end-users who are experienced in the use of the technology also trailing behind. However, the technology seems to be reaching a critical mass: as the number of data centres who are interested in using triplestores increase, the number of suitable ontologies increase, and the number of users who are experienced in the use of them, and expect data centres to provide them as a means to expose their data, increase. This in turn leads to more data centres providing SPARQL endpoints, and encourages ontology authors to provide richer and higher-quality ontologies.

This presentation will show how the BODC has implemented SPARQL endpoints for two distinct use-cases over the past year. First, I will show how BODC have exposed the vast majority of the metadata describing data series that we hold via an implementation of JENA, an open-source triplestore with an associated API that we have used to ensure that the triples that we store and expose are an accurate representation of the metadata held in the underlying database. As this triplestore is kept in sync with our publicly-available data, the SPARQL endpoint provides a convenient method for end-users to query the data and metadata that we hold. I will explain how the wide range of ontologies (both generic, such as Dublin Core (dc) and Simple Knowledge Organisation System (skos), and subject specific, such as the Ocean Data Ontology (odo) and the Marine Metadata Interoperability Semantic Web Services (MMISW)) that we have chosen to use to describe the data are a useful aid into discovering our data, particularly by users who already have the technical knowledge in using a SPARQL endpoint. I will also explain how we plan to expand the current service by implementing machine-to-machine data exchange. This will enable data to be automatically transferred or downloaded, in a range of file formats due to content-negotiation or URL structure that we will implement, as a result of a single, simple SPARQL query. We hope that such a system will transform the user experience associated with the search and download of specific datasets, either by using a native SPARQL endpoint, or subsequently by using simple user interfaces that have the potential to be built on top of the simple HTTP-based API that such an endpoint provides.

Secondly, I will show how a SPARQL endpoint has been used as the API for a particular application – in this case to support the development of a metadata portal that has been used to aid implementation of the Marine Strategy Framework Directive in the Celtic Seas Region. This has been used to support policy makers, special-interest groups, users of the marine environment, and other interested stakeholders in the legislation, which mandates 'Good Environmental Status' to be maintained or achieved by 2020, through a series of Programme of Measures. The metadata portal has been built to provide a signposting service to datasets that are relevant to MSFD within the Celtic Seas. Although the metadata are stored in a traditional RDBMS, they are exposed as linked data via the D2RQ platform, allowing virtual RDF graphs to be generated. This also ensures that the metadata has the widest possible reach, and allows us to expose the service, and associated metadata, as an API. SPARQL queries can be executed against the published end-point allowing any user to search the metadata.

As with the triplestore which is used to expose BODC's data series, we have again mapped a wide range of relevant ontologies to the metadata. However, in this case, we have used D2RQ's mapping language, based on the turtle format, to generate the mappings. The ontologies used [e.g. The Provenance Ontology (prov-o), Ocean Data Ontology (odo), Dublin Core Elements and Terms (dc & dcterms), Friend of a Friend (foaf), and Geospatial ontologies (geo)] allow users to browse the metadata, either via SPARQL queries or by using D2RQ's HTML interface. In this instance, the metadata were further enhanced by mapping relevant parameters to the NERC Vocabulary Server, itself built on a SPARQL endpoint, therefore allowing federated queries to be written.

Additionally, we have built a custom web front-end to enable users to browse the metadata and express queries through an intuitive graphical user interface that requires no prior knowledge of SPARQL, and therefore enables users to effectively filter, sort, or search the metadata. As well as providing means to browse the data via MSFD-related parameters (Descriptor, Criteria, and Indicator), the metadata records include the dataset's country of origin, the list of organisations involved in the management of the data, and links to any relevant INSPIRE-compliant services relating to the dataset. As the MSFD timeline requires Member States to review their progress on achieving or maintaining GES every six years, the timely development of this metadata portal will not only aid interested stakeholders in understanding how member states are meeting their targets, but will also show how a linked data approach, along with the associated technologies, can be used effectively to support policy makers, scientists, and other end-users.

A relaxed approach to handling semantics in data management

Paolo Diviacco, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), pdiviacco@ogs.trieste.it

Alessandro Busato, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), abusato@ogs.trieste.it

Peter Fox, Rensselaer Polytechnic Institute - RPI (United States), pfox@cs.rpi.edu

A very important aspect of geophysical data management is how to handle heterogeneity of the data. This can stem from different vintages, from differences in acquisition practices, resolution, data type, format and media that result in many issues in accommodating such heterogeneity within a single framework. In addition, the same dataset can be used in different contexts, with different research objectives. Data discovery can then be complex because data “tagging” is generally based upon data usage. If we “cage” data into preset possible cases, every time we have a situation or a need that does not match the pre-cooked cases, it will be rather difficult to hit what we are looking for.

Metadata is of course an important aspect of data discovery, but can “portray” only some of the many end user needs. It is necessary then to add to the traditional practices, other means to get to useful data.

Data, Information and Knowledge are the different layers that structure the environment in which each research takes place. Data management is traditionally relegated to the lower layer, under the assumption that the higher layers do not influence it. This, we think, is a gross mistake.

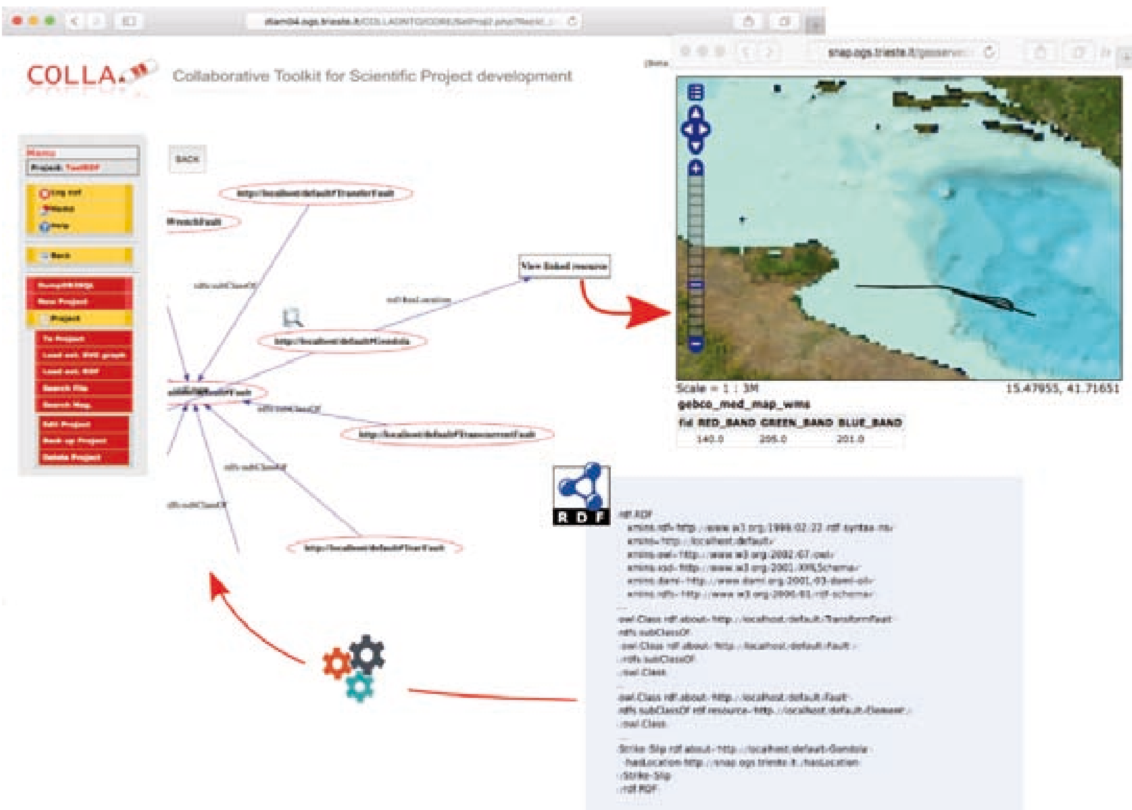
Information introduces meaning into the picture of basic measurement, and knowledge introduces context and cultural issues. A large literature in modern epistemology and sociology of science (see for example: Hanson, 1958; Duhem and Wiener, 1954; Putnam, 1975) informs us that measurements and observation are largely conditioned by theories; experiments conform to previous knowledge, and perception itself is conditioned by cognitive models. The aforementioned conditions must all be considered when providing data management and in particular, data search tools. Consequently, tools like controlled vocabularies and ontologies are filling the gap described above, introducing a semantic level where meaning is encoded. At the same time these very tools can introduce side effects. In fact, since semantic tools refer to a model “portraying” reality they need an explicit formalization of it. Considering that web ontologies are based on a denotative approach, a change in the structure of the model may inevitably lead to a change in the set of results.

Can formalization of a domain vocabulary be achieved? Some authors in the line of view of Polanyi (1966) are rather pessimistic on this.

We propose a more “relaxed” approach that integrates, metadata, visualization, representation, and semantic tools that together can help the end user to establish a path to get to what he or she is looking for.

This is based on the introduction of graphic representations of the domain of knowledge in which data and information are to be used. These can be thought as maps that, following Suchman (1987), show the possible ways to travel from one place (here concepts) to another without conditioning users in their choice. Besides, even if referring to the same data, multiple and concurrent representations (maps) can be designed for the same or different domains of knowledge. Tools have been developed to produce such maps and to organize data and information in them. One implementation we found very promising is based on the use of RDF files, that easilly can introduce in the maps the use of linked data.

Prototypes of the system have already been developed and are currently under study, while first tests show encouraging results.



References

Duhem, P., & Wiener, P. P. (1954). *La Théorie Physique: Son Objet et sa Structure [The Aim and Structure of Physical Theory].* Princeton University Press.

Hanson, N. R. (1958) *Patterns of Discovery.* Cambridge, UK: Cambridge University Press.

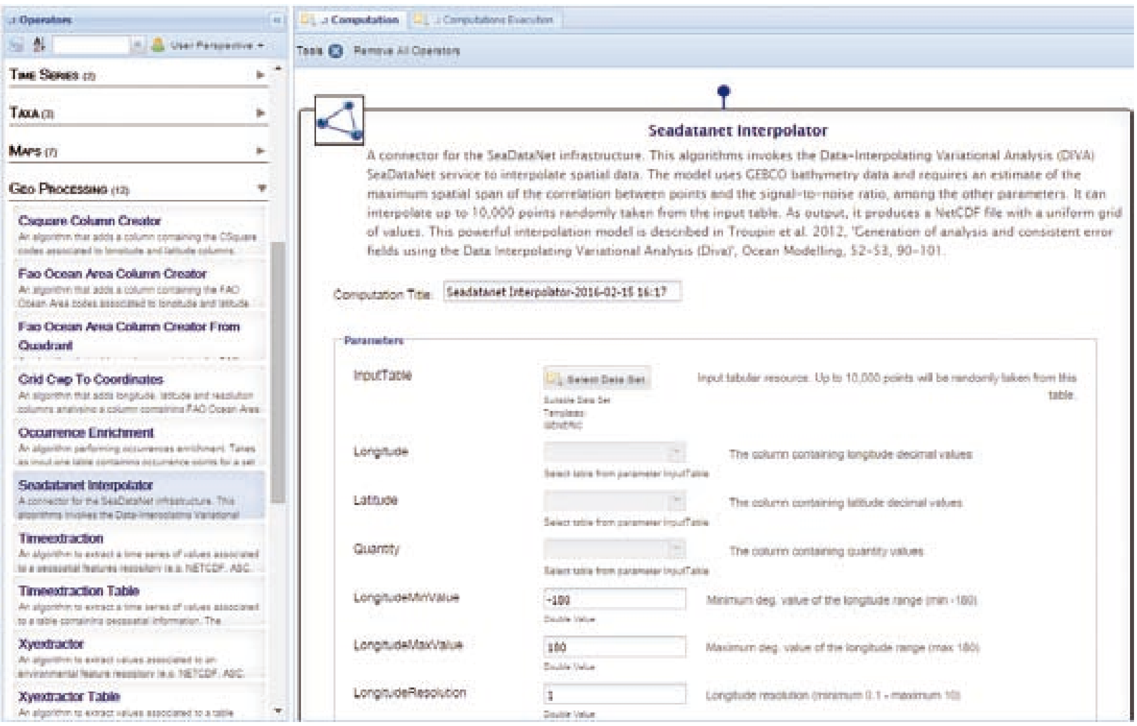
Polanyi, M. (1966). *The Tacit Dimension.* New York: Anchor Day Books.

Putnam, H. (1975). *Mind, language and reality: Philosophical papers (Vol. 2).* Cambridge, UK: Cambridge University Press. doi:10.1017/ CBO9780511625251

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication.* Cambridge, UK: Cambridge University Press.

Bridging environmental data providers and SeaDataNet DIVA service within a collaborative and distributed e-Infrastructure

Gianpaolo Coro, ISTI Consiglio Nazionale delle Ricerche (Italy), gianpaoloninited de.coro@isti.cnr.it
Pasquale Pagano, ISTI Consiglio Nazionale delle Ricerche (Italy), pasquale.pagano@isti.cnr.it
Umberto Napolitano, ISTI Consiglio Nazionale delle Ricerche (Italy), umberto.napolitano571@gmail.com



Among its devastating effects, ocean acidification harms organisms whose life depends on shells and on coral reefs. Therefore, it affects important marine ecosytems hosting high biodiversity and food availability. A recent paper published on Nature (Ries, 2012) has highlighted the decrease of calcification of a very important planktonic mollusc, the sea butterfly (*Limacina helicina*). The shell-building capacity of this organism declines with decreasing aragonite saturation, which is due to the increase of average pH in the seas. Monitoring and containing ocean acidification helps preventing coral reefs and shell-building organisms to dissolve, and thus helps preventing ecological disasters. Unfortunately, environmental observations of parameters like aragonite saturation and pH are usually available as scattered *in situ* data, published on restricted-access data e-Infrastructures (e.g. the Copernicus Marine Environment Monitoring Service). On the other hand, interpolation services exist (e.g. the SeaDataNet Data-Interpolating Variational

Analysis service, DIVA) to estimate global, uniform distributions of environmental parameters from scattered observations. However, these services usually require data to be compliant with a non-standard format and cannot accept in situ data formats directly. Furthermore, they do not support facilities neither to communicate nor to publish their results for a large public.

The D4Science e-Infrastructure is a distributed computer system that aims at supporting large-scale resource sharing (both hardware and software) via the definition of Virtual Research Environments (VREs) and allows data to be processed with distributed computing (Coro *et al.*, 2015). D4Science is able to create a bridge between several e-Infrastructures to fill the communication gap between them. This bridge is realised through a set of data storage and processing web services that are easy to extend. As for the ocean acidification monitoring case, we present a D4Science process (*D4Science-SeaDataNet Interpolator* service) that can transform a set of Copernicus *in situ* marine observations of an acidification parameter into a uniform global distribution map, published on a data catalogue under standard geographical representation formats (e.g. Web Map Service, Web Feature Service, Web Coverage Service). This process is hosted on the D4Science computing platform, which uses also European Grid Infrastructure FedCloud resources, and is able to (i) import Copernicus observations, (ii) organize them into a suitable format for the SeaDataNet interpolation service, (iii) execute the interpolation process and (iv) publish or share the output as a gradient map. The process invokes the SeaDataNet DIVA interpolation service using a distributed computing strategy to process both the input and output. Finally, the e-Infrastructure provides a geospatial services network, on which the computing system can store and publish data as GIS maps. The process is also endowed with a graphical user interface, available through a Web portal[1] (i-marine.d4science.org, Fig. 1). Through D4Science, users can establish the access policy for the input and output data, e.g. they can share data either with selected colleagues or with all the participants to a Virtual Research Environment they are involved in.

The integration realised by our service produces several benefits both to the data and model providers of SeaDataNet. We can summarise these benefits as follows:

Multi-users and concurrent load management: D4science services manage concurrent requests for computationally intensive processes and monitor the total allowed requests per user, using accounting facilities.

Virtual Research Environments: D4Science supports Virtual Research Environments that allow domain scientists to work together, exchange results, data etc. The *D4Science-SeaDataNet Interpolator* service can be published for selected users only and the D4Science social platform allows discussing about the results.

Storing output on a high-availability storage system: input and output are stored on a high-availability service offered by D4Science, based on a secure, fault-tolerant and fully replicated storage system.

Automatic generation of a web user interface: D4Science automatically endows integrated algorithms and services with a Web user interface, based on a declaration of the input and the output of the model. This is valid for the Interpolation service too.

Standard execution interface: the Intepolation service is made available through the Web Processing Service (WPS) standard of the Open Geospatial Consortium. This allows invoking

1 The complete process is explained in a video available at http://goo.gl/yX3kww.

models via REST communication and obtaining standard description of input and output. WPS increases the possibility to integrate the DIVA service with other software (e.g. QGIS) or workflows management systems (e.g. Taverna, Galaxy etc.).

Management of provenance information: D4Science tracks the experimental setup, the input and output of the DIVA service and allows other people to reproduce any experiment while getting the same results.

Data sharing and publishing facilities: D4Science users are endowed with a distributed file system accessible through a web interface (Workspace). This system allows sharing folders and files with (i) selected people, (ii) all the participants to a certain VRE, (iii) people outside of the e-Infrastructure (through the generation of persistent public links).

Applicability of the DIVA model to more data coming from direct observations: D4Science takes care of data harmonization and staging. Data from in situ observations are imported as generic CSV files and formatted for the SeaDataNet DIVA service.

Publication of the results as a GIS map: the NetCDF files produced by the DIVA service can be made publicly available under a number of standard representations (WMS, WCS, OPeNDAP, Esri GRID etc.). This increases the accessibility of the results by other systems.

In summary, our bridge expands the possibilities offered by SeaDataNet. In particular, it adds efficient data pre- and post- processing, sharing and publication facilities. Additionally, it increases the applicability of the DIVA service and expands the dissemination of the results. Therefore, it can facilitate producing global information about ocean acidification from scattered sea observations, with respect to the original scenario of independent infrastructures.

References

Ries, J.B., 2012. *Oceanography: A sea butterfly flaps its wings.* Nature Geoscience, 5(12), 845-846.

Coro, G., Candela, L., Pagano, P., Italiano, A., and Liccardo, L., 2015. *Parallelizing the execution of native data mining algorithms for computational biology*. Concurrency and Computation: Practice and Experience, 27(17), 4630-4644.

Noumea: a Model Driven Framework for NetCDF-CF Data Extraction and Analysis

Jean-Philippe Babau, Lab-STICC Universitè de Bretagne Occidentale (France), babau@univ-brest.fr
Jannaï Tokotoko, Lab-STICC Universitè de Bretagne Occidentale (France), tokotokojannai@yahoo.fr
Oumar Kande, Lab-STICC Universitè de Bretagne Occidentale (France), oumarkabde15@yahoo.fr
Abdallahi Bilal, Lab-STICC Universitè de Bretagne Occidentale (France), abdallahy.bilal@gmail.com

Introduction

To improve interoperability, scientific dataset modeling follows abstract standards like the Unidata's Common Data Model (CDM) [1]. To implement CDM, NetCDF [2] proposes a set of software libraries and self-describing, machine-independent data formats. NetCDF supports the creation, access, and sharing of array-oriented scientific. Because CDM and NetCDF are generic purpose model and libraries, scientists use specific standards like Climate and Forecast (CF) [3] or OceanSITES [4]. The standards propose constraints on the metadata declarations to facilitate file exchange and their manipulation.

To check if NetCDF data conforms to a standard, a code-oriented checker is used classically. Thus, the constraints are not formalized and a modification in the standard results in a manual modification of the code. Moreover, when a file does not conform to the standard, the file correction requires manual modifications or another specific tool. Another aspect is that, if NetCDF appears as a standard for storing scientific data, other formats are used for specific applications (sensor acquisition, specific analyzing tools). And for a huge volume of data, solutions based on Big Data technologies have to be investigated. To finish, for data following standards and stored in an efficient way, data visualization and analysis require other dedicated tools.

Face to the implementation complexity, to reduce implementation cost, some geospatial content management systems (CMS) appear (see GeoCMS [5]). In the same idea, this paper proposes to investigate a model-based approach. Modelling is a new challenging approach of software engineering to master the complexity of platforms and to express domain concepts effectively [6]. The idea is to propose adapted model for each data management phase and then to replace the coding activity by a configuration activity, as in a CMS. The project is ambitious considering the following aspects:

- Format verification
- Standard application
- Data storage
- Data visualization and analysis

In this paper, we present first experiments showing the interest of the proposed approach.

Format verification

OceanSITES is a system for gathering and measuring scientific data especially for time series sites. The OceanSITES User Manual holds around 30 pages of constraints expressed in natural language. These constraints are of different forms: naming conventions, possible attribute values, constraints on dimension length and many others. As an example, a CDM *DataSet* should hold instances of *Dimension* called *TIME, LATITUDE, LONGITUDE* and an instance of *Variable* called *TIME*. The *Variable TIME* should be of *double* datatype. To check the data conformity to OceanSITES, a Java tool already exists [7]. Since constraints are not formalized and the tool is hand-coded, there is no guaranty on checking. Furthermore, for each standard, a particular tool should be developed. To avoid constraint edition ambiguity and to reduce conformity tool development, we proposed a dedicated rule-based *CdmCL* language [8].

Format correction

The conventions for CF metadata are designed to promote the processing and sharing of files created with NetCDF. CF proposes constraints in a less restrictive way than OceanSITES. CF is based on guidelines proposing different recommended approaches to express metadata. For example, an attribute axis may be attached to a coordinate variable and given one of the values X, Y, Z or T which stand for a longitude, latitude, vertical, or time axis respectively. Alternatively, the standard_name attribute may be used for direct identification. To consider variants on constraints, we first extend the *CdmCL* language. Then we propose a tool able to extract from a given NetCDF file, the potential coordinate variables (depending on axis, standard_name and unit). The User Interface proposes helpers to choose and to correct, if necessary, coordinates attributes, following the standard recommendations.

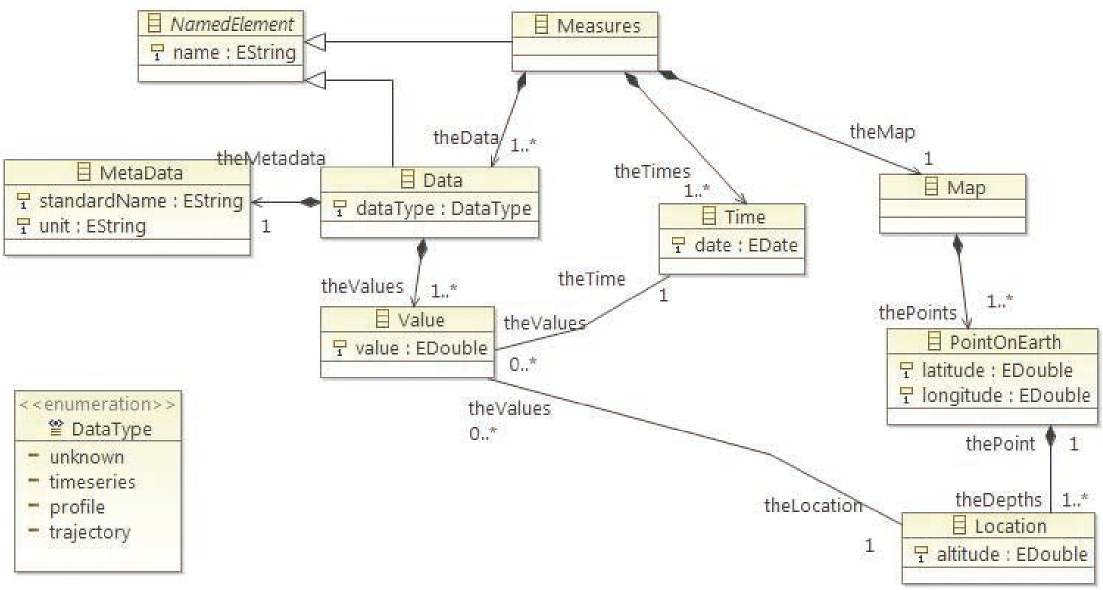


Fig. 1 - The proposed generic spatiotemporal data model.

Generic format and data storage

After correction, we consider data following the proposed unified object-oriented model of spatiotemporal data (see Fig. 1). Data are then stored locally or using a Big Data technology. For this purpose, we reuse the Kevoree project technology [9]. From object-oriented models, Kevoree offers an interface for object persistence and an API to store and to load objects from a data store. Different driver implementations allow to choose the most appropriate data backend for the application (local, database or big data technology) in a transparent way.

Tools

For data analysis tools, we propose to develop specific tool-oriented data model, considering the relevant part of the generic spatiotemporal data model, and tool-oriented parameters. For example, for a time series visualization tool, the *Map* information is reduced to one specific *Location*, and the model is enriched with User Interface parameters, as classically with a CMS. The model is then instantiated for each application. From a model, it is then possible to generate a dedicated visualization tool. For our experiments, we reuse classical visualization JavaScript libraries.

Conclusion

The paper proposes a model-based approach to improve tools development. We target tools dedicated to checking, correction, analyzing, storage and visualization of scientific data. The different proposed models are standard-independent and technology-independent. We experiment the approach on two different standards, using different implementation technologies. We are now working on extending the approach to provide new analyzing tools and to consider new applications.

References

Unidata, Common Data Model www.unidata.ucar.edu/software/thredds/current/netcdf-java/CDM/

NetworkCommon Data Form http://www.unidata.ucar.edu/software/netcdf/

The Climate and Forecast Conventions and Metadata http://cfconventions.org/

OceanSITES User's Manual, Version 1.2. http://www.oceansites.org/

GeoCMS https://github.com/dotgee/geocms

G. Mussbacher and all *"The Relevance of Model-Driven Engineering Thirty Years from Now"* Chapter of Model-Driven Engineering Languages and Systems LNCS Vol 8767, pp 183-200, Oct 2014

OceanSITES file format checker. http://www.coriolis.eu.org/Observing-theocean/ Observing-system-networks/OceanSITES/Access-to-data.

A. Ahmed, P. Vallejo, M. Kerboeuf, J.-P. Babau *"CdmCL, a Specific Textual Constraint Language for Common Data Model"* 14th International Workshop on OCL and Textual ModellingValencia, Spain. 2014

The Kevoree Modeling Framework "https://github.com/kevoree-modeling".

Improved statistical method for hydrographic climatic records quality control

Jérôme Gourrion, CORIOLIS R&D, Centre National de la recherche scientifique (France),
jerome.gourrion@ifremer.fr

Tanguy Szekely, CORIOLIS R&D, Centre National de la recherche scientifique (France),
tanguy.szekely@ifremer.fr

Climate research benefits from the continuous development of global in-situ hydrographic networks in the last decades. Apart from the increasing volume of observations available on a large range of temporal and spatial scales, a critical aspect concerns the ability to constantly improve the quality of the datasets.

In the context of the Coriolis Dataset for ReAnalysis (CORA) version 4.2, a new quality control method based on a local comparison to historical extreme values ever observed is developed, implemented and validated.

In this talk, we present how temperature, salinity and potential density validity intervals are directly estimated from minimum and maximum values from an historical reference dataset, rather than from traditional mean and standard deviation estimates. The concept is presented through examples and empirical validation results.

Such an approach avoids strong statistical assumptions on the data distributions such as unimodality, absence of skewness and spatially homogeneous kurtosis. As a new feature, it also allows addressing simultaneously the two main objectives of a quality control strategy, i.e. increase the number of good detections while reducing the number of false alarms.

The reference dataset is presently built from the fusion of 1) all ARGO profiles up to early 2014, 2) 3 historical CTD datasets and 3) the Sea Mammals CTD profiles from the MEOP database. All datasets are extensively and manually quality controlled. In this communication, the latest method validation results are also presented.

The method has been implemented in the latest version of the CORA dataset and will benefit to the next version of the Copernicus CMEMS dataset.

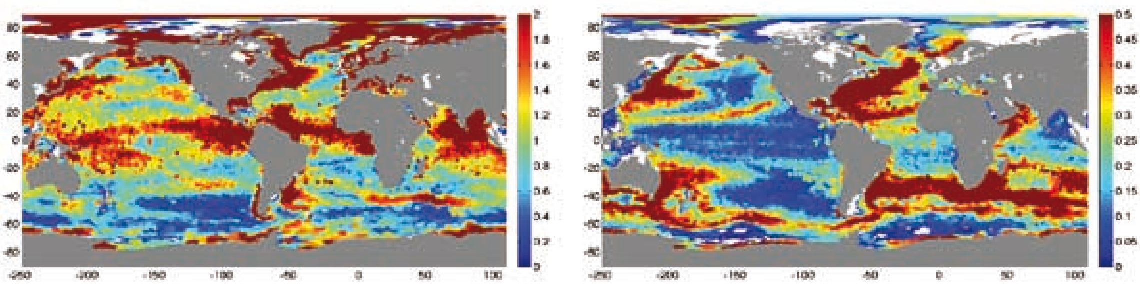


Fig. 1 - Examples of surface (left) and 500 dbar (right) minimum / maximum salinity interval width used as reference fields for outlier detection.

A single DOI for Argo; a generic approach to making datasets that grow and evolve with time citable on legacy infrastructure

Justin Buck, National Oceanography Centre (United Kingdom), juck@bodc.ac.uk
Thierry Carval, Institut français de recherche pour l'exploitation de la mer (France), thierry.carval@ifremer.fr
Thomas Loubrieu, Institut français de recherche pour l'exploitation de la mer (France), thomas.loubrieu@ifremer.fr
Frederic Merceur, Institut français de recherche pour l'exploitation de la mer (France), frederic.merceur@ifremer.fr

Abstract

The Argo dataset grows and evolves with time and changes in the expectations on the citation of Argo data and traceability of data citations has driven a 5 year effort to make Argo data citable via a single DOI. This has now been implemented by Ifremer on the Argo dataset for the first time using an approach that enables citation for the Argo data at monthly snapshots without requirement for significant enhancement to the Argo data infrastructure. The approach presented is readily applicable to other data infrastructures and enables Argo to partly meet the recommendations of the Research Data Alliance Dynamic Data Citation working group.

Paper

Unambiguous citation of data used in academic publications is crucial for the transparency and reproducibility of science especially when results are used as evidence to underpin national and international policy. Data citation of static datasets is well established and documented. However, the citation of dynamic data is an area of active development. When data are cited it should also be possible to unambiguously resolve the data. A key milestone in this development was when DataCite defined three methods for citing dynamic data; cite times slices, snapshots, or the dataset with an access date. This thinking was progressed by the RDA data citation working group in 2015 with the publication of an ideal method for citing dynamic data. These recommendations have not been implemented on real data yet and full implementation is not possible for legacy data systems where there are insufficient resources to upgrade software and infrastructure. In addition to this the syntax for including dynamic data information has not been agreed or ratified with DataCite and CrossRef. The IPCC 5th assessment used evidence from data collected from observing programmes, high impact reports such as this need the data that underpin contributing papers to be unambiguously citable.

To address outstanding dynamic data citation issue a single DOI has been implemented on the Argo dataset that enables long term resolution of previous versions of the Argo dataset with a monthly resolution. This has enabled the citation of Argo data through a single DOI (example: http://dx.doi.org/10.17882/42182). This DOI is associated to a default Landing Page. By default, this landing page would presents the last monthly snapshot with links for displaying all the previous monthly snapshots. Each of these snapshots will be associated to a unique key.

In this default landing page, it is possible to find:

- the general metadata
- a "How to cite" section
- a "Related publications" section automatically filled with Megan Scanderbeg bibliographic survey

some links to export the metadata in 3 different formats including the RIS format to include the citation in software like Endnote For dynamic data citation especially, a keystore is managed by the publisher infrastructure (Ifremer/Seanoe) so to handle snapshot citation for this same enriched DOI (example: http://dx.doi.org/10.17882/42182#41950). Then the landing page will be automatically adapted to offer metadata and data to the snapshot requested:

- The title, the date and the citation is adapted.
- The data section presents the snapshot as default but the link will still be available to list the others snapshots.

This solution is ready applied to other legacy data systems where the amount of available storage does not constrain the storage of monthly snapshots. A significant advantage of this solution is that is an economic way to implement dynamic data citation on legacy datasets where resource is insufficient for the full implementation of the RDA dynamic data recommendations.

Managing Australian Ocean Data Network (AODN) Vocabularies

Kim Finney, Integrated Marine Observing System (Australia), kimtfinney@gmail.com
Sebastien Mancini, Integrated Marine Observing System (Australia), Sebastien.Mancini@utas.edu.au
Roger Proctor, Integrated Marine Observing System (Australia), Roger.Proctor@utas.edu.au
Natalia Atkins, Integrated Marine Observing System (Australia), Natalia.Atkins@utas.edu.au

Using Controlled Vocabularies In The AODN Data Infrastructure

The Australian Ocean Data Network (AODN) comprises of data captured by Facilities of Australia's Integrated Marine Observing System (IMOS) and other organisations active in the coastal and marine realms. IMOS itself is a multi-institutional collaboration led by the University of Tasmania (UTAS). It deploys a wide range of observational equipment and has been capturing and publishing marine data since 2008.

These data are documented using an agreed community (ISO 19115 compliant) metadata schema and are served through a common infrastructure – the AODN Portal, established and hosted by IMOS. To enable the reliable discovery and consumption of AODN data via this infrastructure, dataset descriptions encompass controlled vocabularies. Portal data indexing and facet-based searching harness vocabulary terms to improve the precision and recall of search results. These same terms should also be used inline (e.g. within self describing netcdf files) and work is already underway within IMOS to establish automated procedures for checking Facility submitted NetCDF files for syntactic and semantic compliance.

Research Vocabularies Australia (RVA) - Semantic Tooling For Australia's Research Groups/Institutions

Until late 2015 the creation and management of AODN vocabularies was a labour intensive activity, involving relational databases, spreadsheets and custom scripting. Vocabulary publication was restricted to periodic posting of static ISO 19135:2005 registers and community participation in vocabulary development was restricted. Since then, a partnership between IMOS and the Australian National Data Service (ANDS), has delivered a national set of services that not only help the AODN with its vocabulary management, but any Australian agency wishing to create and publicly deploy vocabularies of any flavour.

These freely available ANDS hosted vocabulary services are bound together as a suite of interoperating tools, comprising of a skos editor (academic license of Pool Party Semantic Suite); a SPARQL-enabled semantic data store (Sesame); a REST-based API (SiSSVoc) and a Catalogue Portal (based on ANDS custom software). These tools have been deliberately loosely coupled so that inevitable advances in semantic technologies can be capitalised upon easily without breaking offered functionality. Vocabularies created using the Pool Party editor can be accessed directly from the ANDS Catalogue Portal, versioned, uploaded to the sesame store and then published. For human-centric use-cases published vocabularies can be searched and browsed in the Catalogue and downloaded via a user interface (see Fig. 1). For use-cases requiring machine-to-machine transactions there is a SPARQL end-point and an API for filtering a vocabulary to access individual terms. This API is also augmented by a user interface (UI).

During testing of the tool suite a number of limitations were encountered with the Pool Party technology, despite it being one of the best commercial editors available. To address some of these short-comings ANDS, with input from IMOS, has developed some stand-alone utilities that execute against the Pool Party data store, by-passing the Pool Party UI. For more intrinsic deficiencies ANDS and IMOS have been influencing the Semantic Web Company (Pool Party vendors) to improve Pool Party functionality for scientific and educational users. To date the company has been positive and responsive to suggestions.

Managing published vocabularies is performed through a ‘MyVocabs’ administration console which is part of the ANDS Catalogue. User accounts are provisioned using Australian Access Federation (AAF) credentials. AAF is a combination of technology and policy that provides a trust framework. Subscribers, which are largely the university and education sector, get a national single sign-on that allows individuals across many different organisations to collaborate and access online resources, within a trusted environment.

AODN Vocabularies In Action

Now that basic tooling is in place to create, manage and publish controlled vocabularies experimenting has begun to develop a shared governance model for vocabulary management amongst AODN partners. Greater participation by partners in curating and moderating community vocabularies should engender stronger ownership, expand the range and number of vocabularies and their terms and make it easier for data providers to deliver conformant metadata mark-up. How to best implement updates to metadata records submitted to the AODN central GeoNetwork-based metadata catalogue, when vocabularies are re-versioned is also under trial. The over-riding aim in these endeavours is to ensure that controlled vocabulary usage does not become a burden, or barrier to participation, for AODN data providers.

Importantly, the basic tooling required to create and use vocabularies, particularly in Linked Data scenarios is now in place. This means more time can be spent on developing innovative ways in which to use the vocabularies, rather than an excessive focus on their curation. To this end IMOS is currently working with a UTAS research team who are developing a national shelf rocky reef habitat classification scheme and associated spatial products. The intent is to develop the classification scheme using RDF and the ANDS tooling in a manner that adds value to the research and makes it easier for future reef data providers to embellish the national products. By using an encoding such as RDF it will be possible to align this classification scheme with past schemes no longer in vogue opening the path for an intelligent re-use of existing data.

(Near) real-time data publication for coastal and deep-sea observing system using OGC Sensor Web Enablement (SWE) standards

Elena Partescano, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), epartescano@ogs.trieste.it

Alberto Brosich, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), abrosich@ogs.trieste.it

Alessandra Giorgetti, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), agiorgetti@ogs.trieste.it

Marina Lipizer Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), mlipizer@ogs.trieste.it

Vanessa Cardin, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), vcardin@ogs.trieste.it

Introduction

The aim of this paper is to describe the system developed at OGS to share in (near) real-time validated data from two meteo-oceanographic buoys (MAMBO1 in the North Adriatic Sea and E2M3A in the South Adriatic Sea) using OGC Sensor Web Enablement (SWE) standards aligned with several Oceanographic communities (EUROFLEETS2, FixO3 , Jerico-Next, ODIP II, RITMARE and SeaDataNet). The Web Client developed by 52° North is adopted for data visualization and it is accessible from NODC/OGS web portal.

Marine (near) real-time data are collected in multiple data format and converted into homogeneous and standard formats, using XML and OGC (Open Geospatial Consortium) standards for data transport and representation. The adoption of Sensor Web Enablement (SWE) specifications enables real time integration of data and all metadata related to the measuring sensors, data processing and calibration, data quality control, allowing a full description of the available data.

Devices

Two marine observatories acquiring meteo-oceanographic data in (near) real-time are currently maintained by OGS: the meteo-marine buoy Monitoraggio AMBientale Operativo (MAMBO1), placed in the Gulf of Trieste and the E2-M3A located in the South Adriatic Sea. The coastal station is equipped with a meteorological station and multi-parametric probes. The deep-sea observatory system is hosting a meteo station aimed to collect air-sea interaction measurements and a deep mooring with sensors for physical and biochemical parameters. Sensors (listed in Table 1) and data are described using Sensor Model Language (SensorML) and Observations and Measurements (O&M) schemas, respectively.

Data-flow

The data management in (near) real-time, is composed by (Fig. 1): two meteo-marine stations; the RT-Loader (Real-Time Loader), permits to store into a relational database real-time heterogeneous data and to validate them applying a procedure to qualify the data values using DB-Validator (Database Validator); the DB (Relational Database); the Web Service, used for data distribution in NetCDF format and a RT-Web (Real-Time Web Interface) to discover and extract data.

Table 1 - List of sensors managed using SWE standards.

E2-M3A	
Oceanographic parameters	Sensor
Temperature and salinity	Seabird SBE37-SM
pHT	Sunburst SAMI pH
pCO2	Pro Oceanus CO2-Pro
pressure	Seabird SBE37-ODO
O2	Seabird SBE37-ODO
Temperature and salinity	Seabird SBE37-SM
Meteorological parameters	Sensor
Atmosferic pressure	Young 61202
Wind speed and direction	Young 04106
Air temperature	Young 41372
Relative humidity	Young 41372
Solar radiation	Eppley PSP
IR radiation	Eppley PSP

MAMBO1	
Oceanographic parameters	Sensor
Atmosferic pressure	Young mod. 61201
Air temperature	Young mod. 41372VC
Wind speed and direction	Young Wind Monitor
Relative humidity	Young mod. 41372VC
Temperature and salinity	Seabird SBE37-SM
pH	Sunburst SAMI pH Pro
pCO2	Pro Oceanus CO2 Pro
O2	Seabird SBE37-ODO
Temperature and salinity	Sea Bird16
O2	Sea Bird SBE43
pH	Sea Bird SBE18
Fluorescence	Wetlab Eco-AFL/FL
Turbidity	Wetlab Eco-NTU
Solar radiation	Wetlab Eco-PAR
Temperature and salinity	Sea Bird16
pH	Sea Bird16
O2	Sea Bird16

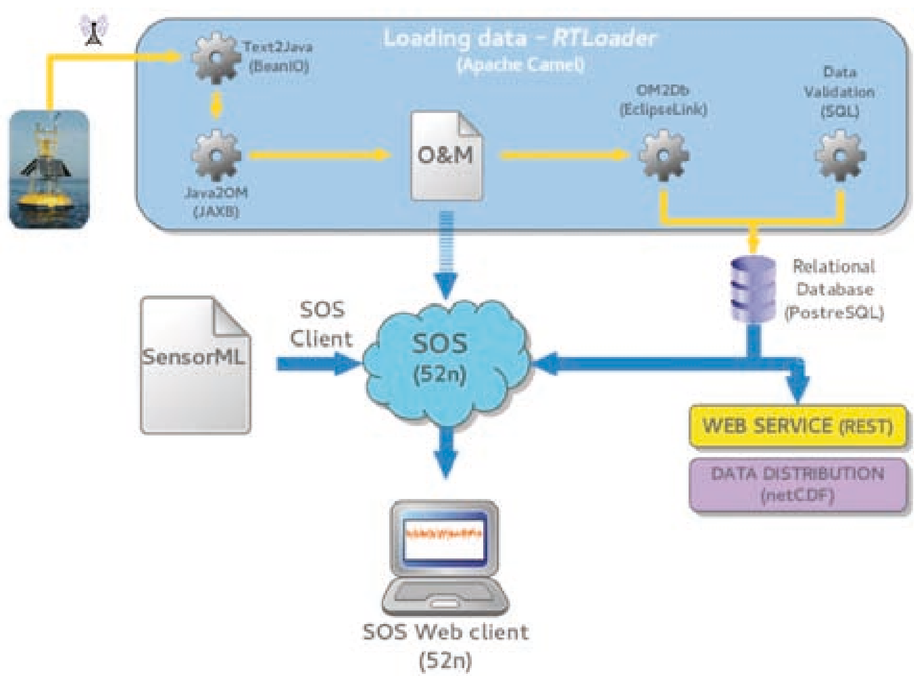


Fig. 1 - Real-Time Oceanographic Data Management System.

We adapted the previous management system to use the OGC standards such as Sensor Model Language (SensorML) and Observations and Measurements (O&M). We developed a SensorML and O&M profiles using BODC vocabularies, that guarantees both a standard way to exchange information (OGC's standard) and, at the same time, semantic interoperability. To realize it we use a Sensor Observation Service (SOS), developed by 52° North. In details, we use an application to insert measurements from relational database to SOS service and a SOS Client to load information about sensors. In the future, we will insert data directly from RT-Loader (dotted arrow in Fig. 1). At the end, a Web Client gives us the opportunity to visualize easily, data and metadata (Fig. 2).

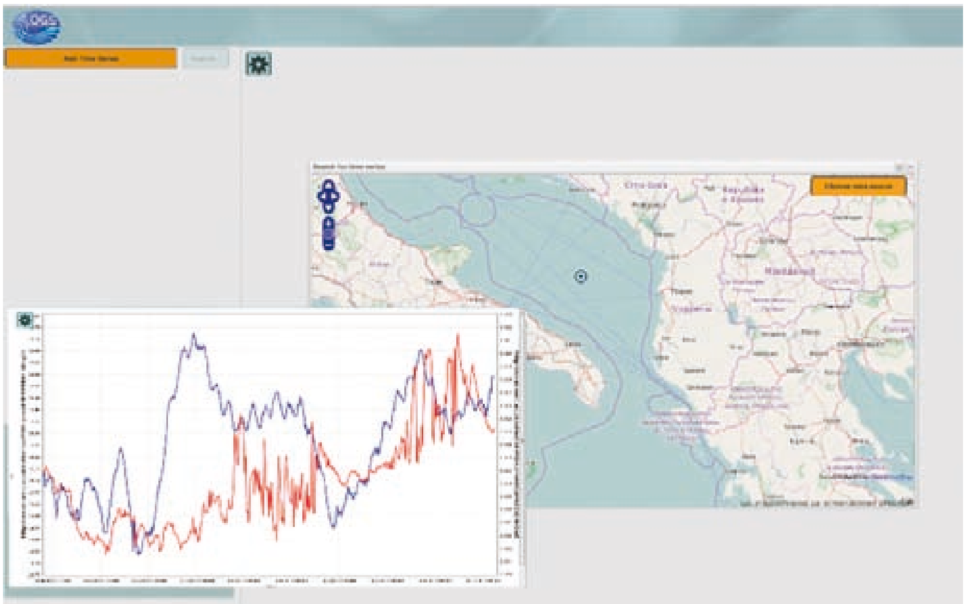


Fig. 2 - Sensor Observation Service - Web Client.

Conclusions and future development

The main objective of the approach presented in this paper is to share data in (near) real-time using OGC's Sensor Web Enablement (SWE) protocols to guarantee a standard system to share data, as well as, BODC vocabularies to ensure semantic interoperability. A critical analysis of this approach, point out as a positive aspect is the adoption of international standard for data, metadata and management procedures, giving the opportunity to connect with other marine data management infrastructures (e.g. SeaDataNet and EMODnet). On the other hand, considering that a data flow already existed, including a dedicated database, an effort has been necessary to include this new system. In addition, others negative aspects are: few choices of servers, client software and libraries available and the lack of a user friendly automatic management interface, for metadata compilation. But we expect that a direct link to newly developed sensors speaking O&M, soon or later, it will be available.

As future perspective, we plan, as first step, to adopt the same standards gradually for all observing system to manage others different kinds of platforms such as gliders and floats, using the same standard in (near) real-time and to have a common system where data and parameters acquired by different platforms can be jointly compared.

Making HELCOM Eutrophication Assessments Operational

Hjalte Parner, International Council for the Exploration of the Sea (Denmark), hjalte.parner@ices.dk

In 2014-2015 the project on making HELCOM Eutrophication Assessments operational (EUTRO-OPER) was carried out.

The project aimed toward a regularly updated high-quality thematic assessment of eutrophication status, produced through an operational and streamlined process. It was a continuation to the CORE EUTRO process, stemming from the EUTRO-, EUTRO PRO- and TARGREV projects, which have since 2005 developed the HELCOM core set of eutrophication indicators, with boundaries of good environmental status and assessment methodology, ending up in the latest update of eutrophication status in the Baltic Sea in 2007-2011.

The EUTRO-OPER project piloted the production of assessment products through efficient data flow processes. During the project, the entire assessment process, from monitoring and data aggregation to assessment calculation, was defined and documented, together with the protocols as well as responsibilities of QA/QC guidance and review. The project continued to improve the quality of the existing eutrophication status core indicators through enabling use of remote sensing and ship-of-opportunity data. Gaps in the present set of core indicator were investigated and new indicators were proposed. In addition, steps toward coordination of harmonizing the coastal and open sea eutrophication assessment were taken.

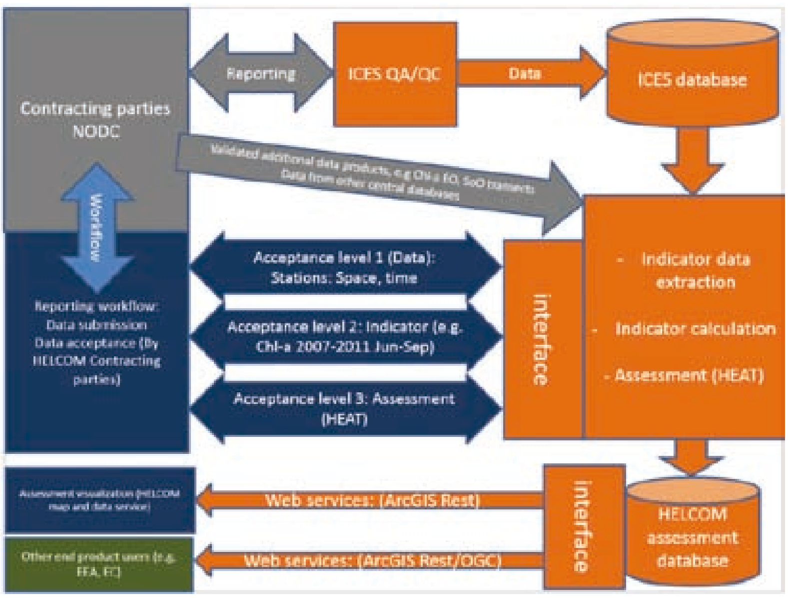


Fig. 1 - Data and information flow. The colour of the items indicate the actor/host: Gray = Contracting Parties, Blue = HELCOM portal (hosted at the Secretariat), Orange = ICES, Green = Other end-users.

HELCOM Eutrophication Assessment Tool (HEAT)

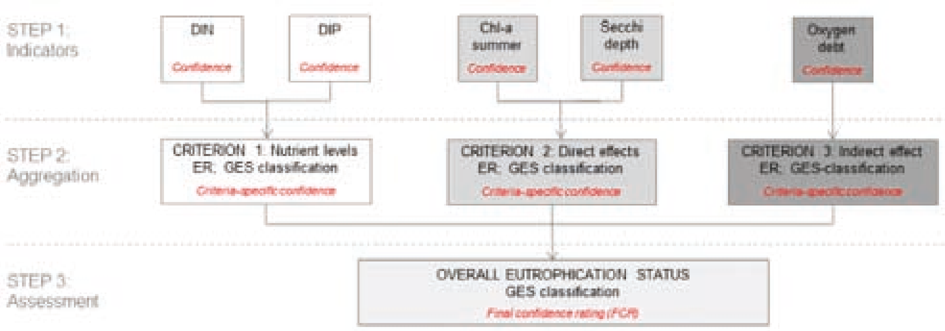
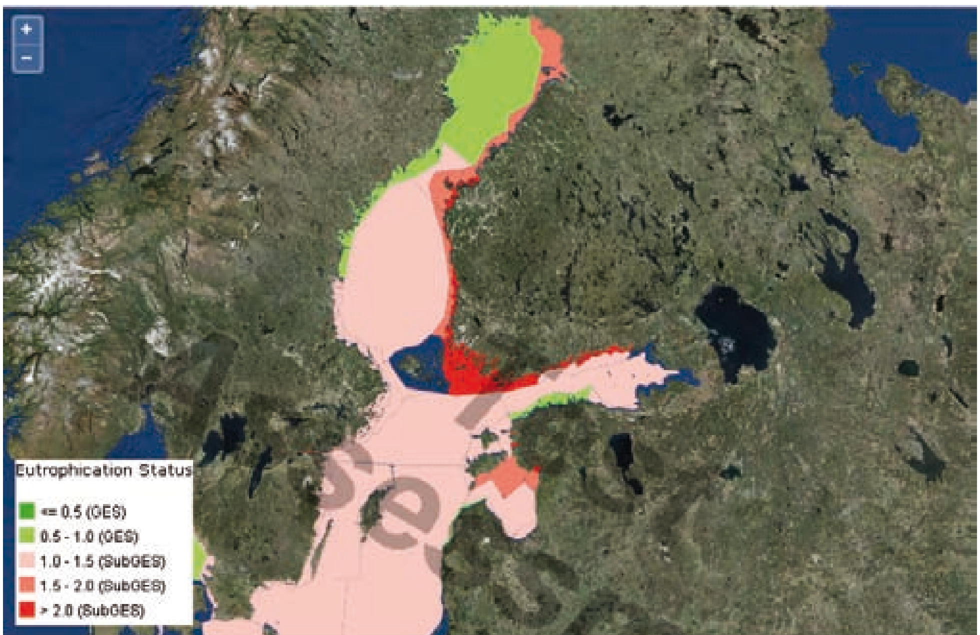


Fig. 2 - Schematic visualization of how the eutrophication status assessment is produced in open-sea assessment units using HEAT 3.0 and the core indicators.

The HELCOM Eutrophication Assessment Tool (HEAT) has now been made available online for the contracting parties which enable them in a transparent and streamlined way to evaluate the data behind the each of the indicators, view the result of the indicator calculations and finally see the full assessment result taken all indicators into account.



Assessment Unit		Period	ES	STD	N	ET	ER	ES-SCORE	ET-SCORE	I-SCORE
ID	Description									
SEA-001	Kattegat	20072011	7.67	0.90	128	7.60	0.99	100	50	75
SEA-002	Great Belt	20072009	6.57	1.03	6	8.50	1.29	50	50	50
SEA-004	Kiel Bay	20072011	5.74	0.94	45	7.40	1.29	100	50	75
SEA-005	Bay of Me..	20072011	5.11	0.83	103	7.10	1.39	100	50	75

Fig. 3 - The indicator view, showing core indicator Secchi depth as an example.

OceanBrowser: on-line visualization of gridded ocean data and in situ observations

Alexander Barth, University of Liege (Belgium), a.barth@ulg.ac.be
Sylvain Watelet, University of Liege (Belgium), swatelet@ulg.ac.be
Charles Troupin, Balearic Islands Coastal Observing and Forecasting System (Spain), ctroupin@socib.es
Alvera Aida, University of Liege (Belgium), A.Alvera@ulg.ac.be
Giorgio Santinelli, Deltares (The Netherlands), Giorgio.Santinelli@deltares.nl
Gerrit Hendriksen, Deltares (The Netherlands), Gerrit.Hendriksen@deltares.nl
Alessandra Giorgetti, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), agiorgetti@ogs.trieste.it
Jean-Marie Beckers, University of Liege (Belgium), JM.Beckers@ulg.ac.be

The ocean is generally a highly stratified fluid, where the seawater properties vary strongly with depth. While the ocean can be seen as a system made up of individual layers, the vertical movements (coastal upwelling due to wind for instance) and mixing couple the different layers. The ocean is also non-stationary as it varies on a broad range of temporal scales. Observational networks usually capture a limited range of these temporal scales. It is necessary to integrate the depth and time dimensions to provide a full 4D view of the ocean.

A web-based visualization allows the integration with other datasets, which facilitates the interpretation of the dataset in conjunction with results from different domains. For instance, physical and biogeochemical variables are strongly related due to the impact of ocean currents. The web interface should allow users to easily find a useful dataset and provide a mechanism to download it (entirely or a subset) and to be interoperable with other visualization platforms.

The web-interface OceanBrowser aims to respond to those challenges and requirements by implementing the following features:

- Horizontal sections of the 4-dimensional fields (longitude, latitude, depth and time) can be visualized at a selected depth layer and time. The climatological fields can also be interpolated and visualized on arbitrary vertical sections.
- The maps displayed in the browser are created dynamically and therefore several options are made available to the user to customize the graphical rendering of those layers. Layers can be displayed either using interpolated shading, filled contours or simple contours and several options controlling the colour-map are also available.
- The horizontal and vertical sections can be animated in order to examine the temporal evolution.
- Images can be saved in raster (PNG) and vector image formats (SVG, EPS and PDF). They can also be saved as a KML file so that the current layer can be visualized in programs like Google Earth and combined with other information imported in such programs.

- The underlying 4-dimensional NetCDF file can be either downloaded as a whole from the interface or only as a subset using the OPeNDAP protocol.
- The web-interface can also import third-party layers by using standard Web Map Service (WMS) requests. The user needs only to specify the URL of the WMS server and its supported version.

Interoperability depends critically on the use of open standards. Beside the web-standards, OceanBrowser uses OGC WMS and NetCDF. Limitations in the WMS standards to implement the defined requirements are also highlighted and the adopted extensions are discussed.

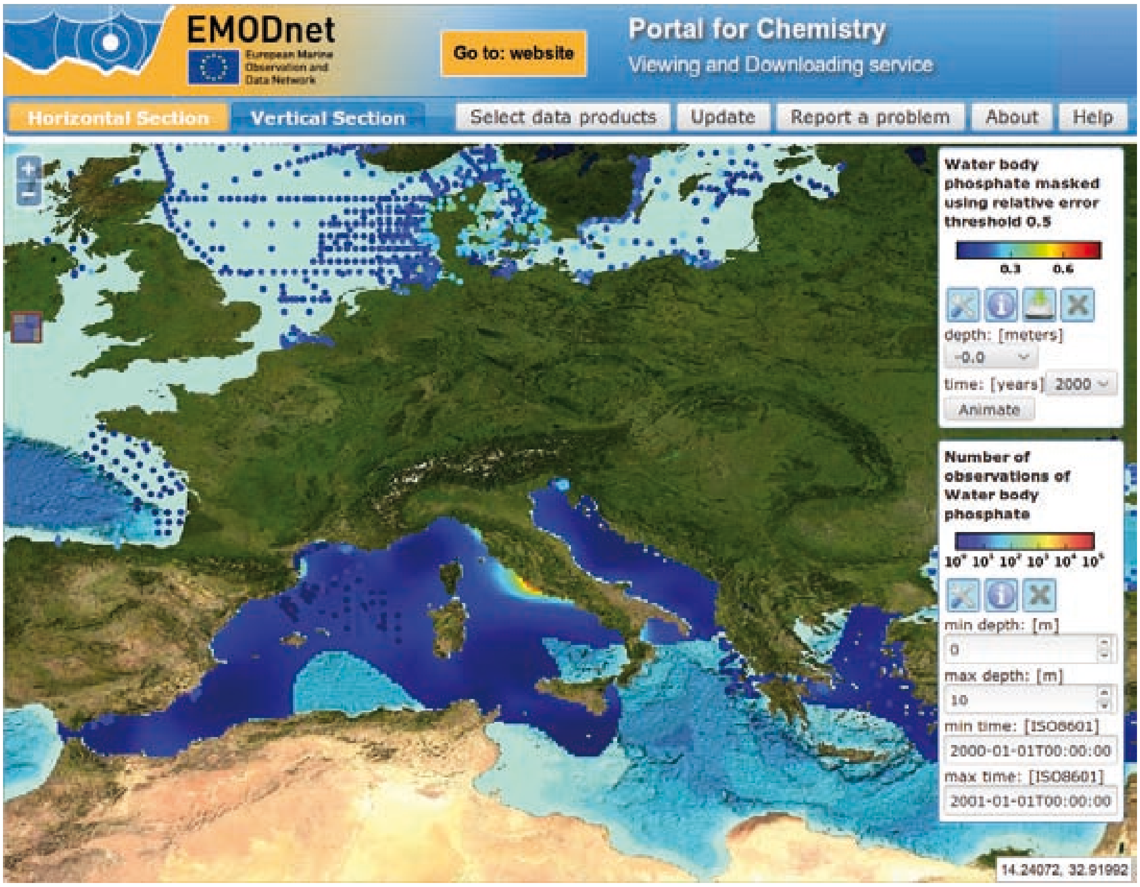
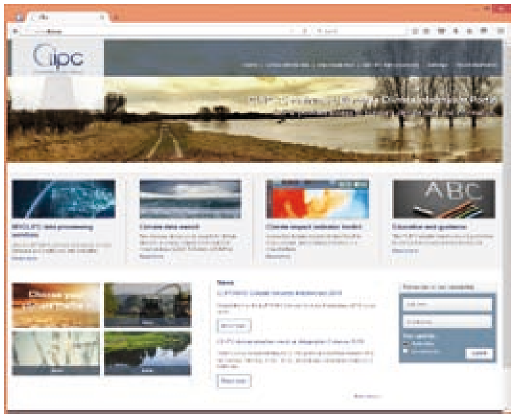


Fig. 1 - OceanBrowser installation for EMODNET Chemistry http://ec.oceanbrowser.net/emodnet/.

CLIPC: Tools and visualisation services for climate datasets to assess climate change

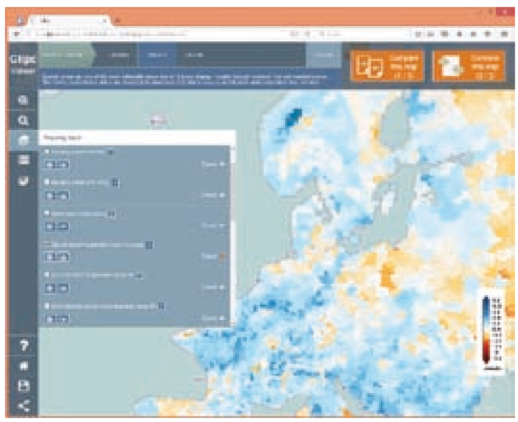
Peter Thijsse, Marine Information Service (The Netherlands), peter@maris.nl
Wim Som de Cerff, Koninklijk Nederlands Meteorologisch Instituut (The Netherlands), wim.som.de.cerff@knmi.nl



CLIPC Climate Information portal for Copernicus

The CLIPC project has developed an integrated platform of Climate Data Services to provide a single point of access for authoritative scientific information on climate variability and change, and climate change impacts. This project supports the Copernicus Earth Observation Programme for Europe, which will deliver a new generation of environmental data for Europe's citizens, decision-makers in the public and private sector, and academics.

The CLIPC portal: www.clipc.eu



Viewing and processing climate data

The CLIPC project aims to enhance access to climate and climate impact information, and support the translation, post-processing and visualisation of climate data in order to tailor climate and climate impact information to user needs.

The presentation will focus on one of the main components of the CLIPC portal to achieve this climate data visualisation: The CLIPC climate indicator toolkit. This toolkit has been developed technically by MARIS in close cooperation with KNMI and is based on concepts of other partners in the project. This toolkit includes:

- A map viewer to view the climate datasets and impact indicators (large NetCDF files) in a map interface with quick response
- Per dataset options to select the timestamp, view description, uncertainty, timeseries, histograms, change legendtype
- An explorative tool for users to compare and combine indicators, along with guidance on how to interpret the produced combinations

- A MyCLIPC space where users have a personal environment to:
  - Store datasets and map viewer settings
  - Process datasets to create their own indicators
  - View previous jobs.

The data and information presented meets the highest feasible standards in terms of credibility (evidenced by clear metadata), legitimacy (through an internationally supported mechanism) and linked to user needs.

Technically the toolkit software runs on Open Standards where possible: OpenLayers 3, WMS, WCS, (Geo)Json, (Py)WPS, webservices, WCS, Geonetwork and PHP are the main software libraries used. The user interfaces for CLIPC run on MARIS servers and are separated from the processing and storage server side. Currently the strong KNMI (Dutch MetOffice) servers are used for processing and storage, but this could well be expanded towards a full cloud solution.

Guidance for different users is important

The CLIPC aims to enhance access to climate and climate impact information for a variety of users including climate scientists, impact researchers, intermediary (or boundary) organisations (e.g. consultants, EEA) and societal end users (including policy makers and NGOs). Different users have different requirements, needs and skills, and usually look for more than just mere access to data – they need information that is relevant, robust and legitimate.

A unique aspect of CLIPC is the expert/or qualitative uncertainty analysis of impact indicators for which a quantitative analysis is not always possible.

The portal has several types of guidance, including:

- Use /user case descriptions to illustrate what specifically can be done via the portal;
- Traffic lights and other types of 'warnings' / recommendations;
- Glossary of terms used in the portal, taken from trustworthy sources like IPCC.
- FAQ's;

Parallels to marine data domain

There are many parallels between meteorological data and marine data, especially when it comes to metadata, large quantities of data, processing capacity for e.g. climatologies, and demand for services on top of the data with sufficient guidance to users.

Metadata is important, as well as harmonisation of metadata and data exchange formats. In CLIPC BODC has been involved to use and expand the NERC Vocabulary Server to facilitate the metadata harmonisation and support data discovery. Processing capacity next to the data in the form of cloud solutions are foreseen (and demanded by users). A toolbox approach seems very promising, especially when developed in open standards so services can be used by several applications. Googlemaps has paved the way, and since many years viewing services are a « must » in order to convince users. WMS, WCS combined with WPS processing and OpenLayers3 provide a strong combination. Guidance is key in all communities. Users have to understand what is quality of all input data, the processing services, and output data in order to use the results.

Data processing and visualization at the Balearic Islands Coastal Observing and Forecasting System (SOCIB)

Charles Troupin, Balearic Islands Coastal Observing and Forecasting System (Spain), ctroupin@socib.es
Joan Pau Beltran, Balearic Islands Coastal Observing and Forecasting System (Spain), jbeltran@socib.es
Biel Frontera, Balearic Islands Coastal Observing and Forecasting System (Spain), bfrontera@socib.es
Sonia Gómara, Balearic Islands Coastal Observing and Forecasting System (Spain), sgomara@socib.es
Miquel Gomila, Balearic Islands Coastal Observing and Forecasting System (Spain), mgomila@socib.es
Andreas Krietemeyer, Balearic Islands Coastal Observing and Forecasting System (Spain), akrietemeyer@socib.es
Cristian Muñoz, Balearic Islands Coastal Observing and Forecasting System (Spain), cmunyoz@socib.es
Miquel Àngel Rújula, Balearic Islands Coastal Observing and Forecasting System (Spain), mrujula@socib.es
Imma Serra, Balearic Islands Coastal Observing and Forecasting System (Spain), iserra@socib.es
Joaquín Tintoré, Balearic Islands Coastal Observing and Forecasting System (Spain), jtintore@socib.es

Multi-platform observing system

The Balearic Islands Coastal Ocean Observing and Forecasting System (SOCIB, http://www.socib.es), is a multi-platform Marine Research Infrastructure that provides free, open, standardized and quality-controlled data from near-shore to the open sea.

In order to properly capture oceanographic processes taking place at different spatial and temporal scales, a multi-platform approach is necessary. In our case, this involves a coastal research vessel, a high-frequency (HF) radar system, weather stations, tide gauges, moorings, drifting buoys, ARGO profilers, and gliders (autonomous underwater vehicles). Recently the system has recently begun incorporating oceanographic sensors attached to sea turtles, providing trajectories provided by the animals. High-resolution numerical models provide forecast for hydrodynamics (ROMS) and waves (SAPO).

Data processing

The Data Center is responsible for all the steps between the data acquisition to their visualization and access by the users. In order to cope with a wide range of platform, automatic management and processing are necessary.

In this work we will present some the applications developed to perform the oceanographic data management. The main tools are:

- *Instrumentation,* a database that contains the inventory of materials, the activities performed with them and the processing applied on the collected datasets.

- *Processing,* an application designed to extract meta-data of the deployed equipment from *Instrumentation* and to perform the data ingestion, processing, quality control and standardization.
- Glider toolbox (https://github.com/socib/glider_toolbox), a complete set of MATLAB/Octave scripts that automates glider data processing function, including thermal lag correction, quality control and graphical outputs.

Visualization

The general objective is to allow the scientific and the general public to visualize and explore the data without having to download them, through a set of tools, among which:

- DAPP (Deployments application, http://apps.socib.es/dapp/), a web application to display information related to mobile platform trajectories.
- LW4NC2 (http://thredds.socib.es/lw4nc2), a web application for multidimensional (grid) data from netCDF files (numerical models, HF radar).
- LEAFLET TIMEDIMENSION (https://github.com/socib/Leaflet.TimeDimension): a free, open-source Leaflet.js plugin that enables visualization of spatial data with a temporal dimension.

Applications

Based on the available data and using a set of web services, several applications were build:

- SEABOARD (http://seaboard.socib.es), a dashboard combining different sources of information in real time for different types of users.
- Smart-phone apps to access data, platform trajectories and forecasts in real-time.
- "Medclic: the Mediterranean in one click" (http://www.medclic.es/en/), a web dedicated to the Mediterranean Sea monitoring, with scientific and an outreach components.

Oceanographic system to control the impact of construction works in the sea in a sensitive coastal area - Case Study

Vlado Dadić, Institute of oceanography and fisheries (Croatia), dadic@izor.hr
Damir Ivanković, Institute of oceanography and fisheries (Croatia), ivank@izor.hr
Stipe Muslim, Institute of oceanography and fisheries (Croatia), smuslim@izor.hr
Dalibor Jelavić, Institute of oceanography and fisheries (Croatia), jelavic@izor.hr
Gordana Beg Paklar, Institute of oceanography and fisheries (Croatia), beg@izor.hr

The construction work at the sea close to the protected natural area can have a very bad influence on the environmental conditions in that area. One of the typical cases of possible negative influence on the very sensitive protected natural area Pantan is construction works on the bridge between land and Island of Čiovo in the town of Trogir, Adriatic Sea. The one km long bridge is being built at a relatively shallow water with a maximum depth of 5 meters so that the structure of the bridge rests on eight supporting pillars buried in the bottom of the sea. Bridge is located very close (500 meters) to the protected area Pantan that belongs where valuable marine and land flora

Fig. 1 - Location of bridge, Pantan area and measuring.

and fauna resides. In specific oceanographic conditions, sediments from the sea bottom during excavation can be easily transported towards protected area of Pantan and badly influence on the flora and fauna. Therefore, the contractor as part of the license for the construction of the bridge has to perform “on-line” monitoring of standard meteorological and oceanographic parameters at the station positioned close to the middle of the bridge (Fig. 1).

Data acquisition and transmission from station to the management centre must be performed at one-minute intervals by sending automatic warning and SMS message. This short interval is needed to have an ability of quick response on oceanographic condition. According to licence rules, reconstruction of currents field in the wide area of the bridge and calculation of sediments transportation to the protection area of Pantan must be done by its measuring data and numeric model. Beside this, control measurements of the optical parameters and chlorophyll_a have be performed every 15 days at the 8 stations in the wide are of the bridge including protected area of Pantan. Basically, the automatic monitoring system must instantly warn bridge authorities to stop construction work in the case when sediment transportation can reach the area of Pantan.

For these purposes, IOF was developed automatic measuring warning system for early warning of the management staff of the bridge construction, which includes automatic measuring oceanographic station with data transmission in real time to the RDBMS located in the Institute, where data are processed, assimilated into the numerical model.

Besides accessories (anchoring system, night light, solar cells, etc.) measuring station consists of embedded processor with GSM module and two sensors: Vaisala WXT520 meteorological multiparameter sensor (wind speed and direction, air temperature, humidity and pressure, and rain) mounted at the mast three meters above sea surface and RDI ADCP Doppler acoustic currents profiler (u, v and z current components of 6 cells along water column from the bottom to the surface). Output results from measuring system can be retrieved from IOF web page: http://faust.izor.hr/automjerenja/ciovo?

Results

All acquired data is stored locally in SD card and automatically sent to the MEDAS database located in the Institute of oceanography and fisheries by FTP protocol. Upon finishing data uploading in database automatic procedure starts and validates received data mainly correcting values outside of the broad range of specific parameters. For data visualization and data quality assurance High Charts and JavaScript API is used for presentation of all parameters except for visualization of multilevel sea currents, for which special JavaScript procedure was developed (Fig. 2).

Next step is starting of procedure for testing of meteorological and oceanographic conditions in the area under consideration from acquired data and in advance defined an arbitrary number of rules, which can be adjusted via web pages. These rules define maximal and average values and directions of wind and currents in all measuring layers and their duration, which represent the edge between conditions suitable for construct works or not. If averaged measured values exceed values in predefined arbitrary rules table, the database automatically sends a warning to the management of bridge construction to stop with work by sound flashing red header and alarm sound at the web page at the local computer and send SMS message to responsible persons.

Wind and currents data from Station are automatically uploaded as input data in numerical model ROMS (physical dynamics in the sea) and its output results is upload to SIM model (for

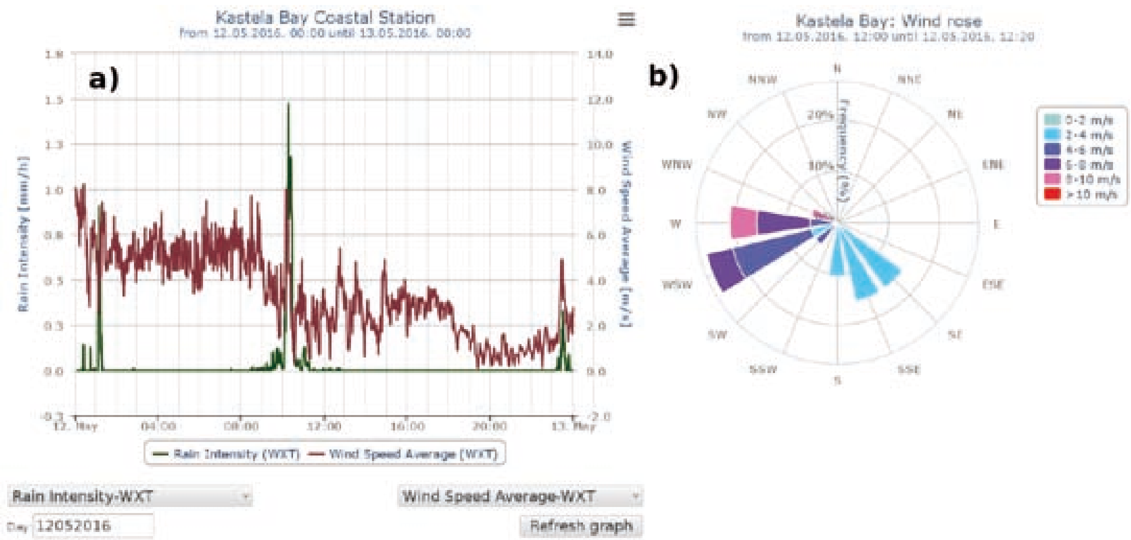


Fig. 2 - Example of view of two measured parameters and wind speed and direction (wind rouse).

simulation of sediments movements), which is operated for the west part of the Kaštela Bay (Fig. 3).

For the period of the system operation (from July 28, 2015 to May 13, 2016) more than 99.93 % of measurements data have been successfully transmitted and processed in receiving centre. In this period alarm was fired in 333 times with average duration about 91 minutes. According the alarm about 7.1 % of time was not suitable for construction works in the sea. These events occur due the fact that deeply and wide part of the Kaštela Bay ended by shallow and narrow channel at the west, in which during pass, stronger winds generate relatively strong currents, which can transport sediments from area of bridge to the protected area of Pantan in a short time (less than half hour).

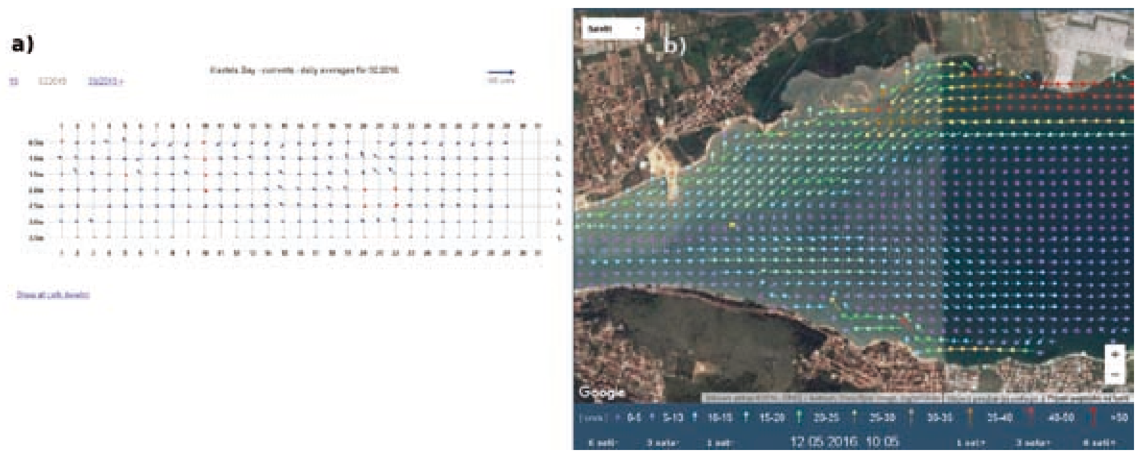


Fig. 3 - Daily average of the sea currents measured at the Station (a)
and sea currents field reconstructed by numerical model (b).

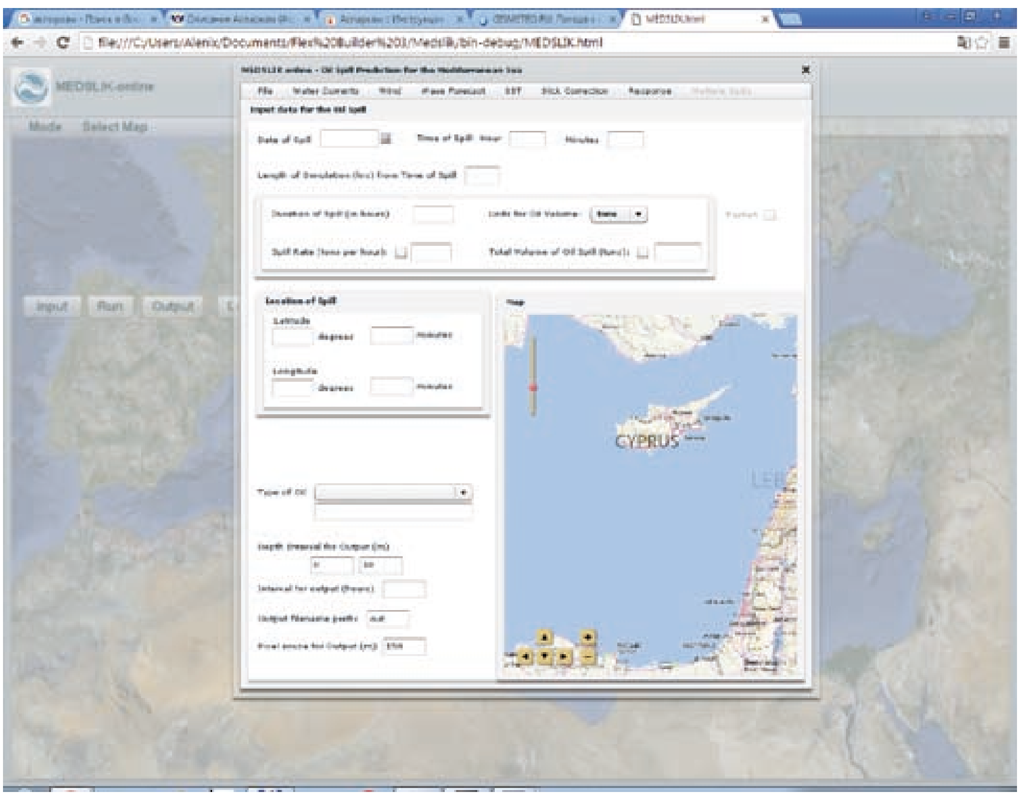
Web interface for the Oil Spill prediction software

Elena Zhuk, Marine Hydrophysical Institute RAS (Russian Federation), alenixx@gmail.com
George Zodiatis, Cyprus Oceanography Centre, UCY (Cyprus), gzodiac@ucy.ac.cy
Stavros Stylianou, Cyprus Oceanography Centre, UCY (Cyprus), stylians@ucy.ac.cy
Robin Lardner, Cyprus Oceanography Centre, UCY (Cyprus), lardner@cytanet.com.cy

The well known MEDSLIK oil spill model has been developed as a powerfull desctop software[1]. The MEDSLIK first web interface was developed and presented at IMDIS-2013[2]. The current paper deals with the improvements of the MEDSLIK web interface.

The online-MEDSLIK system allows authenticated users to use the main functions of the MEDSLIK model via an interactive and user friendly web application without the need of installing any software locally. The application offers three main functionalities:

- Input interface.
- Simulation.
- Visualization interface.



System architecture

The web application is based on a number of different technologies, such as:

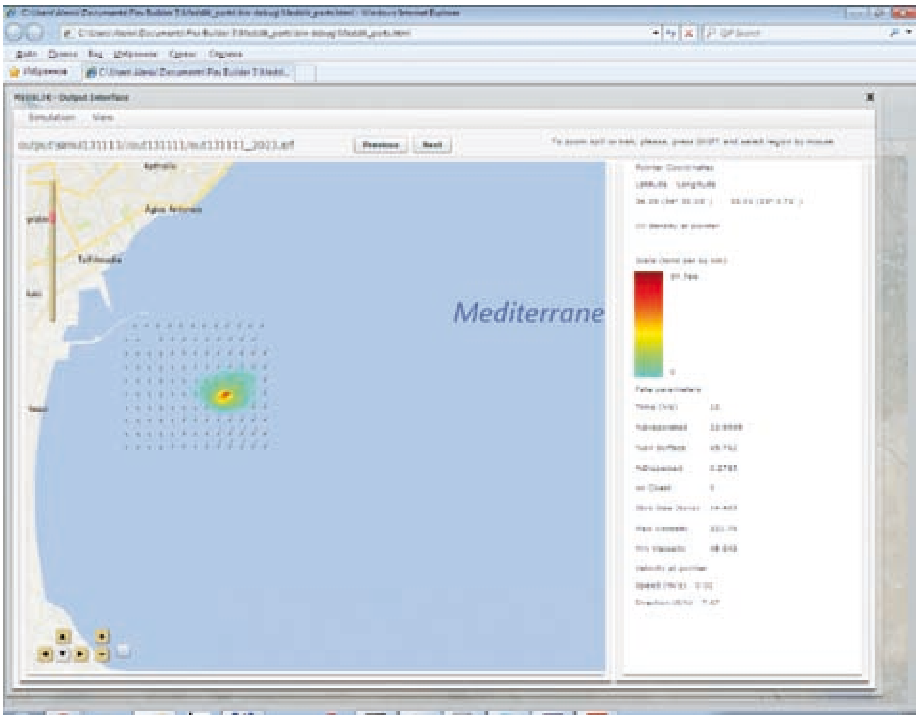
- Flex Rich Internet application used for the presentation of the input interface.
- PHP used for providing data exchange between Client and Server parts.
- MapServer chosen to be a map service.
- Python used for result processing.

The input interface was developed to be a similar to the MEDSLIK descktop interface. It gives opportunity to input all necessary oil spill parameters (Fig. 1) and run oil spill simulation. Each user owns a directory at the online MEDSLIK server. This directory keeps the user's input data and the simulation results which can be visualized.

The two last technologies consistitute the improvemant of the online MEDSLIK system, both allow to process and visualize the results more effectively. To perform oil spill prediction MEDSLIK uses CYCOFOS and Copernicus data.

Also the online-MEDSLIK was adapted to include high resolution domains for 3 ports in Cyprus. The Fig. 2 presents the result of oil spill prediction in the broader area of the Limassol Port, located at the souther coast of Cyprus.

In future, other regions such as Black Sea, will be included in MEDSLIK online.



References

LARDNER, R., ZODIATIS, G., HAYES, D., PINARDI, N., 2006. *"Application of the MEDSLIK Oil Spill Model to the Lebanese Spill of July 2006"*. European Group of Experts on Satellite Monitoring of Sea Based Oil Pollution. European Communities.

ZHUK E., ZODIATIS G., STYLIANOU S., LARDNER R., 2013. *"A web-based GIS for oil spill prediction"*. IMDIS 2013 Book of Abstracts. Vol 54. p.221.

POSTERS

A Web application to publish R scripts as-a-Service on a Cloud computing platform

Gianpaolo Coro, ISTI Centro Nazionale delle Ricerche (Italy), gianpaolo.coro@isti.cnr.it
Giancarlo Panichi, ISTI Centro Nazionale delle Ricerche (Italy), giancarlo.panichi@isti.cnr.it
Pasquale Pagano, ISTI Centro Nazionale delle Ricerche (Italy), pasquale.pagano@isti.cnr.it

Prototype scripting is the base of most models in computational biology and environmental sciences. Scientists making prototype scripts (e.g. using R and Matlab) often need to share results and make their models used also by other scientists on new data. To this aim, one way is to publish scripts as-a-Service, possibly under a recognized standard (e.g. the Web Processing Service of the Open Geospatial Consortium). Unfortunately, prototype scripts are not generally meant to be transformed into Web services, which require managing multi-tenancy, concurrency etc. Often, porting prototype scripts to more efficient programming languages is not affordable, because this operation demands for time, competencies and money. For this reason, Web services are becoming smart enough to integrate prototype scripts directly and possibly make them run efficiently (e.g. WPS4R)[1].

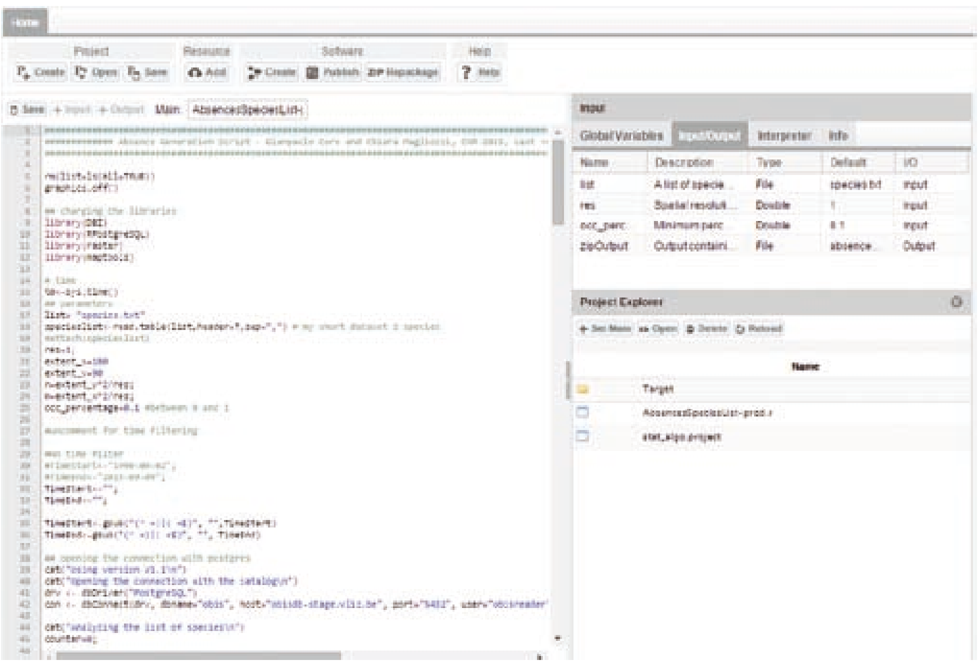


Fig. 1 - Interface of the D4Science Statistical Algorithms Importer.

1 52North, WPS4R for creating WPS processes based on annotated R-scripts, https://wiki.52north.org/bin/view/Geostatistics/WPS4R.

In this paper, we present an interface (Statistical Algorithms Importer, SAI) that allows scientists to easily and quickly import R scripts onto a distributed e-Infrastructure, which publishes the scripts as-a-Service and manages multi-tenancy and concurrency. Additionally, it allows scientists to update their scripts without following long software re-deploying procedures each time. SAI relies on the D4Science e-Infrastructure, a distributed computer system supporting large-scale resource sharing and Cloud computing, via the definition of Virtual Research Environments (VREs). VREs define groups of scientists working together in the same domain and are endowed with social networking and collaborative facilities. VREs are the main collaboration and sharing facilities offered by D4Science. SAI produces algorithms that run on the D4Science Cloud computing platform and are accessible via the WPS standard. This platform can manage multiple versions of the R interpreter by selecting the most appropriate execution environment given the script's requirements. SAI has been developed in the context of European projects having fisheries, biodiversity and environmental science experts as target users (BlueBRIDGE, i-Marine, D4Science).

The SAI interface (Fig.1) resembles the R Studio environment, a popular IDE for R scripts. Our aim was in fact to make SAI friendly to our script providers. The Project button allows creating, opening and saving a working session. A user uploads a set of files and data on the workspace area (lower-right panel). Upload can be done by dragging and dropping local desktop files. As next step, the user indicates the "main script", i.e. the script that will be executed on D4Science and that will use the other scripts and files. After selecting the main script, the left-side editor panel visualises it with R syntax highlighting and allows modifying it. Afterwards, the user indicates the input and output of the script by highlighting variable definitions in the script and pressing the *+Input* (or *+Output*) button: behind the scenes the application parses the script strings and guesses the name, description, default value and type of the variable. This information is visualised in the top-right side *Input/Output* panel, where the user can modify the guessed information. Alternatively, SAI can automatically fulfill the same information based on WPS4R annotations in the script. Other tabs in this interface area allow setting global variables and adding metadata to the process. In particular, the *Interpreter* tab allows indicating the R interpreter version and the packages required by the script and the *Info* tab allows indicating the name of the algorithm and its description. In the *Info* tab, the user can also specify the D4Science VRE the algorithm should be available to.

Once the metadata and the variables information has been fulfilled, the user can create a D4Science as-a-Service version of the script by pressing the *Create* button in the *Software* panel. With the term "software", in this case we mean a Java program that implements an as-a-Service version of the user provided scripts. The Java software contains instructions to automatically download the scripts and the other required resources on the server that will executed it, configure the environment, execute the main script and return the result to the user. The computations are orchestrated by the D4Science Cloud computing platform that ensures the program has one istance for each request and user. The servers constitute a distributed services system that manages concurrent requests by several users and executes code in a closed sandbox folder, to avoid damage caused by malicious code. Based on the SAI Input/Output definitions written in the generated Java program, D4Science automatically creates a Web GUI. In the creation phase, SAI checks the presence of at least one input and one output, the algorithm name, its description and the interpreter's version. A package is created under the *Target* folder that includes a Java archive

(JAR), metadata information and a zip file containing the script and the accessory resources. By pressing the *Publish* button, the application notifies D4Science system administrators that a new process should be deployed. Upon approval, the D4Science staff, assisted by appropriate tools, puts the JAR on the computational platform, but not the zip file. The obfuscated-compiled Java code inside the JAR contains an HTTP link to the zip file, which remains in the user's private project area. Thus, D4Science will not own the source code, which is downloaded on-the-fly by the computing machines and deleted after the execution. This approach meets the policy requirements of those users who do not want to share their code. The *Repackage* button re-creates only the zip package, which will substitute the previous file and inherit the same HTTP link. Thus, the JAR file on the computational platform will be using the new version of the script. The repackaging function allows the user to modify the script and to immediately have the new code running on the computing system. This approach separates the script updating and deployment phases, making the script producer completely independent on e-Infrastructure deployment and maintenance issues. However, deployment is necessary again whenever Input/Output or algorithm's metadata are changed.

To summarise, the SAI Web application relies on the D4Science e-Infrastructure and enables an R script, provided by a community of practice working in a VRE, with as-a-Service features. SAI reduces integration time with respect to direct Java code writing. Additionally, important functions that D4Science offers include (i) multi-tenancy and concurrent access, (ii) scope and access management through Virtual Research Environments, (iii) output storage on a distributed, high-availability file system, (iv) graphical user interface, (v) WPS interface, (vi) data sharing and publication of results, (vii) provenance management and (viii) accounting facilities.

References

Candela, L.; Castelli, D.; Manzi, A.; Pagano, P. *Realising Virtual Research Environments by Hybrid Data Infrastructures: the D4Science Experience. International Symposium on Grids and Clouds (ISGC)* 2014, Proceedings of Science PoS (ISGC2014) 022.

Coro, G., Candela, L., Pagano, P., Italiano, A., & Liccardo, L., 2015. *Parallelizing the execution of native data mining algorithms for computational biology.* Concurrency and Computation: Practice and Experience, 27(17), 4630-4644.

Coro, G., Gioia, A., Pagano, P., & Candela, L., 2013. *A service for statistical analysis of marine data in a distributed e-infrastructure.* Bollettino di Geofisica Teorica e Applicata, 54(1), 68-70.

DORIS Software : Processing and management of Sound Velocity Profiles for echosounding applications

Jean-Marc Sinquin, Institut français de recherche pour l'exploitation de la mer (France), jean.marc.sinquin@ifremer.fr

Xavier Lurton, Institut français de recherche pour l'exploitation de la mer (France), xavier.lurton@ifremer.fr

Christophe Vrignaud, Service Hydrographique et Océanographique de la Marine (France), christophe.vrignaud @shom.fr

Grégory Mathieu, Service Hydrographique et Océanographique de la Marine (France), gregory.mathieu@shom.fr

Hervé Bisquay, GENAVIR (France), herve.bisquay@genavir.fr

Context

Recent improvements in the performance of acoustic systems for hydrography and oceanography require today a high degree of accuracy in processing and managing sound velocity profiles (SVP). Current physical sensors for this purpose are usually sufficient for today's needs and are not addressed here. The SVP processing software tools available today and used currently for acquisition and post-processing, may result from developments by academic teams, operational services, or industry companies specialized either in sonar manufacturing or in software development; obviously these various products correspond to different requirements. There is certainly today a need for a unified tool fulfilling both operational and scientific needs, joining a rigorously scientific methodology with a user-friendly approach, proposing a number of functionalities for both at-sea acquisition and in-lab processing, and compatible with the greatest number of physical sensors and acoustic systems. Starting from this context analysis, Ifremer and SHOM have designed and developed such a SVP-processing software suite, according to their current operational and scientific needs, and based on their previous experiences in the fields of ocean physics, underwater acoustics, and software development for oceanography. Developed in the context of an academic and institutes partnership, *DORIS* is proposed today as an answer to the expectations of both operational and scientific teams.

General presentation

DORIS was designed as a part of an on-board configuration for at-sea operations; however it is usable as well in a data post-processing environment. Its general workflow articulates around five main functionalities (1):

DORIS is available as a freeware, directly downloadable as a trial version excluding the export of the processed profile. The full version, including all available functionalities, is provided under a free-of-charge concession of the software license. The source codes are not included in the delivery, and remain property of Ifremer and SHOM.

Fig. 1 - General structure of the *DORIS* workflow

Calculation of Regional QC ranges of temperature and salinity for Korean waters

Sung-Dae Kim, Korea Institute of Ocean Science and Technology (Rep. of Korea), sdkim@kiost.ac.kr
Hyuk-Min Park, Korea Institute of Ocean Science and Technology (Rep. of Korea), hyukmin@kiost.ac.kr

Establishment work of quality control procedures for the ocean data produced by a variety of national research projects was conducted in order to set up a national ocean data sharing system. Because 12 data items are being collected during the concerning projects, we set up 12 QC procedures for physical, chemical, biological and geological ocean data items (Table 1). At first, we prepared draft version of QC procedures after analyzing existing international and domestic QC methods. The proposed procedures were reviewed and revised by experts in the filed of oceanography. The QC procedure for temperature and salinity data was set up by referring the manuals published by GTSPP (Global Temperature and Salinity Profile Programme), ARGO and IOOS QARTOD (Quality Assurance of Real Time Ocean Data). It consists of 16 QC tests applicable for vertical profile data and time series data obtained in real-time mode and delay mode. Three regional range tests to inspect annual, seasonal and monthly variations were included in the procedure and three programs were developed to provide regional ranges to data managers. The programs can calculate upper limit and lower limit of temperature and salinity at depth from 0 to 1550m by using statistical data of World Ocean Atlas 2013 (WOA13) released by NOAA National Centers for Environmental Information (NCEI). When users input location, time (season or month) and depth to the programs, they extract mean, standard deviation and number of data from WOA13 data set and calculate regional ranges with three standard deviations. They display regional ranges calculated by statistical data of 3 kind of grid systems (5° grid, 1° grid and 0.25° grid) and finally provide recommendation ranges (Fig. 1). Users can use different range from suggested range if users know well the regional characteristics of the area, because it is known that the sparse data can cause bias of the statistic data in some areas around Korean peninsula. It is possible to provide better regional QC range if the experts who know well Korean waters examine data carefully and use more precise data. It is planed to produce new statistical data and regional range by analyzing unpublished new data and reanalyzing existing data.

Table 1 - Data items for QC procedure.

Part	Data Item
Physical	TS, ADCP, Wave
Chemical	DO, CO_2, Nutrient
Biological	Phytoplankton, Algae, Fish
Geological	Surface Sediment, Core Sediment, Shallow Seismic Wave

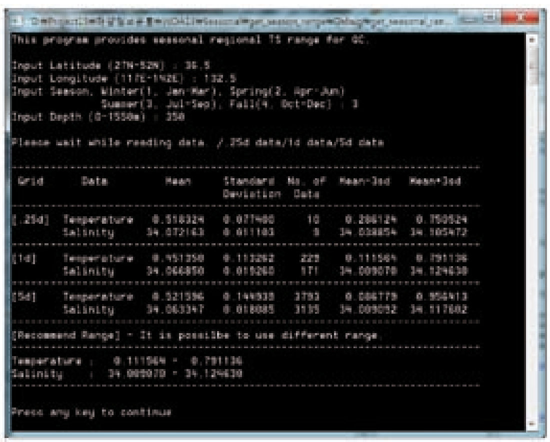


Fig. 1 - Program to calculate seasonal regional QC range.

Desktop QC software test version

Alexey Khaliulin, Marine Hydrophisical Institute RAS (Russia), khaliulin@yahoo.com
Eugeny Godin, Marine Hydrophisical Institute RAS (Russia), godin_ea@mail.ru
Andrey Ingerov, Marine Hydrophisical Institute RAS (Russia), ingerov@rambler.ru
Elena Zhuk, Marine Hydrophisical Institute RAS (Russia), alenixx@gmail.com
Elena Isaeva, Marine Hydrophisical Institute RAS (Russia), isaeva-ea@mail.ru
Isaac Gertman, Israel Oceanographic and Limnological Research (Israel), isaac@ocean.org.il

QC procedure description

The approaches on creating a unified data QC procedure had been proposed in the framework of various national and international projects. Our QC procedure is based on the experience of such projects as WOD; MEDAR/MEDATLAS II; NATO TU BLACK SEA, SEADATANET, Black Sea Scene, EMODNET, and others. We also paid attention to recommended by IOC UNESCO procedures for validation of oceanographic data. The way of increasing reliability of the quality control procedure is to use climatic arrays prepared by specialists in regional oceanography of the Black and Mediterranean seas as well as standard check arrays. The QC procedure includes metadata and data quality tests.

The metadata tests are:

- location check,
- date/time (including velocity and chronology) check,
- sea depth/last sounding value check.

The data tests include

- sounding value check (including order check),
- climatic check (if climatic arrays for the parameter are available),
- statistic check (if statistic arrays for the parameter are available),
- range check (if ranges for the parameter are available),
- fixing density inversions for hydrological data,
- fixing spikes (if corresponding thresholds for the parameter are available),
- gradient check (if corresponding thresholds for the parameter are available).

The QC software implements a procedure of data quality control for SDN and ODV spreadsheet files, assigns QC flags to metadata and data values, creates a log file with a list of possible data and metadata errors, provides metadata and data errors navigation and visualization, and gives a possibility of erroneous value correction.

The QC software offers interactive tools for setting of the QC procedure. They allow

- selecting QC tests and their parameters;
- adding shore line and relief data arrays for metadata quality control;

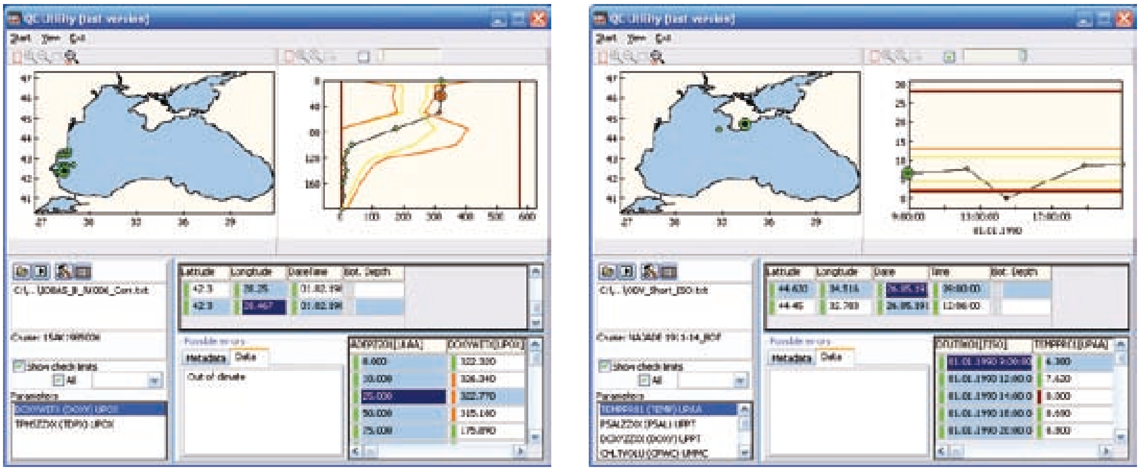


Fig. 1 - The examples of QC results visualization for profile and time series.

- adding sub-regions and check arrays for data quality control;
- adding connections between parameters and check arrays as well as units conversion coefficients.

Climatic and statistic arrays

The QC procedure uses mostly check arrays which are widely applied for data quality control in international oceanographic research projects. At the same time temperature, salinity and oxygen climatic arrays for the Black sea region were prepared by specialists of Marine Hydrophysical Institute. Climatic characteristics (mean and mean standard deviation) were calculated for temperature, salinity and oxygen. The sub-region scheme of 40×60 minute (40×40 mile) squares accounting spatial variability of hydrological parameters was used. The sub-regions for the Mediterranean Sea were taken from the MEDAR.MEDATLAS II and SeaDataNet2 projects. Certain squares with considerable spatial variability were divided into 4 parts and the climatic characteristics were calculated for each of them. The climatic arrays for temperature and salinity were prepared for each month and for oxygen – for each season. Moreover, a number of statistic annual averaged arrays for several parameters in the Mediterranean and Black seas were prepared.

BYTHOS online visualization of ocean forecast

Elena Zhuk, Marine Hydrophisical Institute RAS (Russian Federation), alenixx@gmail.com
George Zodiatis, Oceanography Center, UCY (Cyprus), gzodiac@ucy.ac.cy
Stavros Stylianou, Oceanography Center, UCY (Cyprus), stylianou.stavros@ucy.ac.cy
Andria Karaolia, Oceanography Center, UCY (Cyprus), andriak@ucy.ac.cy
Andreas Nikolaidis, Oceanography Center, UCY (Cyprus), andreas.nikolaidis.cut.cy@gmail.com

As a rich internet application the BYTHOS data base initially was developed to search, visualize and download oceanographic in-situ data [1], [2]. The recent improvements of the BYTHOS system is the ability to access and visualize the CYCOFOS forecasting data and overlay the forecasting fields with observing data. BYTHOS can be accessed at:

http://www.oceanography.ucy.ac.cy/BythosV2_ms/BythosV2_MapServer.swf

BYTHOS architecture

The software for data access and visualization consists of two parts, the server and the user interface.

The server includes an oceanographic database, a map service, php-modules providing interaction between server and client applications and python-modules processing the netCDF files with the CYCOFOS forecasts.

MapServer was chosen to be a map service.

To operate the database, MySQL data management system was chosen.

Presentation of data is achieved with the help of Flex Rich Internet application, used for the creation of BYTHOS front end. The front end is responsible for the search and retrieval of data from database, OPENDAP server and visualization through the map, currently served by MapServer.

BYTHOS in-situ database access and visualization

Data visualization is achieved through the use of Mapserver and is performed in three steps:

- The first step allows the user to perform search and filtering of data based on the type of data, the data center providing the data, the region of interest and a date range.
- The second step shows to the user a list of data collections available

Fig. 1 - Example of vertical profile and metadata from a glider station in the Levantine Basin.

based on the performed search. The user may then select one or more data collections to be plotted on the map.

- The final step is the visualization of the data collections on the map as points. For each point representing the geographical location, vertical profile for a given data set is plotted on a separated window. On-line interactive capability of the system allows to view the values of the data versus depth. Metadata information is also available for each data set with ability to download the data set for a single location/station or the entire set of the data for a given cruise.

BYTHOS forecasting data access and visualization

At present BYTHOS provides access and visualization of forecasting and satellite remote sensing data which are obtained from the CYCOFOS-Cyprus Coastal Ocean Forecasting System. CYCOFOS provides a variety of operational predictions such as high and medium resolution hydrodynamic ocean forecasts in the Levantine Basin and the Eastern Mediterranean Sea, sea state forecasts in the Mediterranean and the Black Sea, tide forecasts in the Mediterranean and the Black Sea, ocean remote sensing in the Eastern Mediterranean.

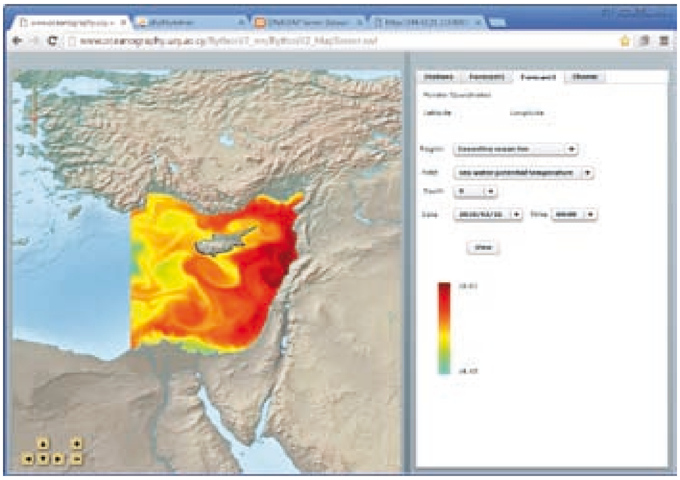


Fig. 2 - Sea surface velocity from CYCOFOS Levantine model.

The CYCOFOS data are kept at OPENDAP Server in netCDF format. To search, process and visualize it, dedicated php and python scripts were developed. The CYCOFO data visualization is achieved through Mapserver.

The BYTHOS interface to access the CYCOFOS forecasting data allows to search the forecasting parameters, by region, depth and time. Also it provides the capability to superimpose different forecasting and observing data, that can be used for complex analyze of sea basin aspects.

To access forecast map it is necessary to select Region, Field, Depth, Date and Time and click View. Then user can click at the map to see a profile.

References

Zodiatis George, Daniel Hayes, Stavros Stylianou, Iacovos Constantinou, Andria Karaolia, Ermis Koufou, Stavros Michael and Elena Zhuk, 2011. *"Bythos: A new online Data Management System"*, IOC/IODE 50th Anniversary International Conference, Book of Abstracts, p.48.

Zodiatis George, Daniel Hayes, Stavros Stylianou, Iacovos Constantinou, Andria Karaolia,Ermis Koufou, Stavros Michael, and Elena Zhuk, 2011. *"Towards a new generation of BYTHOS system for data management in the Mediterranean and Black Sea"*, EGU General Assembly, Geophysical Research Abstracts, Vol. 13, EGU 2011-10411-1.

WebODV: Providing Ocean Data View Services over the Internet

Sebastian Mieruch, Alfred Wegener Institute (Germany), sebastian.mieruch@awi.de
Reiner Schlitzer, Alfred Wegener Institute (Germany), reiner.schlitzer@awi.de

Ocean Data View (ODV) is a widely used software package for the analysis, exploration and visualisation of oceanographic and other environmental data. One strength of ODV is the flexible and efficient handling of diverse and large data collections often consisting of millions of stations on standard computer hardware. By design, ODV is a stand-alone application that has to be installed on local computers and operates on locally stored datasets. This concept is optimal for cases where datasets are not widely shared and do not change often. Also, because software and data both reside on the end-user's computer, ODV can be used during expeditions and in remote regions, when Internet connections are slow or non-existing. However, the stand-alone configuration is not ideal when users want to work with large community datasets that are changing constantly because of corrections of existing data or the addition of new data. The disadvantages are that users have to download the typically large datasets to their local computers and that the local copies of the datasets quickly become outdated because changes in the dataset master copy are not transmitted to the local copy. Users would have to download the modified master dataset again to synchronize with the latest version.

To address these issues we are planning to develop an on-line version of the ODV software called *WebODV*, and provide typical ODV functionality in form of Web services. This will be done in the context of the ongoing and planned EMODNET and SeaDataCloud projects. The idea is to provide clients with *ODV-like* interfaces in their web-browser and access datasets that are centrally maintained and administered on a server. Obviously, this setup requires Internet connection. Users will always work with the latest version of the datasets and will not have to download and store the data on the local computer. This will allow analysis and exploration of marine data on less capable hardware, such as tablets and smartphones.

The new *WebODV* service will be designed and implemented with emphasis on intuitive design and responsiveness. In a first step, we address the typical and basic user request of filtering existing datasets by region, time or data availability and extracting subsets of the data. Emphasis will also be placed on minimization of network traffic, speed and an intuitive web interface, adapted for the scientific user's needs. The way ODV is designed enables full functionality on the command-line and control via XML (Extensible Markup Language) view files, allowing usage of ODV on the server-side implementation. Thus, it is expected to offer a high degree of compatibility to the stand-alone version, which comprises e.g. graphical output, but also the exchange of XML files between ODV and *WebODV* is conceivable.

On the client-side we will use a range of functions of the jQuery framework to provide a fast dynamical interaction between the user and the *WebODV* front-end, which comprises e.g. image manipulation routines. The client-server interaction will be realised using AJAX (Asynchronous JavaScript and XML), which enables asynchronous data exchange between client and server, and thus dynamically changing web content without reloading the complete page.

WebODV will be dynamically evolving, and additional functionality will be implemented step by step. Examples are upload and download of data, on-line quality control, data plotting on maps, as time series or section profiles and more.

The long-term vision is to establish *WebODV* as an easily operated Web service for the analysis of oceanographic and other environmental data, without installing software or storing data locally. This will enable comfortable, reliable and efficient working conditions on large community data sets (and private datasets) from any computer of the world, which is connected to the Internet.

Moving controlled vocabularies into the semantic web

Rob Thomas, British Oceanographic Data Centre (United Kingdom), room@bodc.ac.uk
Alexandra Kokkinaki, British Oceanographic Data Centre (United Kingdom), alexk@bodc.ac.uk
Roy Lowry, British Oceanographic Data Centre (United Kingdom), rkl@bodc.ac.uk

One of the issues with legacy oceanographic data formats is that the only tool available for describing what a measurement is and how it was made is a single metadata tag known as the parameter code. The British Oceanographic Data Centre (BODC) has been supporting the international oceanographic community gain maximum benefit from this through a controlled vocabulary known as the BODC Parameter Usage Vocabulary (PUV). Over time this has grown to over 34,000 entries some of which have preferred labels with over 400 bytes of descriptive information detailing what was measured and how.

A decade ago the BODC pioneered making this information available in a more useful form with the implementation of a prototype vocabulary server (NERC Vocabulary Server - NVS) that referenced each ‘parameter code’ as a URL. This developed into the current server (NVS V2) in which the parameter URL resolves into a RDF document based on the SKOS data model which includes a list of resource URLs mapped to the ‘parameter’.

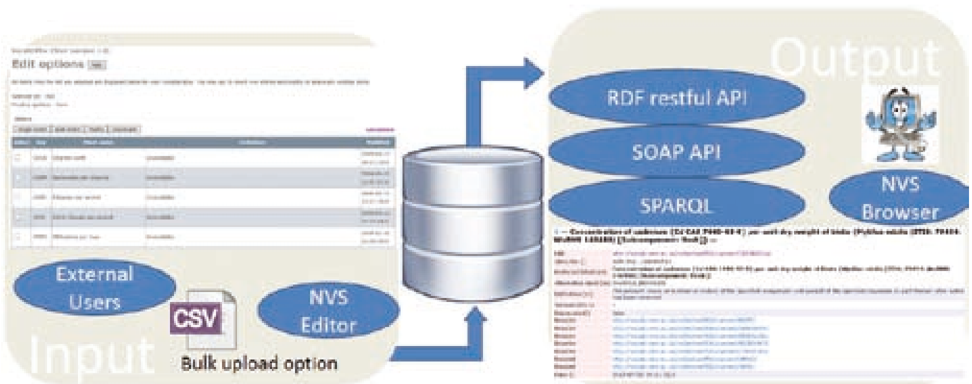


Fig. 1 - Overview of NVS user accessibility for inputs (via the Editor) and outputs as machine-to-machine or web browser content (NVS Browser).

The NVS contents are held within an Oracle RDBMS. A web based editor allows vocabulary content to be manipulated by authorised users outside BODC.

The NVS contents are then available via web services. Whilst this is ideal for technical users and those fluent in SPARQL, most users require something a little more user-friendly and so the NVS browser2 was developed over the end point to allow more users to query the vocabularies and navigate the NVS ontology.

Exposing the semantic model that has underpinned the construction of PUV terms (green bubbles in Fig. 2) using the Linked Data paradigm gives increased searchability by a method analogous to a fruit machine. The new Vocabulary Builder, released by BODC in summer 2016, allows the end user to build faceted style queries using subcomponent vocabularies through a user friendly web interface. Conceptually each wheel on the fruit machine representing a component of the semantic model enhances the Vocab Builder's search functionality. Further this paradigm allows users to construct and submit new combinations for addition to a vocabulary by using the underlying semantic model that combines terms from other vocabularies.

Fig. 2 - Illustration of the key vocabulary linkage from the PUV, the underlying semantic model components (in green) and the links to external authoritative vocabularies (in red).

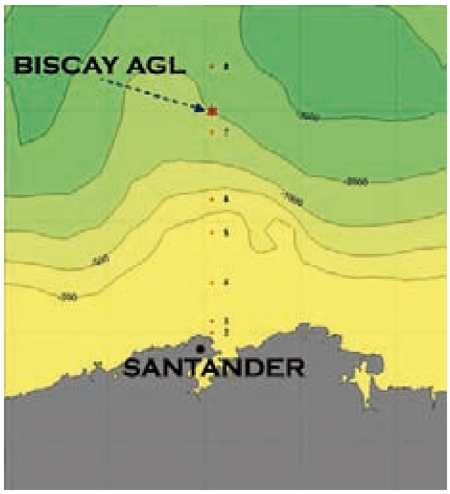
In addition by providing links via SKOS mappings into external ontologies (red bubbles in Fig. 2) the information captured in a 1980s parameter code now conforms to the Linked Data paradigm of the Semantic Web, vastly increasing the descriptive information accessible to a user. For example the parameter code for a contaminant in biota, such as 'cadmium in *Mytilus edulis*', carries RDF triples leading to the entry for *Mytilus edulis* in the WoRMS and for cadmium in the ChEBI ontologies.

This presentation will demonstrate the new user functionality provided by tools created in the last 12 months:

- NVS Editor
- NVS Browser search tool
- External mappings
- NVS Vocabulary Builder.

New developments on Biscay-AGL Observatory. From derived products to sensor networks

Daniel Cano, Instituto Español de Oceanografía Santander (Spain), daniel.cano@st.ieo.es
Elena Tel, Istituto Español de Oceanografia Madrid (Spain), elena.tel@md.ieo.es
César González-Pola, Istituto Español de Oceanografia Gijón (Spain), cesar.pola@gi.ieo.es
Raquel Somavilla, Istituto Español de Oceanografia Gijón (Spain), raquel.somavilla@gi.ieo.es
Amaia Viloria, Istituto Español de Oceanografia Santander (Spain), amaia.viloria@st.ieo.es
Carmen Rodríguez, Istituto Español de Oceanografia Santander (Spain), carmen@st.ieo.es
Manuel Ruiz Villareal, Istituto Español de Oceanografia Coruña (Spain), manuel.ruiz@co.ieo.es
Alicia Lavín, Istituto Español de Oceanografia Santander (Spain), alicia.lavin@st.ieo.es



Biscay-AGL Observatory

Since 1991, shelf and slope waters of the Southern Bay of Biscay are regularly sampled in a monthly hydrographical section north of Santander to a maximum depth of 2500m, as part of the IEO Radiales program. On June 2007, an ocean-meteorological buoy (AGL) was moored at the end of Santander standard section, 22nm north at 2850m depth, to complete the ocean information with the ocean-atmosphere interaction. All of them are part of IEO Observing System (1). The integrated system of AGL and its nearest hydrographic station (2600m depth) is named Biscay-AGL observatory. It is also one station for the EU FixO3 project. Joint resources and systematic analysis lead to a powerful tool, which is much more than the combination of the AGL buoy and the hydrographical samplings.

Data Access

All AGL buoy collected data are added to the local database sited at IEO-Santander in real-time and, after rutinary automatised quality controls, they are immediately available through its dedicated webpage (www.boya_agl.st.ieo.es).

Monthly CTDO2 data from the hydrographic section are lab-calibrated in order to obtain acurated values of salinity, dissolved oxygen and density, and added to the long-term time series.

Biscay-AGL data are quality controlled, flagged and formatted according internationally agreed standards (2, 3) and routinely sent to the IEO Datacenter. This added-value controlled data are incorporated to the IEO permanent archive and made freely available through the SeaDataNet infrastructure for data access and discovery.

Derived products

Data acquired by Biscay-AGL may be displayed as timeseries as usual, but end-users are benefited by derived products which provide direct information. A recently developed software tool produces not only timeseries of several parameters at different time resolutions but also derived products, both real and delayed time. The real time products are generated without human intervention and with every new data arrived at the reception station. Derived products from this buoy include, but not only, annual cycles as well as anomalous values, air-sea heat fluxes, salinity and water temperature anomalies, subinertial currents series, chlorophyll surface series, estimations of the mixed layer depth and wind and currents roses.

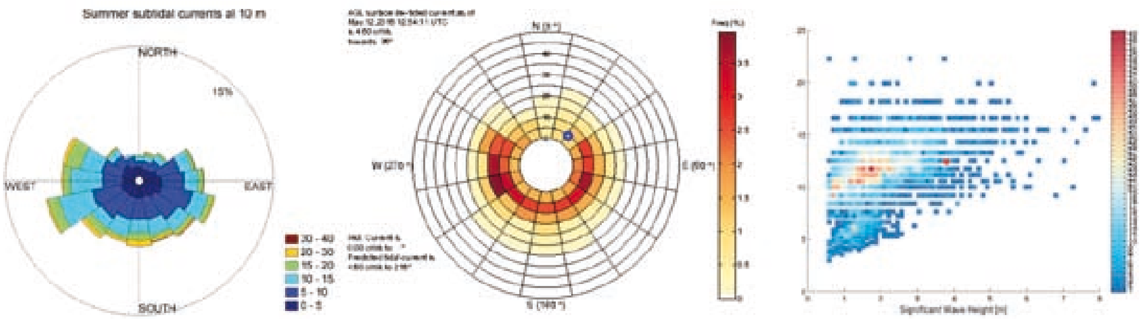


Fig. 1 - Examples of new derived products developed: subtidal currents, instantaneous current and significant wave height.

Sensor Web Enablement (SWE)

Sensor Web infrastructures are setup to access real-time data observed by sensors. This system has been implemented in AGL buoy sensors in order to simplify the retrieved events and alerts triggered through sensors. All those functionalities of the Sensor Web are provided in an interoperable way, following the standards stablished by the Open Geospatial Consortium (OGC). By defining standardized service interfaces, these services hides the heterogeneity of the sensor network, its communication details, enabling the access to archived sensor data that can be discovered and accessed using open and standard protocols and application programming interfaces.

Acknowledgements. The RADIALES program and AGL buoy are mainly supported by the IEO, but also partially by VACLAN (PN I+D+I REM2003-08193-CO3-007MAR), COVACLAN (PN I+D+I CTM2007-646007MAR). FixO3 (FP7-INFRA-2012-1.1.11/312463), SeaDataNetII(FP7/2007-2013/283607).

References

Tel, E., Balbin, R., Cabanas, J.-M., Garcia, M.-J., Garcia-Martinez, M. C., Gonzalez-Pola, C., Lavin, A., Lopez-Jurado, J.-L., Rodriguez, C., Ruiz-Villarreal, M., Sánchez-Leal, R. F., Vargas-Yáñez, M., and Vélez-Belchí, P.: *IEOOS: the Spanish Institute of Oceanography Observing System*, Ocean Sci., 12, 345-353, doi:10.5194/os-12-345-2016, 2016.

SeaDataNet Quality Control Procedures v2.0 http://www.seadatanet.org/content/download/18414/119624/file/SeaDataNet_QC_procedures_V2_%28May_2010%29.pdf

SeaDataNet data file formats: ODV, MEDATLAS, NetCDF (deliverable D8.5) http://www.seadatanet.org/content/download/16251/106283/file/SDN2_D85_WP8_Datafile_formats.pdf

The Global Sea Level Observing System (GLOSS) Data Rescue plan

Elizabeth Bradshaw, British Oceanographic Data Centre (United Kingdom), elizb@bodc.ac.uk

Lesley Rickards, British Oceanographic Data Centre (United Kingdom), ljr@bodc.ac.uk

Andy Matthews, Permanent Service for Mean Sea Level (United Kingdom), psmsl@noc.ac.uk

Svetlana Jevrejeva, Permanent Service for Mean Sea Level (United Kingdom), psmsl@noc.ac.uk

The Global Sea Level Observing System (GLOSS) is an international programme conducted under the auspices of the WMO-IOC Joint Technical Commission for Oceanography and Marine Meteorology. It was set up in 1985 to collect long-term tide gauge observations and to develop systems and standards "for ocean monitoring and flood warning purposes".

Historical sea level data are rare and unrepeatable measurements with a number of applications in climate studies (sea level rise), oceanography (ocean currents, tides, surges), geodesy (national datum), geophysics and geology (coastal land movements) and other disciplines. However, long-term time series are concentrated in the northern hemisphere and there are no records at the Permanent Service for Mean Sea Level (PSMSL) global data bank longer than 100 years in the Arctic, Africa, South America or Antarctica.

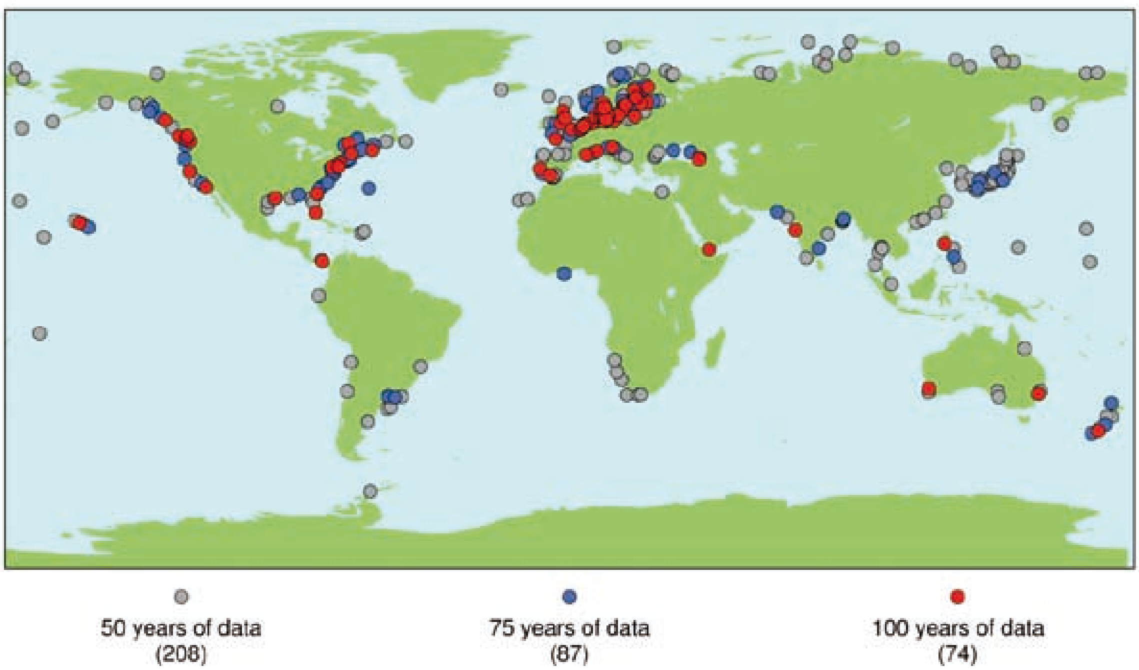


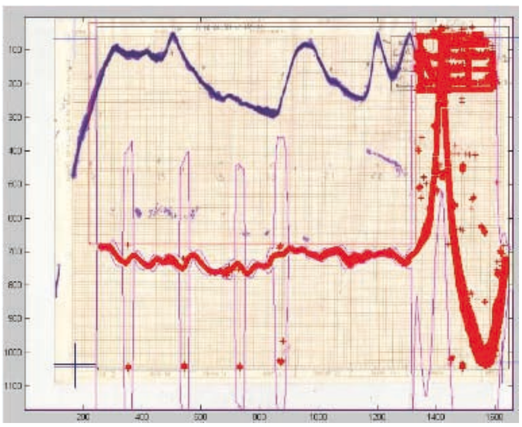
Fig. 1 - Longest records in the PSMSL.

At the GLOSS Group of Experts (GLOSS-GE) XIV Meeting in 2015, a number of data rescue action items were agreed upon, to be developed in the next two years.

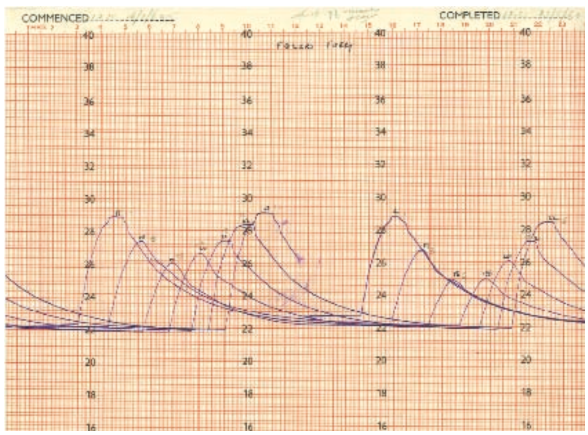
These included:

- Exploring mareogram digitisation applications, including NUNIEAU (more information available at: http://www.mediterranee.cerema.fr/logiciel-de-numerisation-des-enregistrements-r57.html) and other recent developments in scanning/digitisation software, such as IEDRO's Weather Wizards program, to see if they could be used via a browser.
  and
- Proposing "Guidelines" for rescuing sea level data.

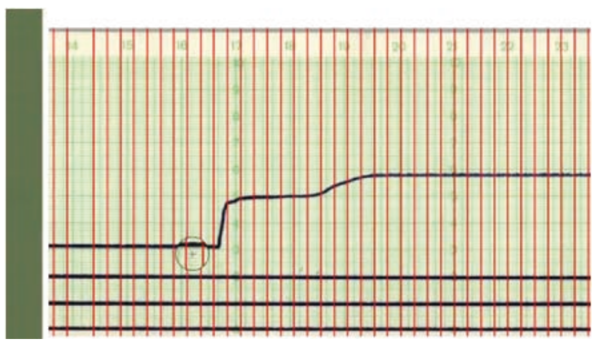
NUNIEAU, CETE Méditerranée



Analogue chart



Weather Wizards, IEDRO



Several GLOSS members have already developed software to automatically digitise charts and there are other freely available software packages that convert images to numerical values. One activity GLOSS could coordinate is a comparison study of the various different digitising software programs by sending the same charts to participating organisations and asking everyone to digitise them using their own procedures. GLOSS could then provide recommendations to the community, create a software repository and even make software available that functions within a web browser.

The GLOSS-GE will work with other international data rescue groups such as the Research Data Alliance Data Rescue Interest Group to produce guidelines for rescuing historic data.

These action items will aid the discovery, scanning, digitising and quality control of analogue tide gauge charts and sea level ledgers and improve the quality, quantity and availability of long-term sea level data series. Data archaeology activities will help fill in the gaps in the global dataset and improve global sea level reconstruction.

Subsurface Water Sampling Network of the Instituto Español de Oceanografía: from data acquisition to final users

Gonzalo González-Nuevo, Instituto Español de Oceanografía (Spain), gonzalo.gonzalez@vi.ieo.es
Elena Tel, Instituto Español de Oceanografía (Spain), elena.tel@md.ieo.es
Agueda Cabrero, Instituto Español de Oceanografía (Spain), agueda.cabrero@vi.ieo.es
Pablo Otero, Instituto Español de Oceanografía (Spain), pablo.otero@md.ieo.es
Manuel Ruiz-Villareal, Instituto Español de Oceanografía (Spain), manuel.ruiz@co.ieo.es
Jose Manuel Cabanas, Instituto Español de Oceanografía (Spain), jm.cabanas@vi.ieo.es

The Instituto Español de Oceanografía (IEO) has implemented a subsurface water sampling network on its fleet like part of its observing system: IEOOS (1). This ship-based network was formed by two local vessels (B/O Navaz, B/O Lura), three regional ones (B/O Ramon Margalef, B/O Angeles Alvariño, B/O Francisco de Paula Navarro) and an oceanic one (B/O Miguel Oliver). Each of them was instrumented with a SeaBird 21 thermosalinograph (TSG) and a Turner 10 Fluorometer. The vessels work all around the Spanish seawaters: the Iberian Peninsula, Balearic and Canary Islands. The TSG data are daily sent to a processing centre. An automatic data processing system was developed to manage all the information generated in quasi-real time by this subsurface sampling network.

The daily quality control is performed in order to detect geopositioning errors, spikes, etc. This quality control includes the assignment of quality flags based on international criteria established in the frame of SeaDataNet European projects (2) and adds information about the data reliability.

Daily controlled data are stored in a Thematic Realtime Environmental Distributed Data Services server (THREDDS, http://centolo.co.ieo.es:8080/thredds/) for operational oceanography purposes. This infrastructure facilitates the data access by scientific community and its visualization by means of Open Geospatial Consortium (OGC) standard services. Nowadays an automatic data storage system based on Postgres/PostGIS database is being developed in order to make easier the implementation of a user-friendly web service to visualize them. The metadata generation is carried out following INSPIRE (2007/2/EC) directive, allowing the interoperability of the database and making easier the development of end-user services based on it.

Delayed mode quality control and permanent data archive at IEO data centre, includes a monthly validation and their incorporation to the SeaDataNet infrastructure, from where they are also accessible under the agreed data policy.

This effort in TSG data gathering and their efficient distribution is helping to use these TSG data for the evaluation of the ocean models that routinely run in the Atlantic area. As examples of the scientific interest of these routinely acquired data, the system is already giving information on the exchange of water between the Galician Rias Baixas (seawater inlets on the NW Spain) and the shelf, the variability in the position of river plume fronts or the spatial variability of chlorophyll concentration.

Acknowledgement. The thermosalinometers data program of IEOOS is supported by the IEO. The IEO data centre has been partially funded by SeaDataNet II (FP7/2007-2013/283607).

References

Tel, E., Balbin, R., Cabanas, J.-M., Garcia, M.-J., Garcia-Martinez, M. C., Gonzalez-Pola, C., Lavin, A., Lopez-Jurado, J.-L., Rodriguez, C., Ruiz-Villarreal, M., Sánchez-Leal, R. F., Vargas-Yáñez, M., and Vélez-Belchí, P.: *IEOOS: the Spanish Institute of Oceanography Observing System*, Ocean Sci., 12, 345-353, doi:10.5194/os-12-345-2016, 2016.

Seadatanet Quality Control Procedures v2.0 http://www.seadatanet.org/content/download /18414/119624/file/SeaDataNet_QC_procedures_V2_%28May_2010%29.pdf

A user-centric approach to validate the EVER-EST Virtual Research Environment infrastructure for the Earth Science

Federica Foglini, ISMAR Consiglio Nazionale delle Ricerche (Italy), federica.foglini@bo.ismar.cnr.it
Francesco De Leo, ISMAR Consiglio Nazionale delle Ricerche (Italy), francescodeleo64@yahoo.it
Jose Manuel Gómez, Expert System (Spain), jmgomez@expertsystem.com
Valentina Grande, ISMAR Consiglio Nazionale delle Ricerche (Italy), valentina.grande@bo.ismar.cnr.it
Rosemarie Leone, European Space Agency ESOC (Germany), rosemarie.Leone@esa.int
Fulvio Marelli, Terradue (Italy), fulviomarelli@me.com

The EVER-EST Virtual Research Environment (VRE)

The EVER-EST project aims to develop a generic Virtual Research Environment (VRE) tailored to the needs and validated by the Earth Science domain. To achieve this the EVER-EST VRE provides earth scientists with the means to seamlessly manage both the data involved in their computationally intensive disciplines and the scientific methods applied in their observations and modellings, which lead to the specific results that need to be attributable, validated and shared within the community e.g. in the form of scholarly communications.

Central to this approach is the concept of Research Objects (ROs) as semantically rich aggregations of resources that bring together data, methods and people in scientific investigations. ROs enable the creation of digital artifacts that can encapsulate scientific knowledge and provide a mechanism for sharing and discovering assets of reusable research and scientific assets as first-class citizens. Though several e-laboratories are incorporating the research object concept in their infrastructure, the EVER-EST VRE is the first RO-centric native infrastructure leveraging the notion of ROs and their application in observational rather than experimental disciplines and particularly in Earth Science (Fig.1).

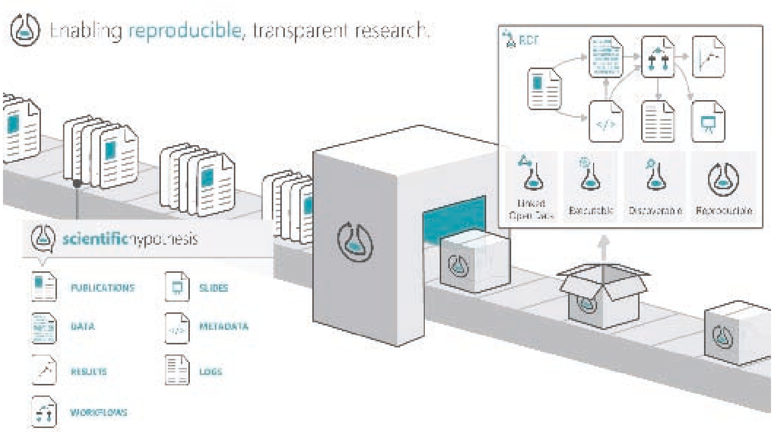


Fig. 1 - Example of RO schema.

By deploying research object technologies in Earth Science, we aim to enhance the portability of scientific materials and methods, the collaboration capabilities within and betweenteams of earth scientists, the quality and long-term preservation capabilities of the methods, and the means to com-municate, validate and measure the impact of scientific outcomes.

The EVER-EST VRE offers a framework based on advanced services which are delivered both at the infrastructure and domain-specific level, with the objective of supporting each described phase of the Earth Science Research and Information Lifecycle. It provides innovative e-research services to Earth Science user communities for new forms of scholarly communication, cross-validation and the sharing of knowledge and science outputs.

The EVER-EST Virtual Research Communities (VRCs)

The project follows a user-centric approach. Real use cases, taken from pre-selected communities covering different Earth Science research scenarios drive the implementation of the proposed system's services and capabilities. The EVER-EST VRE user community ranges from marine researchers (Sea Monitoring use case), to Common Foreign and Security Policy institutions and stakeholders (Land Monitoring use case), natural hazards forecasting systems (Natural Hazards use case), and disaster and risk management teams (Supersites use case). The resulting requirements for data management (e.g. accessibility of heterogeneous data sources), preservation of data and data-intensive methods and workflows, user experiences, data exploitation and other important e-research services for communication, cross-validation, and sharing of science products, will be addressed through the proposed solution.

Each identified Virtual Research Community (VRC), represents a different research area and is constantly involved in the project for the purposes of validating the EVER-EST infrastructure. This user-centric approach allows an assessment to be made of the capability for the proposed solution to satisfy the heterogeneous needs of a variety of Earth Science communities for more effective collaboration, and higher efficiency and creativity in research. Each of the identified VRCs involves different categories of users both within and outside the Earth Science domain (e.g. general public, policy and decision makers). This allows the piloting and testing of the infrastructure in real-life scenarios which may significantly differ from each other. Data providers and value adders are constantly involved in this process during the project to maximize its scientific and societal return.

The Sea Monitoring VRC

The Institute of MARine Science (ISMAR-CNR) is a scientific partner providing the EVER-EST project with a direct link to the targeted users with the definition of real use cases to be carried out through the VRE solution. In particular, ISMAR-CNR provides useful and applicable contributions to the identification and definition of variables indicated by the European Commission in the Marine Strategy Framework Directive (MSFD) to achieve the Good Environment Status (GES). ISMAR-CNR is willing to deliver practical methods, procedures and protocols to support coherent and widely accepted interpretation of the Descriptors 1, 2 and 6. The three use cases deal with 1. the Posidonia meadows along the Apulian coast, 2. the deep-sea corals along the Apulian continenatal slope and 3. the jellyfish abundance in the Italian water.

Research objects play a significant role in key aspects related to data-intensive science like long-term data preservation. The SeaMonitoring VRC created specific RO for bathymetric data implementing a data preservation plan and a specific vocabulary for metadata. The metadata elements are defined taking into account several international metadata standards such as the PREMIS, Dublin Core standards and vocabularies and Long Term Preservation of Earth Observation Space Data - Definitions of Acronyms and Terms.

Marine Science Videos

Catherine Borremans, Institut français de recherche pour l'exploitation de la mer (France), catherine.borremans@ifremer.fr
Frédéric Merceur, Institut français de recherche pour l'exploitation de la mer (France), frederic.merceur@ifremer.fr

More and more research work on marine environment is based on the analysis of imagery. Thanks to technological development, we can now monitor ocean ecosystems using video cameras deployed on different platforms. Video acquisition provides information on organisms' life and on environmental conditions. The sources of video data and the subjects of interest are many and varied: films from submersible dives in deep-sea environment, programmed video sequences of undersea rotary video stations in lagoons, high-frequency video data on deep-sea observatories, video observations recorded by fixed and autonomous camera devices on fishing gears, and others. The information contained in those images is essential to understand species biology and the functioning of these ecosystems, and ultimately inform environmental management policies. Because of the video data specificity (volumes, formats, diffusion…) a dedicated system has to be proposed for displaying and making accessible this huge archive of imagery, still growing every day.

In spring 2015 IFREMER opened a new web portal called "Marine Science Videos" (http://video.ifremer.fr/). The first objective of this video library was to give access to recent and historical images produced during the IFREMER submersibles dives (ROV Victor 6000 and Nautile). Users can view and search videos and photos by keywords or dive information (cruise, submersible, dive, year, zone, camera...).

Indeed, one or more submersibles can be deployed during oceanographic cruises in order to perform several dives lasting from some hours to some days. With their cameras catching videos throughout the dive they produce Tera Bytes of footages that are worth to be proposed to all IFREMER'S partners and to the public. The videos are available within three definitions (Low, Medium and High) in the MP4 H264 proxy format. Metadata relative to submersible dives and acquisition conditions are proposed together with this data: submersible position and attitude, temperature and salinity conditions, camera position and configuration. Comments recorded simultaneously with the videos are displayed if existing.

Furthermore, images acquired during corresponding operations are accessible in addition to the videos:

- Thumbnails: also called «mini-films», these images result from an automatic video sampling at a given frequency.
- Images: these are photos from the submersible digital camera and video snapshots.

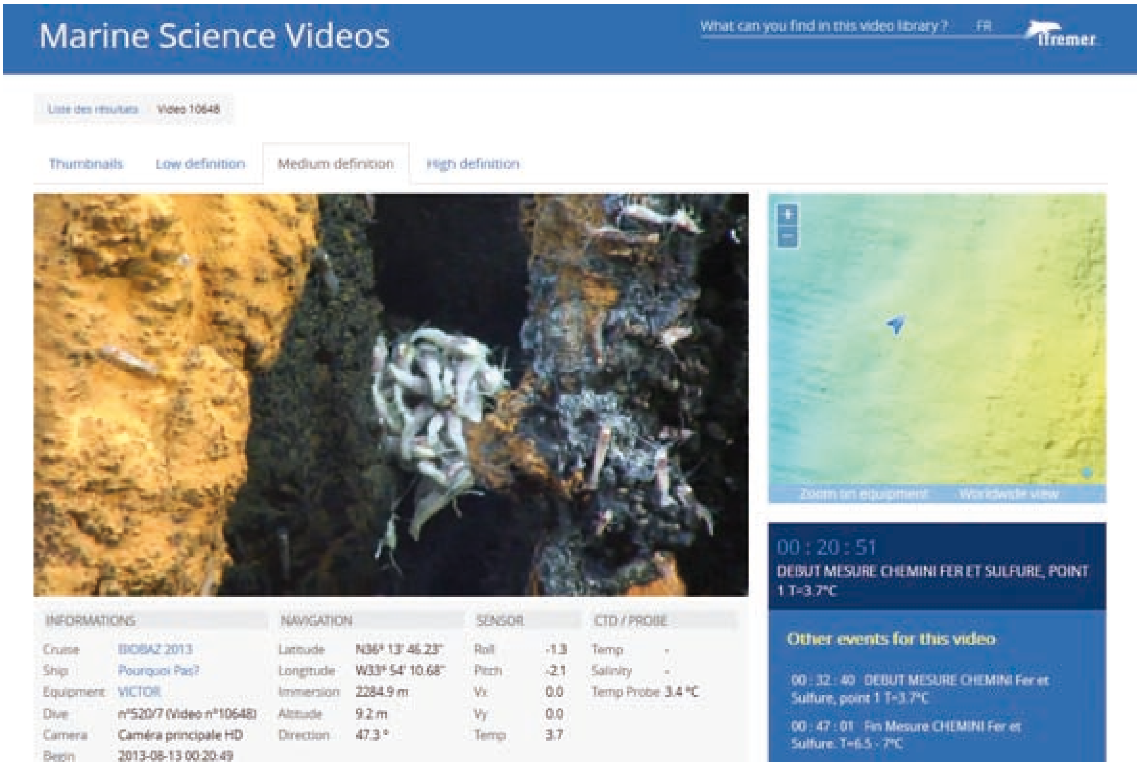


Fig. 1 - Video and related metadata visualisation interface.

Thumbnails, Low Definition images and videos are downloadable without restriction. Medium and High Definition versions can be used by authorised persons or on demand.

At a later stage this portal will be adapted in order to present other video types:

- videos from other IFREMER vehicles or devices (SCAMPI, Pagure, mini ROV, HROV, towed cameras, EROC-VECOC) or from divers,
- videos from IFREMER observatories or fixed points (EMSO, STAVIRO-MICADO),
- similar videos acquired by other institutes using their own equipment.

The integration of scientific-orientated indexing of images (observed taxa, geological features...) and related products and tools (mosaic, picture enhancement...) will be important perspectives for the system as well.

SEXTANT, a marine spatial data infrastructure: implementation of OGC protocols for the dissemination of marine data

Mickael Treguer, Institut français de recherche pour l'exploitation de la mer (France), Mickael.Treguer@ifremer.fr

Catherine Satra Le Bris, Institut français de recherche pour l'exploitation de la mer (France), Catherine.Satra@ifremer.fr

Erwann Quimbert, Institut français de recherche pour l'exploitation de la mer (France), Erwann.Quimbert@ifremer.fr

Julien Meillon, Institut français de recherche pour l'exploitation de la mer (France), Julien.Meillon@ifremer.fr

At national and European levels, in various projects, data products are developed to provide end-users and stakeholders with homogeneously qualified observation compilation or analysis.

Ifremer has developed a spatial data infrastructure for marine environment, called Sextant, in order to manage, share and retrieve these products for its partners and the general public. Thanks to the OGC and ISO standard and INSPIRE compliance, the infrastructure provides a unique framework to federate homogeneous descriptions and access to marine data products processed in various contexts, at national level or European level for DG research (SeaDataNet), DG Mare (EMODNET) and DG Growth (Copernicus MEMS).

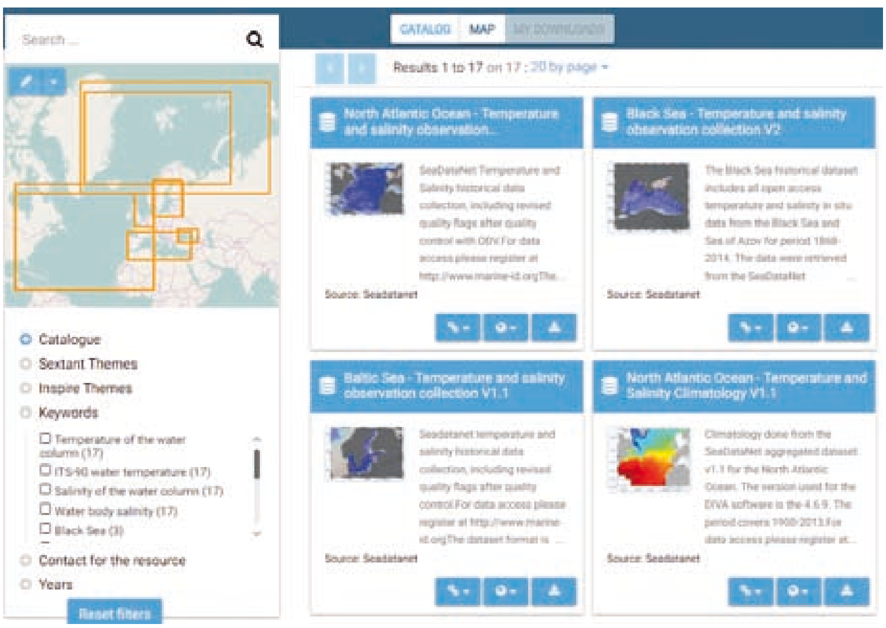


Fig. 1 - SeaDataNet content catalogue.

The discovery service of Sextant is based on the metadata catalogue. The data description is normalized according to ISO 191XX series standards and Inspire recommendations. Access to the catalogue is provided by the standard OGC service, Catalogue Service for the Web (CSW 2.0.2).

Data visualization and data downloading are available through standard OGC services, Web Map Services (WMS) and Web Feature Services (WFS). Several OGC services are provided within Sextant, according to marine themes, regions and projects. Depending on the file format, WMTS services are used for large images, such as hyperspectral images, or NcWMS services for gridded data, such as climatology models.

New functions are developped to improve the visualization, analyse and access to data, eg: data filtering, online spatial processing with WPS services and acces to sensor data with SOS services.

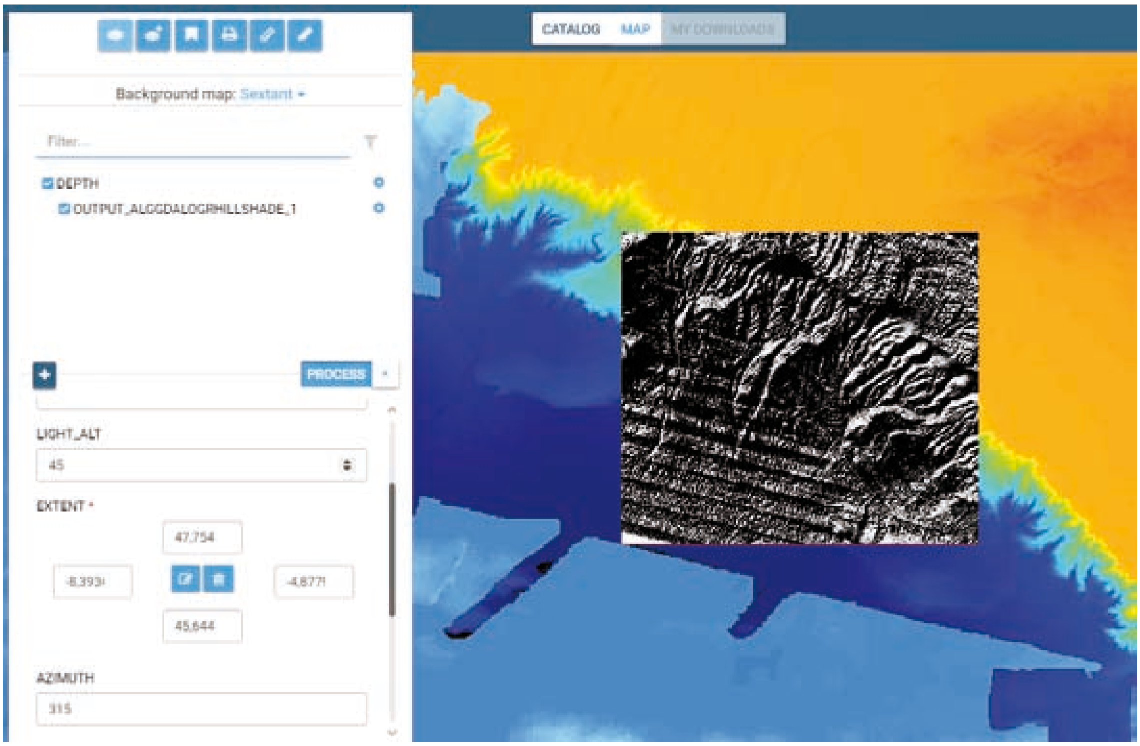


Fig. 2 - Online Process output for hillshade on DEM.

Unlocking the potential for coastal innovation growth using Earth Observation data and cloud infrastructure

Eirini Politi, University College Cork (Ireland), eirini.politi@ucc.ie
Rory Scarrott, University College Cork (Ireland), r.scarrott@ucc.ie
Miguel Terra Homem, DEIMOS (Portugal), miguel.terra@deimos.com.pt
Hervé Caumont, Terradue (France), herve.caumont@terradue.com
Nuno Grosso, DEIMOS (Portugal), nuno.grosso@deimos.com.pt
Antoine Mangin, ACRI (France), antoine.mangin@acri-st.fr
Nuno Catarino, DEIMOS (Portugal), nuno.catarino@deimos.com.pt

Abstract

In the era of Big Data, satellite remote sensing plays a major role in marine and coastal research, by contributing large and long archives of datasets to the scientific community for a wide range of marine and coastal applications. Dating from the 1970s, the generation of polar-orbiting satellite data that cover the entire globe can be as frequent as every 12 hours. The availability of such satellite datasets, some of which now hold an open access status, has made feasible the development of large-scale marine projects, with examples including the European Space Agency (ESA) Climate Change Initiative (CCI) Ocean Colour, Sea Surface Temperature, Sea Ice and Sea Level projects. However, until recently, scientists had to deal with the daunting task of mining large datasets for suitable data, and often downloading EO information from various different sources. In addition, as the datasets increased in volume, the processing has become slower and demanding of better computing facilities. The European Commission (EC) H2020 Co-ReSyF (Coastal Waters Research Synergy Framework) project aims to tackle these issues, by developing a platform for combined data access, processing, visualisation and output in one place. The platform is based on cloud computing to maximise processing effort and task orchestration. Co-ReSyF will address issues faced by inexperienced and new EO researchers, and also target EO experts and downstream users, with main focus on enabling EO data access and processing for coastal and marine applications. Such a platform will revolutionise accessibility to Big EO Data, and help create a new era of EO data processing and exploitation in the coastal and marine environment.

Overview of the Co-ReSyF platform

The Co-ReSyF project kick-off in February 2016 and will implement a dedicated data access and processing infrastructure, with automated tools, methods and standards to support coastal water research applications that use Earth Observation (EO) data. The user-friendly interface will be accessible to inexperienced researchers in EO and coastal research, but also EO experts and algorithm developers.

Co-ReSyF platform facilities and tools, optimised for distributed processing, include EO data access catalogues, discovery and retrieval tools, as well as a number of (pre-)processing toolboxes for manipulating EO data (Fig. 1). Through a collaborative front-end, the Co-ReSyF platform users will be able to upload new algorithms, or use existing ones, compose and configure processing chains for easy deployment on the cloud infrastructure. Users will be able to accelerate the development of high-performing applications taking full advantage of the scalability of resources available in the cloud framework. Advanced users will be able to take full control of the processing chains and algorithms by having access to the cloud back-end, and to further optimize their applications for fast deployment for Big Data access and processing.

Development phases

The goal of the Co-ReSyF project is to develop and provide a service to the coastal and marine community that incorporates EO data and information into multi-disciplinary coastal research applications. This service, in the form of a cloud platform hosted in a distributed environment, will be developed in two major phases. The initial phase (2016-2018) is dedicated to platform evolution based on user feedback and consultations that will guide the definition of requirements. A series of early adopters will support and demonstrate the Co-ReSyF capabilities, and access the platform to develop new research applications in the coastal domain, and hence serve as beta-testers of the system. Towards the end of the project, the Co-ReSyF platform will be publicly released to the research community (second phase).

Research Applications

Co-ReSyF aims to focus on the needs of the coastal waters research community and as a result, six main coastal research applications have been thus far identified, and are currently being

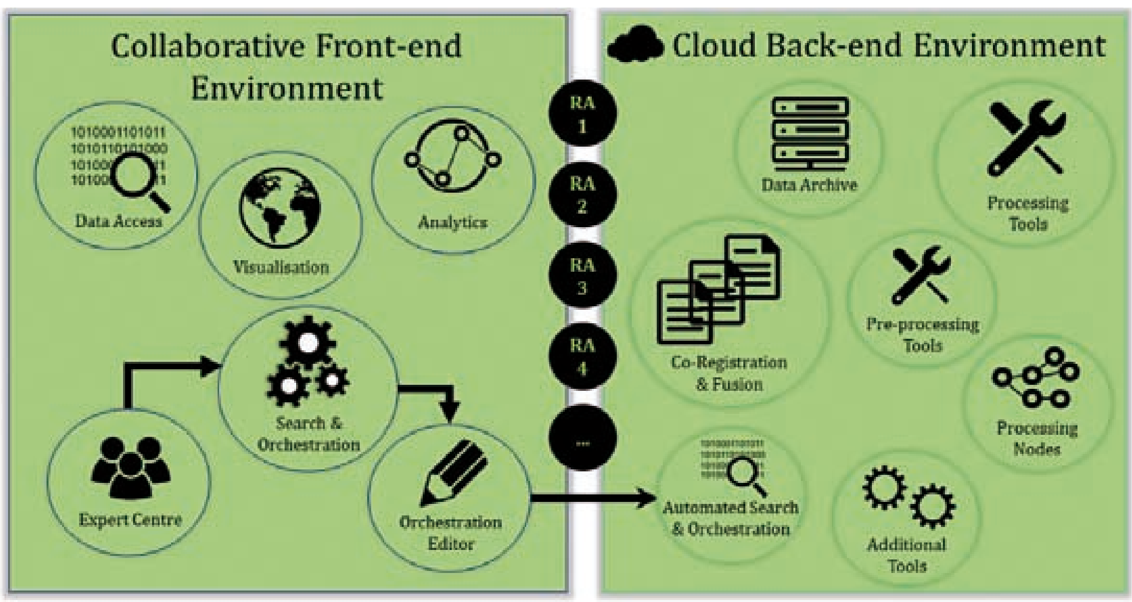


Fig. 1 - Co-ReSyF development and testing environment, including interdependence of platform components (RA denotes Research Applications).

developed by the project partners for integration in the system. These are:

- Coastal water bathymetry mapping
- Coastal water benthic classification
- Coastal water quality
- Marine vessel detection and oil spill detection
- Open water and coastal water altimetry
- Time series processing for hyper-temporal optical data analysis

These applications use a range of optical, thermal and Synthetic Aperture Radar (SAR) satellite data, and will be complemented by new applications developed later on by the platform early adopters and various users.

Conclusions

Co-ReSyF intends to eliminate many of the barriers related to the use of EO data, usually identified by inexperienced users. This will lead to a wider integration of EO datasets in a research context. Such an expansion will raise awareness to the potential of EO data, and encourage innovative thinking and development of new algorithms, EO products and services. Ultimately, the platform, when operational, aims to act as a facilitator for scientific knowledge generation and innovation growth, supporting the advancement of science and allowing ideas to be tested and explored at a scale not previously accessible to all researchers.

Acknowledgement. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687289.

The ICSU World Data System (WDS): An update on progress and current developments

Mustapha Mokrane, WDS International Programme Office (Japan), Mustapha.mokrane@icsu-wds.org
Sandra Harrison, Macquarie University (Australia), sandy.harrison@mq.edu.au
Lesley Rickards, Permanent Service for Mean Sea Level (United Kingdom), ljr@bodc.ac.uk

The World Data System (WDS) is an Interdisciplinary Body of the International Council for Science (ICSU) created by its 29th General Assembly in Maputo, Mozambique in 2008. ICSU-WDS builds on the 50+ year legacy of the World Data Centres and Federation of Astronomical and Geophysical data analysis Services established by ICSU to manage data generated by the International Geophysical Year (1957-1958). It became clear after the International Polar Year (2007-2008) that these bodies were not able to respond fully to modern data needs, and they were thus replaced by the ICSU World Data System in 2009.

This poster describes the development of the WDS since its inception, considering its mission, strategy, membership, working groups and developing partnerships - highlighting successes and considering future activities.

The mission of the World Data System is to support ICSU's vision by promoting long-term stewardship of, and universal and equitable access to, quality-assured scientific data and data services, products, and information across a range of disciplines in the natural and social sciences, and the humanities. ICSU-WDS aims to facilitate scientific research under the ICSU umbrella by coordinating and supporting trusted scientific data services for the provision, use, and preservation of relevant datasets, while strengthening their links with the research community.

The strategy for achieving these objectives is outlined in the current five-year Strategic Plan 2014-2018, structured round five major targets:

1. Make trusted data services an integral part of international collaborative scientific research
2. Nurture active disciplinary and multidisciplinary scientific data services communities
3. Improve the funding environment for trusted data services
4. Improve the trust in and quality of open Scientific Data Services
5. Position WDS as the premium global multidisciplinary network for quality-assessed scientific research data

WDS Members

Member Organizations of ICSU-WDS from wide-ranging fields are the building blocks of a worldwide 'community of excellence' for scientific data. Not only do these Members participate towards advancing WDS goals; their data holdings, services, and products are the cornerstone of the federated data system. As of February 2016, ICSU-WDS had 95 Member Organizations in four different categories. These are:

Type	No.	Definition
Regular	61	Organizations that are data stewards and/or data analysis services
Network	10	Umbrella bodies representing groups of data stewardship organizations and/or data analysis services, some of which may or may not be WDS Regular Members. Usually serve as coordinating agents for nodes that have common characteristics and mostly common disciplines.
Partner	6	Organizations that are not data stewards or data analysis services, but that contribute support or funding to ICSU-WDS and/or WDS Members
Associate	18	Organizations that are interested in the WDS endeavour and participating in our discussions, but that do not contribute direct funding or other material support

Member Organizations are formally accepted into ICSU-WDS by the WDS Scientific Committee, who safeguard the trustworthiness of ICSU-WDS by certifying Regular and Network Members according to internationally recognized standards. Partners and Associates are co-opted.

ICSU-WDS brings its Member Organizations together to coordinate their activities and through that process, to achieve an overall capability that transcends individual ones. Membership also increases exposure to potential users and collaborators internationally. It demonstrates that the Member Organizations have a strong and tangible commitment to open data sharing, data and service quality, and data preservation - all of which are increasingly considered prime requirements by science funders and are high on policymakers' agendas since they benefit the scientific community, economy, and society in general.

WDS Working Groups

Data Publishing

The WDS Working Group on Publishing Data was started in late 2012 with the aims of identifying and defining best practices for data publication, as well as testing its implementation by involving core stakeholders. The WDS concept addresses essential problems and practical issues such that publication of research data can become part of the scholarly record. Because the involved issues are interlinked, the following four WGs were established: Bibliometrics, Cost Recovery, Services and Workflows.

Certification

Data sharing enables reuse of data and ultimately makes science more transparent. To ensure the quality and usability of shared data and to provide long-term preservation and access, sustainable data services are key components of scientific infrastructure. However, it is vital to guarantee the trustworthiness of a service to perform these functions, and certification by an appropriate authority is fundamental in promoting confidence in shared data resources.

Knowledge Network

The Knowledge Network is a Web-based, interlinked repository of relationships between the actors and entities that make up our research landscape: people, institutions, projects, research disciplines and topics, funding sources, and the like. ICSU-WDS has a stated objective to establish an aggregation of WDS Members' metadata holdings so that the extent of data and services offered are discoverable through a single interface.

New Partnerships Established

The Research Data Alliance, ICSU's Committee on Data for Science and Technology (CODATA) and the World Data System have signed Memoranda of Understanding outlining their collaboration. The three organizations have different but complementary foci and strengths, and the MoUs clarify their respective strengths and enable the new Partnership to accomplish more; to advance data practices and science faster and further, and thus increase the benefit of research data for society.

Creating a Common Operating Picture and Improving Efficiency using Open Standards - Practical Examples from UK Ports

Mike Osborne, OceanWise (United Kingdom), mike.osborne@oceanwise.eu

A port is a multifaceted business requiring access to all types of data and information. Traditionally, this data is managed in non-connected applications and proprietary formats making the sharing and exchange of data difficult and thereby increasing management overheads. The use of open standards and systems can help to solve this problem but how these are integrated into existing business and IT environments and applied to pre-existing workflows remains a major challenge.

Over the last 5 years, OceanWise has been working closely with numerous UK ports (e.g. London, Southampton and Liverpool) to provide practical solutions to improve how these ports manage and disseminate their data. These solutions are making essential data more readily accessible to users inside the port and improving the exchange of data with external stakeholders e.g. dredging contractors and licensing authorities. In doing this, ports are now significantly increasing operational efficiency and reducing risk.

This paper uses examples of environmental monitoring and geographic information systems (GIS) maps, combined with business and operational data, to demonstrate the challenges ports face and how these are being overcome. The paper also shows how port personnel are adopting the latest international standards from the OGC, for example, and applying these practically including to maritime specific situations by, for example, port pilots, hydrographers and VTS (Fig. 1) to further increase the safety of navigation.

Fig. 1 - Port of London VTS is increasingly relying on Open Standards to Integrate Datasets from a Range of Disparate Sources.

A System Console allows different web pages to be configured according to users' differing requirements and allows users to easily switch between real time data from multiple sensors to more detailed and historical data from a single sensor depending on the task at hand. Predictions of tide height are combined with observed water level data from tide gauges at multiple locations in and around the port and port approaches. Surge forecast data from the UK Met Office, for example, is added to allow crucial decisions on vessel movement and flood protection measures to be made.

A map interface (Fig. 2) built using OpenLayers provides context and situation awareness. Different types of data are displayed depending on user preference. Overlays provided from source as OGC compliant Web Map or Web Feature Services include a WMS of ENCs (Electronic Navigational Charts), real time and historical vessel data from AIS (Automatic Identification System) and predicted tidal current vectors, and the location and status of environmental monitoring stations and telemetry systems.

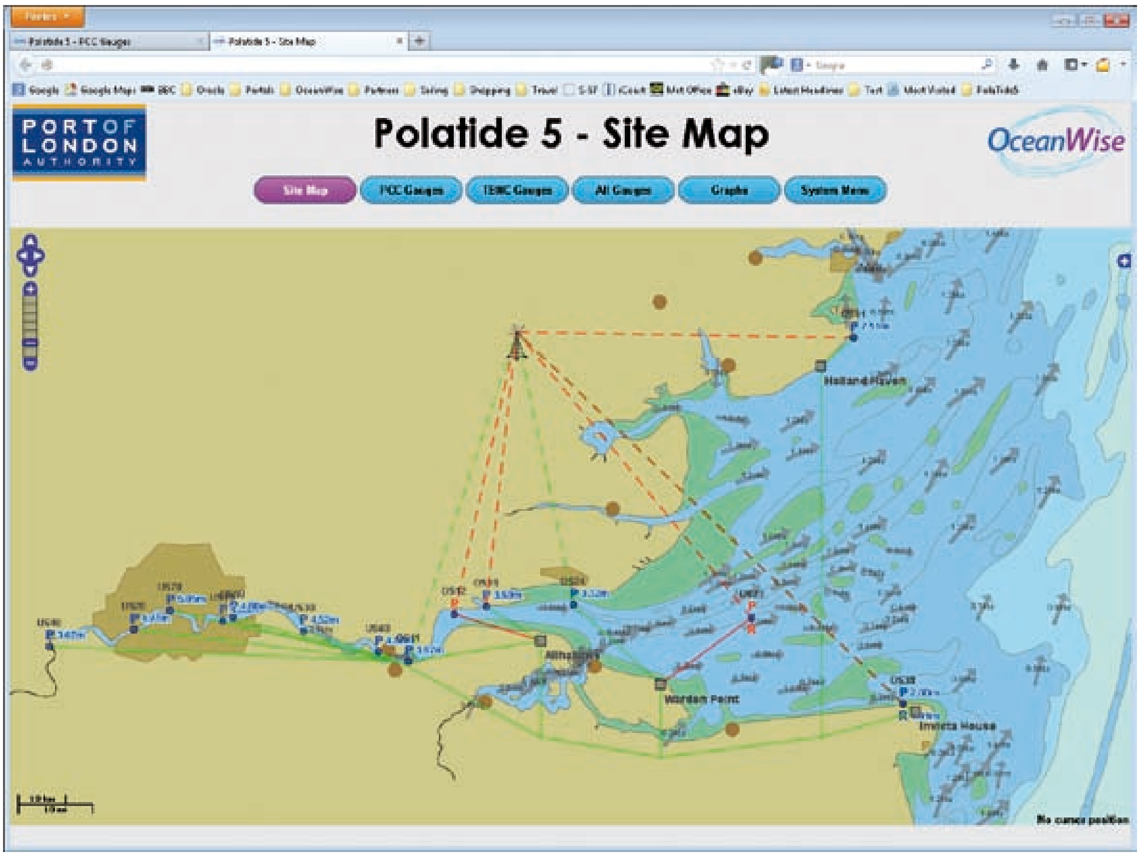


Fig. 2 - Port of London VTS Displays includes Situation Awareness Map Overlays delivered via OGC Web Map Services.

Decision support system (DSS) for disasters based on integrated database – output of possible impacts and recommendations for decision makers

Evgenii Viazilov, Russian Research Institute of Hydrometeorological Informatio - WDC (Russia), vjaz@meteo.ru

In spite of the use of the high-automated systems of observation, collection, storage, processing, prognosis and delivery of information of environment, the size of damage increases from the natural disasters. It is related to that information about environment, delivering on industrial enterprises, is effectively not enough used. For the increase of this efficiency, it is possible:

- to organize automatic notification of persons and a decision-making (DM) about arising up natural disasters through software–agent, working on a mobile internet-device;
- to represent information in a compact kind on the screen of computer;
- to estimate a possible damage from natural disasters and expense on preventive actions;
- to give out information about affecting of environment industrial enterprises, population and to recommendation for DM; to present information as interactive maps, graphics, tables.

Tools of the notification of the DM about the arisen natural disasters through the software-agent are working with use of critical values of indicators of natural disasters for each type of enterprise and a technological process. The software-agent is adjusting by the user under concrete object, region, influencing natural disasters.

Information by means of the compact scheme reflecting the environment parameters influencing industrial enterprises in the form of devices (the thermometer, the aneroid, a speedometer) are visualizing with values of the indicators which exceeded critical values.

DM must understand possible impacts from natural disasters and lean on recommendations, given out by the system for decisions support. Delivery of information about possible impacts of natural disasters and recommendations leans on the knowledge, formed on the base of the accumulated experience. The type of information (observing, forecasting, climatic, after the phenomenon), season of year, climatic zone, level of making decision is thus taken into account.

The testing base of critical values are created, the formalized information on impacts and recommendations for natural disasters are collected.

The principles of information dissemination about natural disasters

- **Personalization of information** - only what need at a moment of natural disaster, for a certain enterprise, technological process, management level, region;
- **Mobility** - delivery of information to any mobile internet-device;
- **Automation of delivering of information** about disasters for decision-makers;
- **Automatic using** of information for each object;
- **The formalization** of business processes which define the organization of works to increase the safety of the population and industrial enterprises;
- **Classification** of impacts and recommendations (before, during, after disasters).

The main idea is next. Knowing conditions of the environment is possible define the list of possible impacts on object in advance, knowing impacts, it is possible to define recommendations for acceptance of preventive actions for various levels of objects management in advance.

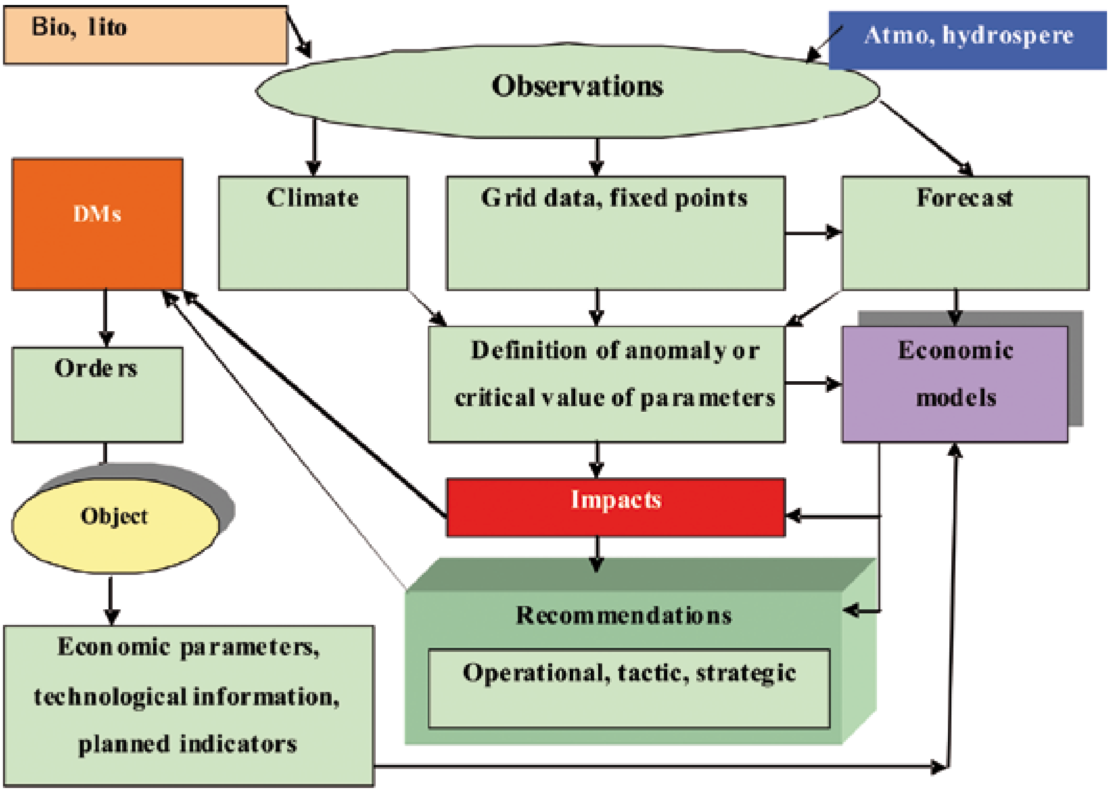


Fig. 1 - Scheme of supporting of decision maker.

For the creation of DSS, the following steps are needed:

- to collect and formalization information on possible impacts and recommendations;
- to develop a database of critical values of environmental parameters for separated objects and technological processes;
- to create of economic models for an assessment of impacts and possible damage, specification of costs on preventive actions.

The knowledge base for decision making is created as tables of impacts and recommendations for natural disasters for different objects, disaster level (disaster level (yellow, orange, red), technological processes, seasons, climate zones. The demonstrational variant of DSS is created for mobile internet device. Automatic detection of natural disasters on the basis of observation and forecast is organized.

SESSION 2

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

Developing a common global framework for marine data management

Helen Glaves, British Geological Survey (United Kingdom) hmg@bgs.ac.uk
Dick Schaap, Marine Information Service (The Netherlands), dick@maris.nl

In recent years there has been a paradigm shift in marine research moving from the traditional discipline based methodology employed at the national level by one or more organizations, to a multidisciplinary, ecosystem level approach conducted on an international scale. This increasingly holistic approach to marine research is partly being driven by policy and legislation. For example, the European Commission's Blue Growth strategy promotes sustainable growth in the marine environment including the development of sea-basin strategies (European Commission 2014). As well as this policy driven shift to ecosystem level marine research there are also scientific and economic drivers for a basin level approach. Marine monitoring is essential for assessing the health of an ecosystem and determining the impacts of specific factors and activities on it.

The availability of large volumes of good quality data is fundamental to this increasingly holistic approach to ocean research but there are significant barriers to its re-use. These are in part due to the heterogeneity of the data having been collected by many organizations around the globe using a variety of sensors mounted on a range of different platforms. The resulting data is then delivered and archived in a range of formats, using various spatial coordinate systems and aligned with different standards. This heterogeneity coupled with organizational and national policies on data sharing make access and re-use of marine data problematic. In response to the need for greater sharing of marine data a number of e-infrastructures have been developed but these have different levels of granularity with the majority having been developed at the regional level to address specific requirements for data e.g. SeaDataNet in Europe, the Australian Ocean Data Network (AODN). These data infrastructures are also frequently aligned with the priorities of the local funding agencies and have been created in isolation from those developed elsewhere. To add a further layer of complexity there are also global initiatives providing marine data infrastructures e.g. IOC-IODE, POGO as well as those with a wider remit which includes environmental data e.g. GEOSS, COPERNICUS etc.

Ecosystem level marine research requires a common framework for marine data management that supports the sharing of data across these regional and global data systems, and provides the user with access to the data available from these services via a single point of access. This framework must be based on existing data systems and established by developing interoperability between them. The Ocean Data and Interoperability Platform (ODIP/ODIP II) project brings together those organisations responsible for maintaining selected regional data infrastructures along with other relevant experts in order to identify the common standards and best practice necessary to underpin this framework, and to evaluate the differences and commonalties between the regional data infrastructures in order to establish interoperability between them for the purposes of data sharing. This coordinated approach is being demonstrated and validated through the development of a series of prototype interoperability solutions that demonstrate the mechanisms and standards necessary to facilitate the sharing of marine data across these existing data infrastructures.

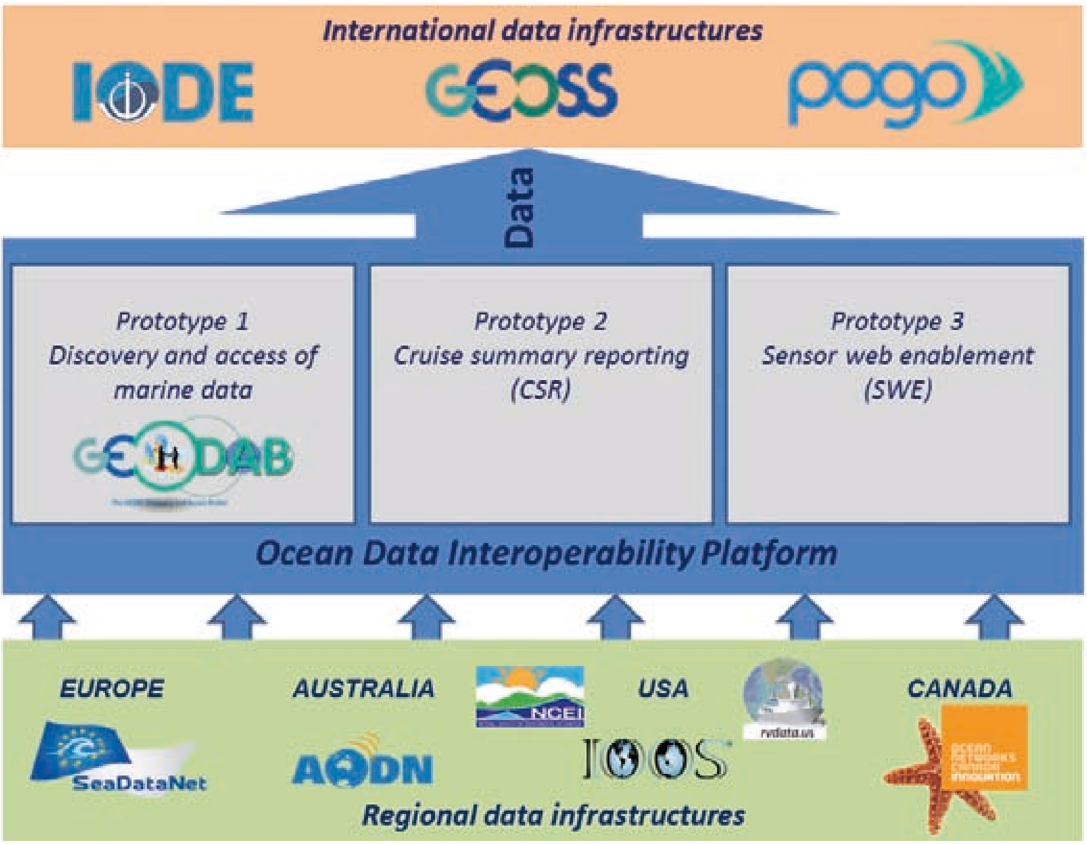


Fig. 1 - Ocean Data Interoperability Platform project.

From SeaDataNet II to SeaDataCloud

Michèle Fichaut, Institut français de recherche pour l'exploitation de la mer (France), michele.fichaut@ifremer.fr

Dick Schaap, Marine Information Service (Netherlands), dick@maris.nl

On behalf of SeaDataNet and SeaDataCloud consortium

Scientists, environmental policy makers, and industry need accurate, homogeneous, integrated marine data sets and reliable services to access these. Oceanographic and marine data include a very wide range of measurements and variables, covering a broad, multidisciplinary spectrum of projects and programmes. They are collected by over a thousand research institutes, governmental organisations and private companies in the countries bordering the European seas. Various heterogeneous observing sensors are installed on research vessels, submarines, aircraft, and moorings, drifting buoys, gliders, floats, fixed platforms and satellites. These sensors measure physical, chemical, biological, geological and geophysical parameters, with further data resulting from the analysis of water and sediment samples for a wide variety of parameters. These data are collected at a very considerable cost. Besides, they are of prime value because they are the reference for any study and, if lost, cannot be replaced. In order to make best use of these data for science and for society, a robust operational infrastructure, based on European and internationally agreed standards, is mandatory. It has to cover data quality and long term stewardship as well as technical and semantic aspects of interoperability.

- The SeaDataNet pan-European infrastructure has been developed by NODCs and major research institutes from 34 countries, in the frame of two European projects (FP6/SeaDataNet – 2006-2011 and FP7/SeaDataNet II – 2011-2015). Over 100 marine data centres are connected and provide discovery and access to data resources for all European researchers. SeaDataNet is a key infrastructure driving several portals of the European Marine Observation and Data network (EMODnet); it complements the Copernicus Marine Environmental Monitoring Service (CMEMS). However, more effective and convenient access is needed to better support European researchers. The standards, tools and services developed must be reviewed and upgraded to keep pace with demand, such as developments of new sensors, and international and IT standards. Also EMODnet and Copernicus pose extra challenges to boost performance and foster INSPIRE compliance. More data from more data providers must be made available, from European and international research projects and observing programmes.

- In this context the new SeadataCloud project, aimed at building upon and expanding the achievements of the SeaDataNet infrastructure, has the following main aims and objectives:

- 1. To enhance and innovate the SeaDataNet standards, products and services offered to an expanded multi-disciplinary community:
  - ❐ To innovate and improve the functionality, performance and quality of the present data

discovery and access services of the SeaDataNet infrastructure by adopting a European cloud environment (EUDAT);

- To expand the range of services of the SeaDataNet infrastructure by specifying, developing and deploying advanced e-services to facilitate individual and collaborative research by using, handling, curating, quality controlling, transforming and processing marine and ocean data into value-added analyses, harmonised data collections, and data products;
- To progress standardisation by adopting ISO, OGC, and W3C standards, in particular by refining existing and formulating new dedicated profiles for marine metadata, and data formats, expanding the SeaDataNet vocabulary services with new lists, developing and deploying enhanced services (Web Processing Services - WPS, Sensor Observation Services - SOS), and Semantic Web applications (e.g. Linked Data) for the SeaDataNet European directories;
- To facilitate the implementation of these standards by offering to data centres a preconfigured and pre-built system including all necessary data management tools, easily deployable and ready to use with minimal setup;
- To develop synergy and tuning between the SeaDataNet network of dedicated ocean and marine data centres, and the upcoming practice of academic data publishing with persistent identifiers (e.g. DOI) for citation purposes;

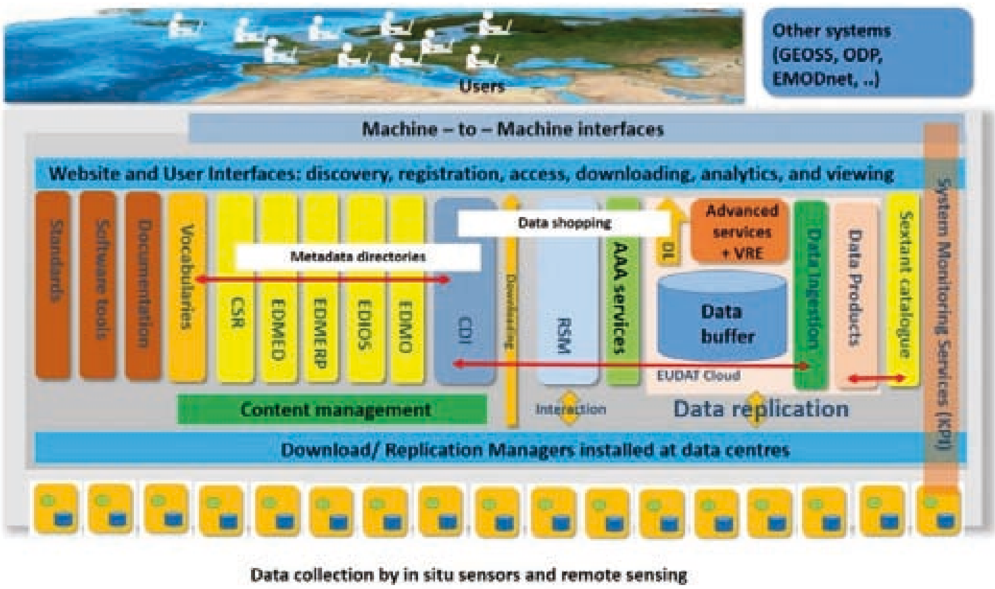


Fig. 1 - Proposed SeaDataCloud architecture with data replication, advanced services and VRE in the cloud.

- 2. To promote the adoption of the protocols and standards developed for interoperability to other key downstream initiatives in the field to expand the communities of data providers and users:
  - To analyse and improve meeting the requirements of lead user communities, such as Copernicus Marine Environmental Monitoring Service (CMEMS), the European Marine Observation and Data network (EMODnet) portals, large marine observing networks (EuroGOOS, AtlantOS, Euro-ARGO, EMSO, JERICO-Next…);

- ❒ To contribute and tune standards with the INSPIRE community;
- ❒ To contribute to global portal initiatives, such as the IOC-IODE – Ocean Data Portal (ODP), GEOSS and the International Council for Science World Data System (ICSU WDS);
- ❒ To increase considerably the number of users and transactions by expanding and engaging potential user communities by promotion, dissemination, development and demonstration of use cases, and organising workshops;
- ❒ To secure more metadata, data input and data circulation in the SeaDataNet infrastructure by further development of the marine data centre network, organizing training workshops, thereby promoting and supporting adoption and implementation of SeaDataNet standards, tools and services;
- ❒ To deploy interoperability solutions and establish exchange with relevant international data management systems in the world, also making use of the Ocean Data Interoperability Platform (ODIP) results, for giving users overview and access to these data sets;

- ■ 3. To present a long-term sustainable arrangement for the integrated SeaDataNet infrastructure and network of the key data centres in Europe for in situ and remote sensing data for marine research (including coastal research) and their resources:
  - ❒ To achieve an overall capacity of all connected data centres on a pan-European scale to provide up-to-date and high quality metadata and data access services for the in situ and remote sensing data managed by their centres;
  - ❒ To formalise the cooperation and to establish an official long-term arrangement with the CMEMS for synergy and mutual exchanges of data and data products;
  - ❒ To develop an exploitation and business plan for long-term sustainability of the integrated infrastructure and take steps towards fostering long-term cooperation and integrated operation of the key data centres.

IODE Ocean Data Portal - platform to build national distributed data systems

Sergey Belov, Russian Research Institute of Hydrometeorological Information WDC (Russia), belov@meteo.ru
Nikolai Mikhailov, Russian Research Institute of Hydrometeorological Information WDC (Russia), nodc@meteo.ru
Tobias Spears, Fisheries and Oceans Canada DFO-MPO (Canada), tobias.spears@dfo-mpo.gc.ca

Introduction

The birth of the Internet has resulted in great change in how science works and the expectations of the scientific community. Although science has always been global, the digital age has improved the visibility of research activities and has made collaboration easier. It is now much easier to access the data and analytical products used by peers in the development of their research publications for other activities. However, in addition to improving the visibility and accessibility of publications via the Web, the data and other analytical products supporting the publication are now expected to be accessible to the global community. Relying on static tables and images is no longer sufficient. In some cases, the accessibility of data and analytical products is a condition placed upon the researcher by the publisher. Although there are increased demands placed upon the researcher, there is the benefit of increased visibility for the researcher and their accomplishments.

Researchers face many challenges when looking to work in this global digital community. Although data and other value added products exist which may advance one's own work, finding and leveraging these resources on the Web presents many challenges. Knowing where to look, how to search, how to access, and how to integrate data and products from many disparate sources can be difficult. Furthermore, contributing one's own data and analytical products for use by others, especially when resources are limited, can present many significant barriers.

The solution to these challenges is a mix of standards, technology, people, and education. Benefits are worth the investment, both to the data provider and the data user. IODE has created the data integration platform though its Ocean Data Portal project that includes a number of components for data connection,creation of metadata, backend components such as service bus, database, portaland web GIS solution and a number of portal services. This platform is aiming to assist its regional activities and also deloping countries to create distributed data systems.

IODE Ocean Data Portal

The International Oceanographic Data and Information Exchange (IODE) was established by in 1961. It strives to enhance marine research, exploitation, and development through exchange of data and information between member states. IODE supports a variety of programs including standards development, technology, data access, capacity building (education). IODE adheres to the IOC Oceanographic Data Exchange Policy – free and unrestricted access to data and information. 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. 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. ODP is developed as a component of the IODE system to provide sustainable data exchange and dissemination infrastructure to achieve the IODE objectives: to facilitate and promote the exchange of 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.

Fig. 1 - IODE ODP global and regional portals.

At present moment IODE ODP has established a global node (NODC of Russia), two regional nodes (China for IODE ODINWESTPAC and Kenya for IODE ODINAFRICA) and one specialized node (SNDM of Argentina) which provides access to more than 200 national, regional and global datasets and products.

In terms of information contents IODE ODP fully depends on the data contributions from IODE Member states and related projects, programmes and initiatives.

Towards the development of an overarching Marine Research Infrastructures information system

Sandra Sá, EurOcean (Portugal), sandra.sa@eurocean.org
Ned Dwyer, EurOcean (Portugal), ned.dwyer@eurocean.org
Marcin Wichorowski, Institute of Oceanology IOPAN (Poland), wichor@iopan.pl

Background

The development, update and maintenance of comprehensive publicly-available online inventories of a wide array of marine research infrastructures is one of EurOcean´s core tasks. Such information has been collected and stored in a number of InfoBases developed in recent years taking advantage of various opportunities and projects, and is an important source of information for the research and operational oceanographic community. The currently active databases are:

- InfoBase on the European Research Vessels - EurOcean_RV.
- InfoBase on Aquaculture Experimental and Research Facilities – EurOcean_AF.
- InfoBase on the Large Exchangeable Instruments - EurOcean_LEXI.
- Marine Research Infrastructures InfoBase - EurOcean_RID.

These InfoBases gather information on over 900 unique pieces of infrastructure, primarily located in Europe. A key objective of EurOcean is to leverage the value of this information by creating synergies with ongoing initiatives and projects. For example, the EurOcean_RV and the EurOcean_LEXI are seamlessly linked to the Virtual Infrastructure in Ocean Research (EVIOR), as part of the FP7 Eurofleets project; the EurOcean_RV is linked directly to the Partnership for Observation of the Global Oceans (POGO) website and the EurOcean_RID is the source of information on Ocean Research Infrastructures for the JPI (Joint Programming Initiative) Oceans.

Opportunity for Rationalisation

Nevertheless the maintainance of separate databases, with slightly varying and sometimes overlapping information became more difficult over time and also as the quantity of information grew. Therefore the challenge was to develop an integrated system in order to harmonise the above four Infrastructure InfoBases, to be compatible with standards and vocabularies deployed by oceanographic organisations, to streamline the update process, to improve the filtering and the displaying functions and to allow for easier evolution of the whole information system. This led to a development of one reference InfoBase, designed as a modular system which is also more coherent in terms of accessibility for all stakeholders in the marine sciences community.

The New database – RID 2.0

The new database is being developed upon the existing Marine Research Infrastructures InfoBase.

The upgraded InfoBase provides a dynamic structure with a core module and additional specialised modules associated with the existing databases. For the first stage three modules are forseen, for the EurOcean_RV, EurOcean_AF and EurOcean_LEXI integration. The attributes of the objects to be available within any new future modules will be proposed by experts and will be deployed to the database model when required.

Other user-friendly capacities are being incorporated in the new database such as:

- Simplification of the categories of infrastructures: this will make it more intuitive and easier to search for users.
- Real-time location for Research Vessels: such live-feeds are appealing and can be very helpful to keep up to date on the status and cruises of the vessels.
- Compatibility with standards, protocols and vocabularies: this will facilitate machine to machine communications and simplify integration and seamless sharing of information with other initiatives.
- Real-time statistics: For a visual insight of the information available in the database (e.g. age of research vessels, operating depths of underwater vehicles)

The data model used for the new InfoBase is being upgraded to ensure coherence of data and compatibility with ISO and other standards being used by the oceanographic community. Reduced redundancy of data across modules and reference to the vocabularies provided by SeaDataNet reduce resources required for the Infobbase management and enhances the overall quality of information provided to users.

Fig. 1 - EurOcean Marine Research Infrastructures Database Interface [EurOcean_RID V1.0].

The challenge ahead towards a RID 3.0

The new database is conceived to be part of an evolving system. The subsequent steps for the development are expected to contain more user-friendly assets that will improve not only the usability of the database but also, the available content. The next set of upgrades will focus on:

- Inclusion of more specialised modules in connection with other European initiatives;
- Generation of tailored reports in accordance with the search made by the user;
- Natural language search.

Conclusion

The streamlining of one of Europe´s premier sources of information on European marine research infrastructures will facilitate the management and future development of the system while simplifying the potential for sharing and synergistic use of its content with current and new initiatives requiring access to comprehensive marine infrastructural information.

SeaDataNet network monitoring services: Features and Statistics

Angelos Lykiardopoulos, Hellenic Centre for Marine Research (Greece), angelo@hcmr.gr
Stavroula Balopoulou, Hellenic Centre for Marine Research (Greece), spmalop@hcmr.gr
Sissy Iona, Hellenic Centre for Marine Research (Greece), sissy@hnodc.hcmr.gr

SeaDataNet (SDN) is a standardized system for managing the large and diverse data sets collected by the oceanographic fleets and the automatic observation systems. The SeaDataNet infrastructure network currently includes national oceanographic data centres of 35 countries, active in data collection. The networking of these professional data centres, in a unique virtual data management system provide integrated data sets of standardized quality on-line. As a research infrastructure, SeaDataNet contributes to build research excellence in Europe and in order to fulfill that it needs to be constantly monitored.

Network monitoring is the use of a system that constantly monitors a computer network for slow or failing components and that notifies the network administrator in case of outages. SeaDataNet has implemented the SDN network monitoring System that monitors the SDN services (web pages, Download Managers, etc.) through Internet to confirm that they are accessible and available, notifies persons in charge to take action in case of service failure and keeps statistics.

The SeaDataNet monitoring plan was divided into nine phases. Firstly, the appropriate software that was going to be used as a base for further development was selected and then a pilot group of partners was set up to test the software. So, the first services were added to the system and the first monitoring data were examined. Plugins were developed (Download Manager plugin) and/or added to the software to improve its functionality. The rest of the partners' local services (Download Managers) were then added gradually to the monitoring system.

A user friendly interface for the local administrators (partners whose services were being monitored) has been developed that uses the SDN CAS authentication system and many features for reporting monitoring results such as service uptime for a selected time period, reports of critical events, histograms, trends and handling of scheduled downtime events. The last implemented feature was the Global Availability Indicator, a percentage value that shows the whole network's uptime for a user-defined period of time.

To improve the reliability of the SDN monitoring system, a false alarm detection system was developed. For this purpose, a second monitoring system that constantly monitors the SDN services was installed and configured. In cases of outages the main monitoring system crosschecks the alarms with the second system then decides if the alarm is true or false and sends notification emails to the local administrators.

The SDN network monitoring system currently monitors 160 services derived from 7 services groups and keeps their statistics. The Global Availability Indicator for the last reported period (September 2014-September 2015) was 99.14% and the average uptime for the SDN Download Managers was estimated to be 98.19%. The above results prove that the SDN Monitoring Systems ensures the SeaDataNet infrastructure's well operability and sustainability which means that all its services are available and accessible by the end-user (scientists, students, data analysts, etc.). The SDN monitoring system will constantly improve to meet the new and forthcoming needs.

Arctic Data Centre, bridging scientific and operational data streams

Øystein Godøy, Norwegian Meteorological Institute (Norway), o.godoy@met.no
Bard Saadatnejad, Norwegian Meteorological Institute (Norway), bards@met.no

Background

The Arctic Data Centre (ADC) is a Data Collection and Production Centre within the WMO Information System (WIS). During the International Polar Year (IPY), the Norwegian Meteorological Institute (METNO) was coordinating IPY operational data streams and Norwegian IPY data management. This bridging of scientific and operational communities has since been a primary focus for the activities undertaken within the context of ADC. It is the end point for project specific data collections hosted by METNO, whether funded by EU (e.g. Damocles and ACCESS) or Research Council of Norway (e.g. DOKIPY).

Content

ADC contains both operational data generated by the Norwegian Meteorological Institute as well as research data generated in research projects nationally or through EU projects. These data range from in situ measurements, through satellite remote sensing products to numerical simulations of the ocean and the atmosphere. Much of the data hosted by the Norwegian Meteorological Institute are not yet included in the catalogue, but there is an ongoing effort to properly document and prepare data for sharing through this catalogue interface as well as to integrate legacy systems.

Functionalities supported

ADC supports metadata exchange (harvest and export), visualisation of data (utilising OGC WMS and OPeNDAP), transformation of data (remapping, reformatting, subsetting utilising OPeNDAP access), automated and ad hoc data upload, file format compliance checking, and subscription to datasets.

A strong focus for ADC technological development has been to standardise ad hoc data delivery in self describing file formats like NetCDF following the Climate and Forecast convention wherever possible. To help users submit properly formatted data, conversion tools, templates and guidance documentation have been developed. This ensures cost effective data management for projects while retaining requirements for long term data preservation through a high degree of automation.

Interoperability

The ADC is developed to integrate with other data management systems. It is an approved WIS DCPC implying that metadata are prepared for exchange with WIS centres. Data hosted locally are primarily served using a THREDDS data server offering access through HTTP and

OPeNDAP as well as OGC WMS for gridded data. Metadata are exposed and can be harvested using OAI-PMH serving GCMD DIF and ISO19115. Translations of harvested metadata are done using XSLT. This also supports translation of content through mapping of controlled vocabularies used by some of the external partners.

Technological development of ADC is done in synergy with developments for WMO Global Cryosphere Watch, ESA/NSC CryoClim, Norwegian Satellite Earth Observation Database for Marine and Polar Research (NORMAP), Norwegian Marine Data Centre (NMDC) and the emerging Svalbard Integrated Arctic Earth Observing System (SIOS) guided by INSPIRE, WIS, GEOSS and SAON/IASC Arctic Data Committee technical requirements where relevant and appropriate.

ADC serves data to the Norwegian Marine Data Centre through exposure of relevant metadata using OAI-PMH and data using OPeNDAP.

Future plans

ADC is currently undergoing a transition to a new technological framework to ensure improved flexibility and modularity. The background for this transition was input from user communities requiring increased flexibility when configuring the user interface. ADC has been supporting a number of scientific projects, e.g. EU FP6 and FP7 projects DAMOCLES and ACCESS, and Research Council of Norway (RCN) projects NORMAP and NMDC. User feedback through these projects and IPY, which except for NMDC all used the same software basis, has been consistent. What the development team thought were important, were not necessarily what the users wanted, Furthermore, an internal evaluation of the current software system revealed insufficient scalability and flexibility, or in other words, the system was unable to serve continuously changing external requirements. In order to achieve greater flexibility and scalability, the new platform places emphasis on standard interfaces using standard protocols between its constituent modules. The constituent modules are therefore loosely coupled through HTTP and protocols based on HTTP. This ensures easier integration or replacement of software components as well as the ability to spread the load on multiple servers if required. Continuing this process, user feedback will be collected through web based tools (issue tracking for feature requests, polls, etc.) as well as through workshops. The intention is to continuously show the users the status of feature requests. Interaction with the user community will be coordinated by two RCN projects focusing on data management and user interaction. These are the Norwegian Scientific Data Network which is an infrastructure project, and GeoAccessNO which is a pilot project evaluating geoscientific data management.

Norwegian Marine Data Centre

Arnfinn Morvik, Institute of Marine Research (Norway), arnfinn.morvik@imr.no
Terry Hannant, Institute of Marine Research (Norway), terry.hannant@imr.no
Armin Pobitzer, Christian Michelsen Research (Norway), armin.pobitzer@cmr.no
Helge Sagen, Institute of Marine Research (Norway), helge.sagen@imr.no

Seamless access to marine data

The main objective of NMDC is to serve the marine science community with seamless access to documented marine datasets covering waters of Norwegian interests. In order to fullfill this objective, a national infrastructure for marine datasets has been established.

NMDC is a distributed infrastructure for data with a central node holding a centralized metadata catalogue. The central node harvests metadata from the partners in the project and a web based data cataloge viewer gives access to the metadata.

The project has 16 national partners and is coordinated by IMR.

System design

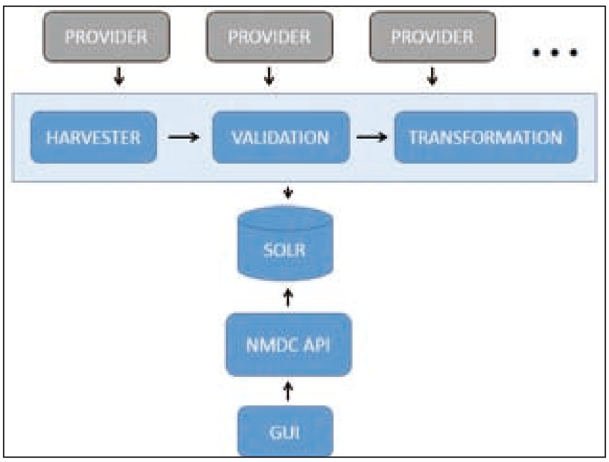


Fig. 1 - Shows the central node and the various data providers.

The central node harvests metadata from the data providers. The data providers are responsible for setting up repositories that expose structured metadata via the protocol OAI-PMH. The central node requests to harvest metadata once every day.

The central node accepts metadata in either ISO 19139 or GCMD DIF format.

After harvesting, the metadata are validated using standard xsd schema. The data providers are notified if their metadata fails to pass the validation.

The last part of the harvesting process is to transform the metadata to an internal NMDC metadata format.

The transformation serves two purposes: All metadata are stored in the same format, and the format is designed to be well suited for geographical search.

Harvesting, validation, and transformation are standalone processes using AciveMQ as messaging server. Programs are written in Java and the JMS API are used to send and receive messages.

The central node uses Solr as its search platform. After transformation, the metadata are imported into Solr and the data are indexed due to a set of search criteria’s defined by the central node. The current setup of the central node runs a Solr v5 installation. There is no immediate plan to move towards v6.

The NMDC API is a REST API available to anyone that wants to build applications on top of the central node. Solr is not directly accessible from the outside, and all searching in the indexed metadata are performed by calling methods through the NMDC API.

Solr can return various different formats, but JSON has been chosen as the only format the NMDC API will return.

Available methods are: getFacets, getMetadataDetails, getData, and search

Applications

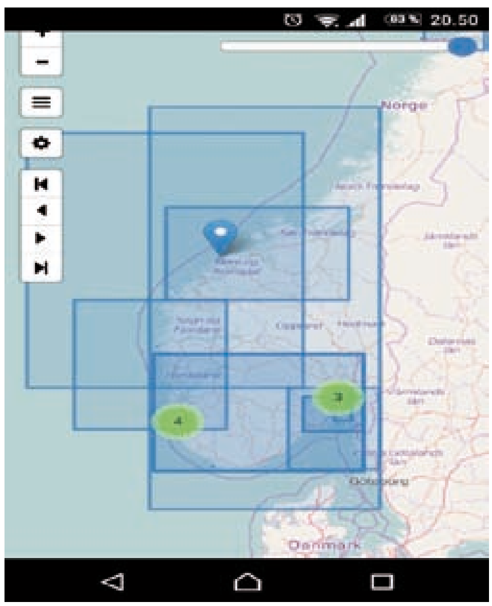
There are currently 2 applications using the NMDC API:

- The data catalogue as a web application on http://prod1.nmdc.no/UserInterface/
- An experimental map-based search interface available at Google Play: NMDC Mobile Portal

With the web application, the user can search for metadata on scientific keywords, data provider, free text, geographical coverage and temporal coverage. All criteria's can be combined in order to narrow the search.

It is possible to view the metadata for each dataset in more detail, and download the relevant dataset following the URL's given in the metadata files.

It will also be possible to download all the data from a selected group of metadata without using the individual download links from each dataset.



The map-based search uses the same NMDC API as the web application. The outline of all the available datasets are visible in a map as a rectangle for datasets with a geographical extent, or as a pin for time series from a point location. The user can select individual datasets to get more detailed information about the selected data.

The number of rectangles visible depends on the current geographical extent of the map. The default setup will only show datasets whose extent is completely within the current map extent. It is possible to change this to allow datasets with extent partially intersecting the map extent to be visible.

The same metadata details as in the web application are available.

CNR-ISMAR in situ observations network: new approaches for an interactive, high performance, interoperable system

Stefano Menegon, ISMAR Centro Nazionale delle Ricerche (Italy), stefano.menegon@ismar.cnr.it
Pierluigi Penna, ISMAR Centro Nazionale delle Ricerche (Italy), pierluigi.penna@ismar.cnr.it
Mauro Bastianini, ISMAR Centro Nazionale delle Ricerche (Italy), mauro.bastianini@ismar.cnr.it
Giuseppe Stanghellini, ISMAR Centro Nazionale delle Ricerche (Italy), giuseppe.stanghellini@ismar.cnr.it
Francesco Riminucci, ISMAR Centro Nazionale delle Ricerche (Italy), francesco.riminucci@ismar.cnr.it
Tiziano Minuzzo, ISMAR Centro Nazionale delle Ricerche (Italy), tiziano.minuzzo@ismar.cnr.it
Alessandro Sarretta, ISMAR Centro Nazionale delle Ricerche (Italy), alessandro.sarretta@ismar.cnr.it

This paper describes the activities in ISMAR (Institute of Marine Sciences[1] of the Italian National Research Council) for the creation of an in-situ observational network able to: gather data from a distributed network; allow an efficient visualization of time-series data; access and share harmonised data through interoperable and standard services.

Fig. 1 - Web interface of the portal.

Context and objectives

During the last years, the need for a unique visible access point for the ISMAR observational network emerged from various research projects and, in particular, in the projects RITMARE, LTER and JERICO. This is also related to the request, by various European-funded projects, that data should be made discoverable, accessible and sharable using standards formats and services (NetCDF, ODV, SOS…).

More in general, the ISMAR observational network aims to ensure archiving and storage of the historical time series data of the Institute (useful to long-term research and climate changes research) and to provide an efficient and convenient support for the operational model (through real-time data streams). These goals are achieved through harmonization of data management,

1 http://www.ismar.cnr.it/index_html-1?set_language=en&cl=en

using standard formats and interoperable services and creating a unique access point (web portal) for all ISMAR data related to in-situ observations, where to visualize and access real-time and archived data; in addition to that, a clear data policy and standard open licences are adopted to facilitate dissemination and reuse of data.

ISMAR manages various marine/atmospheric measurement sites that have had different implementation paths (funds, specialised personnel, IT infrastructures) and that now need to be commonly organised and managed to allow a better use and dissemination of the data collected.

Currently, 5 stations for marine/meteorological observations (Acqua Alta, E1, GB, Telesenigallia, Gargano) and 3 meteorological stations (ISMAR-Venezia, ISMAR-Ancona, ISMAR-Lesina) are active, while 4 more are in plan to be included (buoys and moorings), with a clear focus on the Adriatic Sea region. In addition to standard meteorological parameters, the system collects data about the water column (i.e. temperature, salinity, sea level, oxygen, waves, currents and other water quality parameters).

Other platforms or acquisition systems might be added in the future to extend the geographical scope and the availability of environmental parameters observed.

Architecture and IT solutions

The adopted solution aims to combine approaches based on interoperability standards (i.e. SensorML) with systems that are optimized for handling time series data.

The core of the system is a centralized database implemented with InfluxDB, an open source database created specifically to handle time series data with high availability and high performance requirements. InfluxDB[2] has several interesting features like a simple, high performing write and query HTTP(S) API, a flexible and dynamic data model and is part of a stack of integrated components and tools that allow to implement a complete solution for collecting, storing, visualizing and alerting on time-series data. The Graphical User Interface (GUI) is based on Grafana[3], an open source, feature rich metrics dashboard and graph editor.

Data are collected from the measurement stations and stored without further modification in a "raw data" database. At the same time, a data processing engine computes an automatic data quality control in order to clean the data, underlying statistical anomalies and, if necessary, triggers alerts. The "cleaned" data are stored in a "higher quality" database.

InfluxDB and Grafana are commonly used as part of the ICT network and infrastructure monitoring system and, if on the one hand they are simple to use and very efficient, on the other they do not offer some basic features for an environmental and scientific monitoring network (i. e. Interoperability services like THREDDS and SOS, interchange formats like NetCDF, rigorous description of metadata).

In order to handle such issues we have, first of all, adopted already existing conventions and best practices (i.e. the use of SeaDataNet Agreed Parameter Groups[4]); then, we have extended the platform with specific functionalities and tools, for instance a plugin to export data in NetCDF format (i.e. SeaDataNet NetCDF - CF) and a plugin to harvest and import data from SOS (Sensor Observation Services) servers.

2 https://influxdata.com/time-series-platform/influxdb/

3 http://grafana.org/

4 http://vocab.nerc.ac.uk/collection/P03/current/

Aggregation and processing of data originating from heterogeneous sources for multidimensional analyses

Marcin Wichorowski, Institute of Oceanology, Polish Academy of Sciences (Poland), wichor@iopan.pl
Katarzyna Błachowiak-Samołyk, Institute of Oceanology, Polish Academy of Sciences (Poland), kasiab@iopan.pl

Approach to an aggregated data warehouse

Interdisciplinary analyses, carried out on extensive sets of data, are the new and promising trend, because they can provide valuable information for many various disciplines (e.g., taxonomy, oceanography and ecology). Completing data from different sources and fields is one of the most current trend in ecological and biological sciences. The examples of modern and high-tech databases providing multiple data from many disparate disciplines (i.e., systematic, ecology) in temporal scales are: OBIS (Ocean Biogeographic Information System) and ERMS (European Register of Marine Species). The main aims of such databases are

- Collection of operational data from different research, cruises and years,
- Organization of faunal and environmental data in a standardized form,

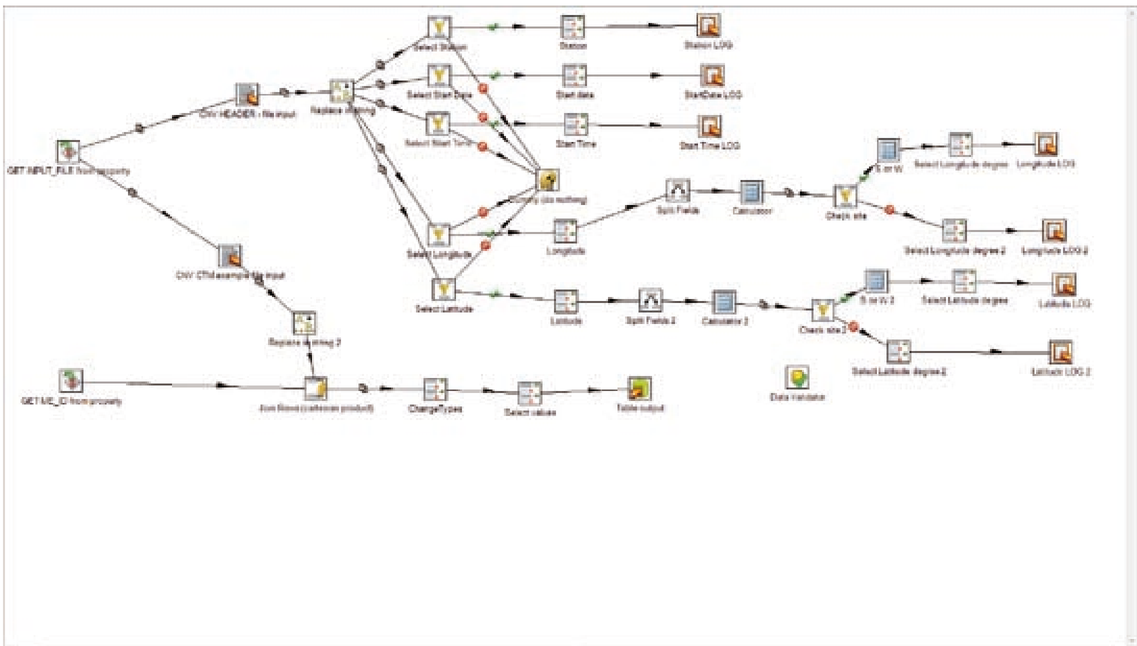


Fig. 1 - Definition of data extraction and transform process.

- Preservation of data for studies of long-term temporal changes and
- Preparing a comprehensive database with easy access to raw data.

Analysis of aggregated data allow to identify phenomenons and patterns hard to distinguish using traditional methodologies of research shifting the whole data management process towards Big Data. Such approach lead to "fourth paradigm" in scientific research. Aggregation of data is possible using automatized processes and ETL (extract, transform and load) tools. Example of definition of process could be provided in graphic form and interpreted than by Pentaho Business Analytics. However this is commercial tool, open tools, community software and open scientific data are most important supporters of modern scientific research.

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. Institue of Oceanology has developed infrastructure to face these identified chalenges and raise operational performance of data processing and efficiency of data analysis.

Integrating Data and Information Across Observing Systems

Kevin O'Brien, University of Washington/JISAO (United States of America), kob@uw.edu

Steven Worely, University Corporation for Atmospheric Research (United States of America), worley@ucare.edu

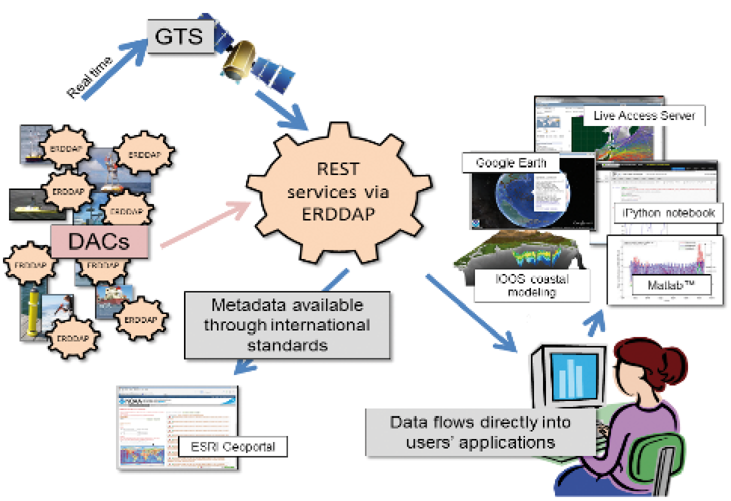
Bob Simons, National Oceanic and Atmospheric Administration (United States of America), bob.simons@noaa.gov

Benjamin Pfeil, Bjerknes Climate Data Centre (Norway), Benjamin.Pfeil@uib.no

Eugene Burger, National Oceanic and Atmospheric Administration (United States of America), Eugene.Burger@noaa.gov

"Integrated data management" is best understood by contrasting it with what has historically been the norm for ocean data management. Ocean observations transmitted through the GTS have very high societal value, due to their potential use in ocean forecast models, climate state estimates, etc. Yet, for the average science data user, the ability to access and use data from the GTS has been such a prominent barrier that it was all but impossible. This creates a situation where much of this data is underutilized, despite its very high scientific value and the large investments made to collect it.

The potential for improved integration among delayed mode global ocean data assembly centers (GDACs) is another significant opportunity. Currently, each platform assembly center typically provides its own web site at which users (humans, but not machines) can select data of interest. Data can then be downloaded in the file format(s) available from that network. A user desiring interior ocean temperatures from both Argo and TAO, for example, must learn to navigate two independently designed (and very distinct) Web sites, each with its own unique technical language and each with its own style of selecting regions, times, platforms, and variables with various imposed groupings. When the use of these distinct Web sites have been mastered, a larger barrier remains – that of unstandardized data formats limiting the fitness of use of the data. Custom software is often needed to read the data from



distinct networks. Variable names and metadata differ between providers. Non-experts would often struggle to pursue such a task to completion.

Rather than having data systems and offerings developed around individual networks, an integrated framework promises interoperable access that crosscuts the networks and offers a more intuitive research view for the users (Fig. 1). In practical terms, this integrated access will allow users to focus on Essential Ocean Variables (EOVs), rather than on platform specific measurements. This framework will create interoperable access to data and metadata, and is not simply another data portal. Data providers and groups will be able to leverage the interoperable framework to build customized data portals, and with little advanced planning will be able to expand their impact with broader access support from the framework. As is, with the myriad of access services provided by the framework, users will already be able to access and use the data with clients they prefer (e.g. MatLab, iPython Notebook, etc.). They will not need to navigate through many different download paths or be subjected to having to reformat data specifically for their favorite client.

In this presentation, we will discuss how we have utilized such a framework to provide interoperable access to real time ocean observations from the GTS. We will also describe our efforts to integrate data available from the heterogeneous platform networks providing ocean observations in support of the Tropical Pacific Observing System (TPOS) 2020 effort. We will discuss how this framework, utilizing a tool called ERDDAP (http://coastwatch.pfeg.noaa.gov/erddap), leverages the OPeNDAP protocol and supports RESTful web services, providing machine to machine access and easy client access. We will also highlight some complimentary efforts already underway within Europe to provide this integrated framework and interoperable data access.

The end of reporting deadlines? Optimising dataflow by harvesting - a case study using open source REST-server technology from SMHI in corporation with ICES

Patrik Strömbergr, Swedish Meteorological and Hydrological Institute (Sweden), patrik.stromberg@smhi.se
Arnold Andreasson, Swedish Meteorological and Hydrological Institute (Sweden), arnold.andreasson@smhi.se
Nils Nexelius, Swedish Meteorological and Hydrological Institute (Sweden), nils.nexelius@smhi.se

The scientist will always need access to the highest quality, largest collection of, and most recent version of the data. Advances in technology can meet this increasing demand via technical solutions. In order to ensure these demands are met, one way forward is to grant open access, and using machine to machine interfaces. **SMHI have developed methods and systems to handle complex marine biogeochemical and abundance types of datasets**. The systems are **sharkweb**. smhi.se (human interface) and **SHARKdata**.se (machine interface). The data on SHARKdata.se are currently being harvested by portals such as EMODNet Biology (and henceforth to EUROBIS/OBIS), and Swedish Lifewatch systems. There is also an ongoing pilot study in cooperation with the ICES data centre. The aim of the pilot is to investigate the possibility for the national submissions of data to be automated. This will: I. reduce the amount of manual labour, II. ensure the most recent version in the ICES DOME (Marine Environment) database,. III) allow the quality checks performed by ICES to be called by the SHARKdata.se system directly, which collectively leads to IV. higher quality of the data. All technology used is open source (the MIT license) and hence for anyone to download and build their own ("SHARKdata") system. For the scientist it is possible to use for example R or Python to set up any type of analysis using the most up to date data from the Swedish Oceanographic data center. Examples are published on the server SHARKdata.se

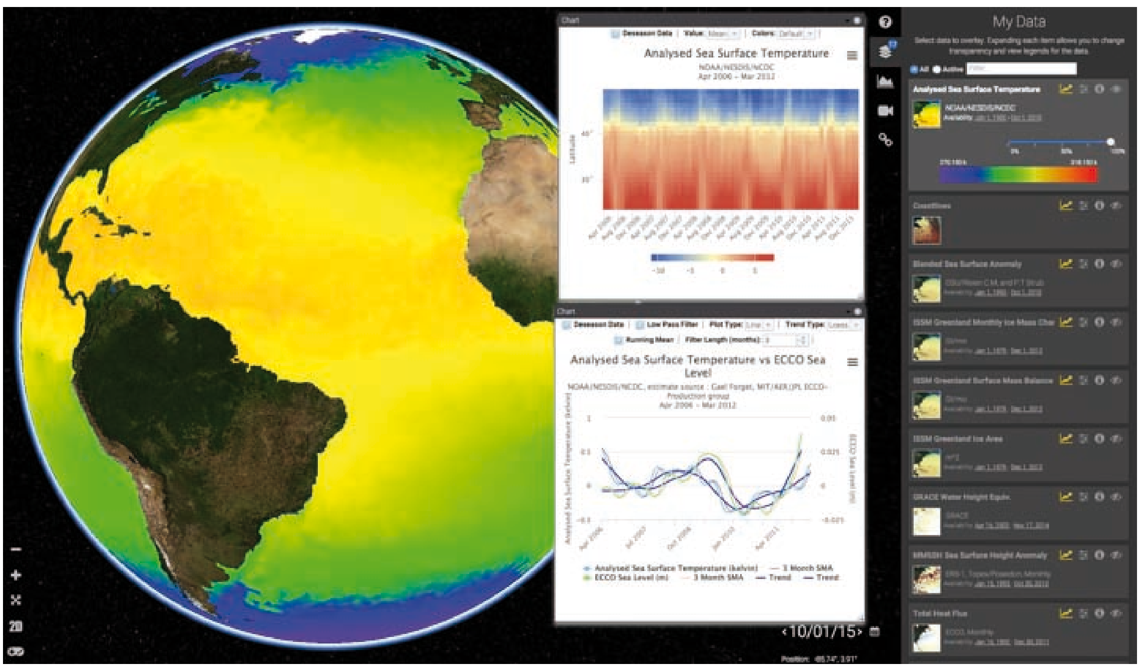
Taking on Big Ocean Data Science

Thomas Huang, California Institute of Technology/Jet Propulsion Laboratory (United States of America), thomas.huang@jpl.nasa.gov

Almost all of the existing data analysis solutions for Earth science are still built around large archives of granule files, which yield poor performance. Common data access solutions, such as OPeNDAP and THREDDS, provide web service interfaces to archives of observational data. They perform poorly because they are still built around the notion of files. It is very time consuming for climate scientists to conduct research that involves the generation of time series over large spatial region and/or over decades of observational data.

Map Reduce is a well-known paradigm for processing large amounts of data in parallel using clustering or Cloud environments. Unfortunately, this paradigm doesn't work well with temporal, geospatial array-based data. One major issue is they are packaged in files in various sizes. These data files can range from tens of megabytes to several gigabytes. Depending on the user input, some analysis operations could involve hundreds to thousands of these files. Moving these files from storage node to computing node is very inefficient where the majority of the time is spent on I/Os.

NEXUS is a Deep Data Platform being developed at NASA's Jet Propulsion Laboratory. It takes on a different approach in handling file-based observational temporal, geospatial artifacts by fully leveraging the elasticity of Cloud Computing environment. Rather than performing on-the-fly file I/Os, NEXUS breaks data artifacts into small data tiles where they are managed by



a Cloud-scaled database with high-performance spatial lookup service. NEXUS provides the bridge between science data and horizontal-scaling data analysis. It provides a workflow to divide science artifacts into small data tiles, store in a cloud-scaled database where they can be quickly retrieved through a high performance spatial search registry. This platform simplifies development of big data analysis solutions by bridging the gap between files and Map Reduce solutions such as Hadoop and Spark.

This presentation discusses applications of NEXUS in three Big Ocean Data Science projects at NASA in relation to the definition of Big Data, which evolves around the 3Vs model, Volume, Velocity, and Variety.

- The NASA Sea Level Change Portal (https://sealevel.nasa.gov) is the official NASA's one-stop information portal for all news and data relevant to sea level rise. The portal has a built-in Data Analysis Tool (DAT) to provide climate scientist high-performance visualization and a suite of on-the-fly analysis tools.
- The NASA OceanXtremes: Oceanographic Data-Intensive Anomaly Detection and Analysis Portal is a cloud-based analytic service that enables execution of domain-specific, multi-scale anomaly and feature detection algorithms across the entire archive of ocean science datasets.
- The NASA's Distributed Oceanographic Match-Up Service (DOMS) is a cloud-based reconciliation of satellite and in-situ datasets

Adding Big Data's Velocity to Oceanographic Data Systems

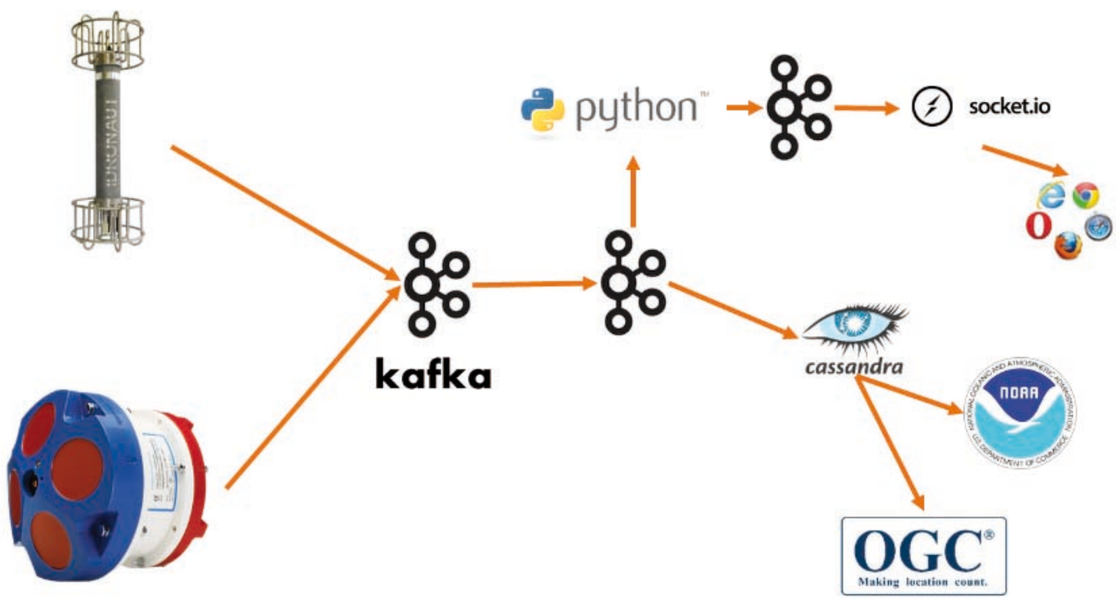
Adam Leadbetter, Marine Institute (Ireland), adam.leadbetter@marine.ie
Damian Smyth, Marine Institute (Ireland), damian.smyth@marine.ie
Rob Fuller, Marine Institute (Ireland), rob.fuller@marine.ie
Eoin O'Grady, Marine Institute (Ireland), eoin.ogrady@marine.ie

Introduction

In August 2015, a new seafloor observatory was depolyed in Galway Bay, Ireland. This sensors on the observatory platform are connected by fibre optic cable to a shore station, and from there a broadband connection allows data to be delivered to the Marine Institute's data centre.

This setup allowed the development of a new data acquisition system which takes advantage of the open source streaming data solutions which have been developed in response to the Big Data paradigm, in particular the "Velocity" aspect. Big Data is commonly defined in the following terms and we also expand on their application to marine data:

- *"Volume"*: Big Data deals with data volumes which cannot be easily managed on a personal computer, stretching away from the megabyte and gigabyte scales towards the terabyte and petabyte scales,
- *"Velocity"*: The speed at which data are acquired, processed and made available is key to Big Data and for oceanographic data represents a move from batch, delayed mode processing through near real-time processing into genuine real-time data availablilty,



- *"Variety"*: There is a range of data types and sources in a Big Data system, as is typical in an oceanographic data system, and incorporates data types which move beyond those easily incorporated into database tables such as images and video,
- *"Veracity"*: The accuracy of the data is important in Big Data applications, as it is in oceanographic data systems as can be seen by the various flagging schemes available (IOC 2013),
- *"Variability"*: The context of a data point is important in interpreting it within the Big Data paradigm. This is also important in an oceanographic context when integrating data from models, remote sensing systems and in-situἀ observations.

The Streaming Data System for Galway Bay

The streaming data system is an analytic computing platform that is focused on speed. This is because these applications require a continuous stream of often unstructured data to be processed. Therefore, data is continuously analysed and transformed in memory before it is stored. Streaming data applications typically manage a lot of data and process it at a high rate of speed. Because of the amount of data, it is typically managed in a highly distributed clustered environment. This satisfies the real-time data delivery requirement of the Big Data paradigm.

There are many open source software components which allow the deployment of such streaming data systems which have been developed by web companies such as LinkedIn. The Galway Bay cable observatory data system makes extensive use of the Apache Kafka messaging queue to relay data messages in a managed way from shore station to data centre and between processing scripts and data store (see Fig. 1). Stream processing software, in particular the Apache Storm software, was investigated but is not mature enough to deploy in a production environemnt. Finally, a time-series optimised Big Data datastore, Apache Cassandra, has been deployed for archiving data.

Evolving Data Standards

Access to the outputs of the Galway Bay observatry data system is provided visually via dashboards; via a file store archive; through an ERDDAP interface for downloading subsets of the data; and using standardised interfaces for programmatic access – namely the Open Geospatial Consortium's Sensor Observation Service and the ISO standard MQTT messaging protocol.

The Sensor Observation Service has been developed as a component of the streaming data system and interacts directly with the Cassandra database, which also provides a backend to the ERDDAP interface. Mindful of the "Born Connected" approach (Leadbetter *et al.,* 2016), the fields of the Observations & Measurements records delivered by the Sensor Observation Service use URLs from the NERC Vocabulary Server where possible. Similarly, the emergence of new JSON-encodings for the Observations & Measurements data model (Cox and Taylor 2015) has also allowed a more consumer-friendly serialisation of the results to be produced.

In order to allow true streaming of the data into client software, a MQTT broker has been deployed. MQTT is an important piece of the Internet of Things architecture as it is currently the only accepted standard for machine-to-machine communications in that domain. The Galway Bay observatory implementation required development of an Apache Kafka to MQTT bridge which has been incorporated into the main software branch and exposes the message queue using a standard protocol and in more a secure manner than simply exposing the Apache Kafka server.

Future developments within the MQTT broker will define standardised representations of the data within the messages exposed, and further the Born Connected concept.

Conclusions

Big Data, streaming data and the Internet of Things are all currently highly important topics in software engineering at enterprise scale. However, they have not been tranlsated well to the Earth and Space Science Informatics realm through the lack of ability to connect them with the pre-existing, well-defined and highly-adopted data standards. In this paper, we have shown that the connection between the two paradigms is possible and that there is further work to do in the standardisation of data messages in the Internet of Things domain. This final challenge may be the precursor to connecting many more instruments much more quickly to the well established data systems of the oceanographic data management world.

References

Cox, S., Taylor, P., 2015. *OM-JSON - a JSON implementation of O&M*. Presentation to the OGC Sensor Web Enablemenr Domain Working Group, Nottingham, UK, September.

Intergovernmental Oceanographic Commission of UNESCO, 2013. *Ocean Data Standards*, Vol.3: Recommendation for a Quality Flag Scheme for the Exchange of Oceanographic and Marine Meteorological Data. (IOC Manuals and Guides, 54, Vol. 3.) 12 pp. (English.) (IOC/2013/MG/54-3)

Leadbetter, A., Cheatham, M., Shepherd, A., Thomas, R., 2016. *Linked Ocean Data 2.0*, in Diviacco, P., Leadbetter, A., Glaves, H., Oceanographic and Marine Cross-Domain Data Management for Sustainable Development, Hershey, PA: IGI Global.

Sensor Nanny: support observation operators with equipment and data management services on the cloud

Thomas Loubrieu, Institut français de recherche pour l'exploitation de la mer (France), thomas.loubrieu@ifremer.fr

Jérôme Détoc, Institut français de recherche pour l'exploitation de la mer (France), jerome.detoc@ifremer.fr

Catherine Borremans, Institut français de recherche pour l'exploitation de la mer (France), catherine.borremans@ifremer.fr

Arnaud Thorel, ASI Informatique (France), athorel@asi.fr

Hamza Azelmat, École nationale d'ingénieurs de Brest (France), hamza.azelmat@ifremer.fr

In marine sciences, the diversity of observed properties (from water physic to contaminants in observed in biological individuals or sediment) and observation methodologies (from manned sampling and analysis in labs to large automated networks of homogeneous platforms) requires different expertise and thus dedicated scientific program (e.g. ARGO, EMSO, Research vessels…). However, all of them requires similar IT services to support the maintenance of their network (instrument model documentation, calibrations, deployment strategy, spare part management...)

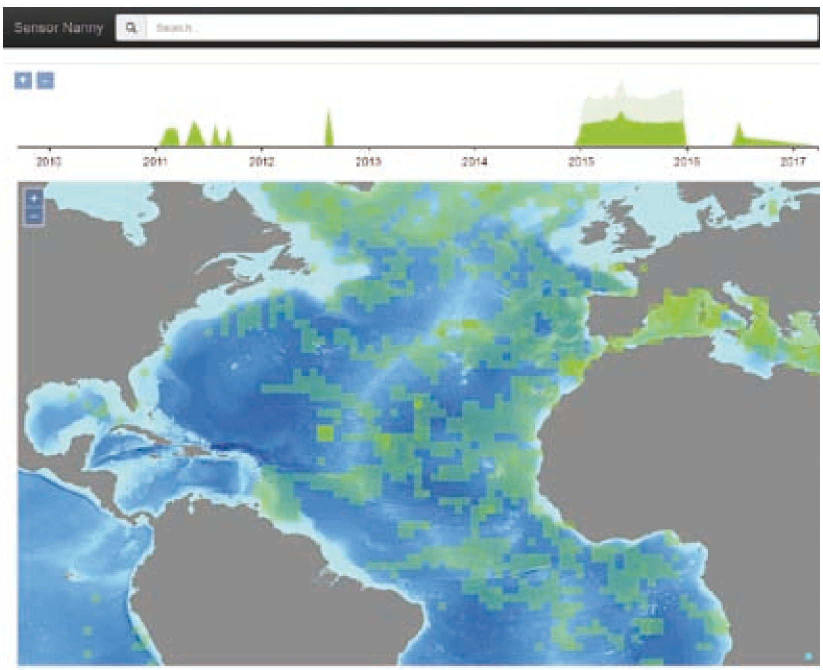


Fig. 1 - Snanny web portal showring observation spatio-temporal density.

and their data management. This is actually also a concern at a trans-disciplinary level (ICOS for carbon, EPOS for solid earth or SIOS for Arctic) identified in a “sensor registry” use case in ENVRIplus project.

Upstream the current scope of the well-established European marine data management infrastructures (NODCs, ROOSs), a range of services are available or under-development to support scientific programs and observation operators in this perspective. They are integrated together by implementing the Sensor Web Enablement standards and will improve the quality, cost-effectiveness, and latency of data integration in the data management infrastructures.

The solution’s corner-stone, sensor nanny, is a collaborative web environment enabling data providers to describe their observatory and drop or synchronize observation results “on the cloud”.

The application provides an on-line editor to graphically describe, literally draw, their observatory. The observatory description is composed by the user from a palette of hundreds of pre-defined sensors or hardware which descriptions is extracted in sensorML from the Fixo3-yellow pages (http://www.esonetyellowpages.com/).

In addition, the data providers can safe-guard their observation results by uploading or synchronising in real-time local data resources. The users can thus share their data on-line with their partners. The native format for the observatory and observation descriptions are sensorML and O&M from the OGC/Sensor Web Enablement suite applying profiles discussed in Ocean Of Tomorrows and ODIP projects.

The observatory descriptions and observation data are indexed so to be very fluently browsed, filtered and visualized in a portal. This has been demonstrated with up to 2.5 million observation points from French research vessels, ARGO profiling floats and EMSO-Azores deep sea observatory.

The key components used for the development are owncloud for the file synchronization and sharing and elasticSearch for the scalable indexation of the observatories and observations.

The foreseen developments aim at handling instrument maintenance support (calibration, spare parts) based on LabCollector. Within JERICO-NEXT and AtlantOS further observation networks (e.g. HF radars) will be integrated. Tools for sharing data in NODCs, ROOSes and publish datasets as DOIs are also being developed.

Enhancing User Access to Oceanographic Data Through Commercial Cloud Services

Roger Proctor, Integrated Marine Observing System (Australia), roger.proctor@utas.edu.au
Peter Blain, Integrated Marine Observing System (Australia), peter.blain@utas.edu.au
Dan Fruehauf, Integrated Marine Observing System (Australia), malkodan@gmail.com

The Integrated Marine Observing System (IMOS) is a national project funded by the Australian government established in 2007 to deliver ocean observations to the marine and climate science community. Its mission is to undertake systematic and sustained observations and to turn them into data, products and analyses that can be freely used and reused for broad societal benefits.

Historically IMOS information infrastructure has relied on national e-research infrastructure cloud services such as those supported by the University of Melbourne and the University of Tasmania. These services, being research grade, suffer from unplanned outages, unexpected loss of service and, at times, data corruption. As IMOS has matured as an observing system, now in its 10th year, and community exposure to IMOS data has increased, expectation on the system's availability and reliability has also increased. IMOS is now seen as delivering 'operational' information. In responding to this expectation, IMOS has relocated its services to the commercial cloud service Amazon Web Services (AWS). The relocation has significantly increased IMOS' measured availability.

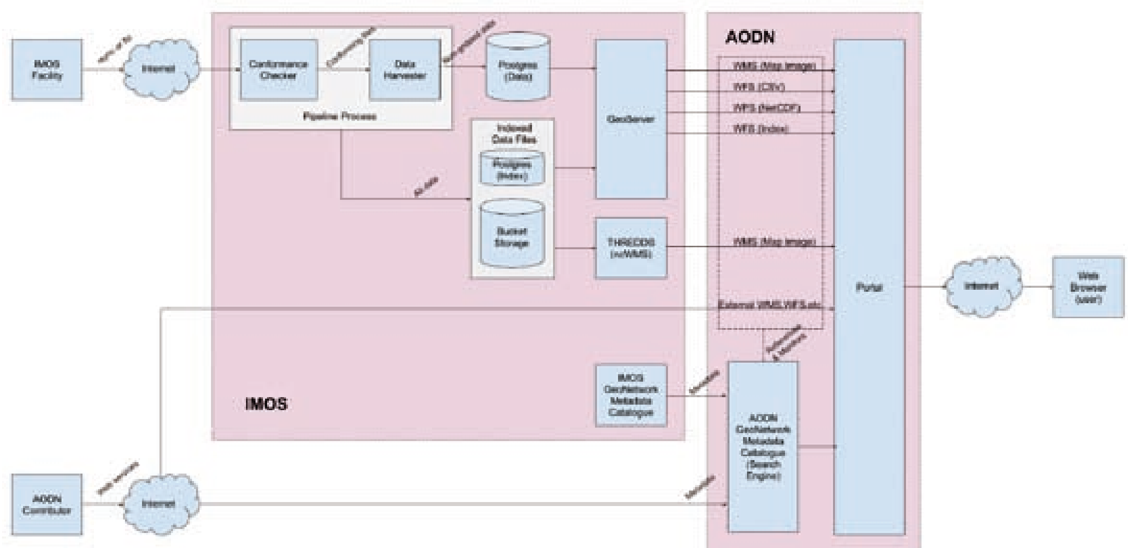


Fig. 1 - IMOS infrastructure at AWS.

Using AWS has enabled IMOS to improve the system architecture utilizing more advanced features like object storage (S3 - Simple Storage Service) as opposed to a traditional file system, Fig. 1. S3 has virtually unlimited affordable storage and supports redundancy (stores copies in multiple facilities and devices) and versioning (keeps tracks of a single object's multiple versions). Redesigning the IMOS system architecture around S3 has improved data availability and resilience while protecting against human errors in data handling and providing a disaster recovery plan.

POSTERS

An Overview of the WMO Information System (WIS)

Patrick Gorringe, EuroGOOS (Belgium), patrick.gorringe@eurogoos.eu
Erik Buch, EuroGOOS (Belgium), erik.buch@eurogoos.eu

WMO and Data Management

The mission of the World Meteorological Organization (WMO) includes facilitating scientific understanding, and the comprehensive exchange of information world wide. Data is collected and exchanged around the clock by the 189 WMO Members--data from sixteen satellites, one hundred moored buoys, six hundred drifting buoys, three thousand aircraft, seven thousand ships, and ten thousand land-based stations.

From the beginning of WMO, timely data exchange has been crucial. Operators have been routing data messages over international lines using systems dedicated to WMO. This is the WMO Global Telecommunication System (GTS). Now WMO is taking a step beyond managing messages. In committing to WIS, WMO moves to an overarching approach based on managing data and information.

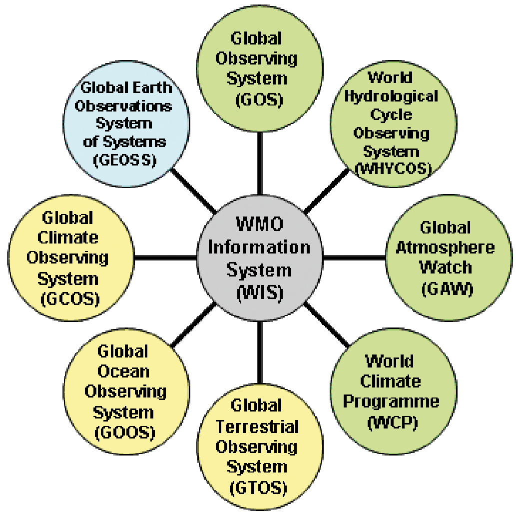


Fig. 1 - WIS Component Systems.

WIS Implementation

In 2003, WMO Members decided that WIS will use international industry standards for protocols, hardware and software; build on the existing WMO GTS, with special attention to a smooth and coordinated transition; provide time-critical data exchange, as well as data access and retrieval services; and support all WMO and related international programmes. The most visible, new facility provided by WIS is the comprehensive Discovery, Access and Retrieval (DAR) catalogue for all of the data, products, and services encompassed by WIS and has been operational since Jan 2012.

WIS Benefits

- WIS enhances the collection of critical data needed to monitor and predict aspects of the environment, including hazards;
- WIS catalogues the full range of data and products, simplifying search and assuring equitable access per WMO policies;

- WIS enhances the availability of time-critical data and products at centres in all nations, ensuring the effective provision of services to their populations and economies;
- WIS opens up WMO's private network (the GTS) to other types of environmental data so that all programmes have stronger infrastructure support; and
- WIS exploits opportunities as they become available with technology innovation.

Cross-cutting Task Team for Integrated Marine Meteorological and Oceanographic Services within WIS (TT-MOWIS)

JCOMM decided in 2015 to establish a Task Team, TT-MOWIS, with the objective to find ways to increase the availability of marine meteorological and oceanographic data in the WIS system. TT-MOWIS will work on:

- Formulating a "Guidance for integrating marine meteorological and oceanographic services within WIS" and having it approved by JCOMM and its mother organizations IOC and WMO;
- Promote the importance of being part of the WIS system to a number of large and important preselected marine datacentres worldwide.

EARS : Repositioning data management near data acquisition

Thomas Vandenberghe, Royal Belgian Institute of Natural Sciences (Belgium), tvandenberghe@naturalsciences.be
Paolo Diviacco, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), pdiviacco@ogs.trieste.it
Jean-Marc Sinquin, Institut français de recherche pour l'exploitation de la mer (France), jean.marc.sinquin@ifremer.fr
Jordi Sorribas, Consejo Superior de Investigaciones Cientificas (Spain), sorribas@utm.csic.es
Raquel Munoz, Consejo Superior de Investigaciones Cientificas (Spain), rcasas@utm.csic.es
Oscar Garcia Sanchez, Consejo Superior de Investigaciones Cientificas (Spain), ogarcia@utm.csic.es
Serge Scory, Royal Belgian Institute of Natural Sciences (Belgium), serge.scory@naturalsciences.be
Yvan Stojanov, Royal Belgian Institute of Natural Sciences (Belgium), ystojanov@naturalsciences.be

The EU FP7 Projects Eurofleets and Eurofleets2 are a European wide alliance of marine research centers that aim to share their research vessels, to improve information sharing on planned, current and completed cruises, on details of ocean-going research vessels and specialized equipment, and to durably improve cost-effectiveness of cruises.

Within this context logging of information on how, when and where anything happens on board of the vessel is crucial information for data users in a later stage. This forms a primordial step in the process of data quality control as it could assist in the understanding of anomalies and unexpected trends recorded in the acquired datasets.

In this way the completeness of the metadata is improved as it is recorded accurately at the origin of the measurement. In the past, the collection of this crucial information was done in very different ways, using different procedures, formats and pieces of software in the context of the European Research Fleet. At the time that the Eurofleets project started, every institution and country had adopted different strategies and approaches, which complicated the task of users that need to log general purpose information and events on-board whenever they access a different platform losing the opportunity to produce this valuable metadata on-board. Among the many goals the Eurofleets project has, a very important task is the development of an "event log software" called EARS (Eurofleets Automatic Reporting System) that enables scientists and operators to record what happens during a survey. EARS will allow users to fill, in a standardized way, the gap existing at the moment in metadata description that only very seldom links data with its history.

Events generated automatically by acquisition instruments will also be handled, enhancing the granularity and precision of the event annotation. The adoption of a common procedure to log survey events and a common terminology to describe them is crucial to provide a friendly and successfully metadata on-board creation procedure for the whole the European Fleet. The possibility of automatically reporting metadata and general purpose data, will simplify the work of scientists and data managers with regards to data transmission.

An improved accuracy and completeness of metadata is expected when events are recorded near acquisition time. This will also enhance multiple usages of the data as it allows verification of the different requirements existing in different disciplines.

Universal database of Unified System of Information on the World Ocean

Alexander Kobelev, All Russian Institute for Hydrometeorological Information World Data Centre (Russia), kobelev@meteo.ru

Evgenii Viazilov, All Russian Institute for Hydrometeorological Information World Data Centre (Russia), vjaz@meteo.ru

To ensure the integrated provision of hydrometeorological information and information on marine activities to users, integration of distributed and heterogeneous data is needed. Integration is implemented with the help of the unified data description model based on 19115/19139 ISO standards, unified vocabulary of parameters, common codes and classifiers. For data held in various sources various types of physical storage are typical: factographic, object (images and documents), spatial, service. There are various types of data logical presentation: points, profiles, grids. The same attributes of data may be presented in various units of measurement and numerical systems.

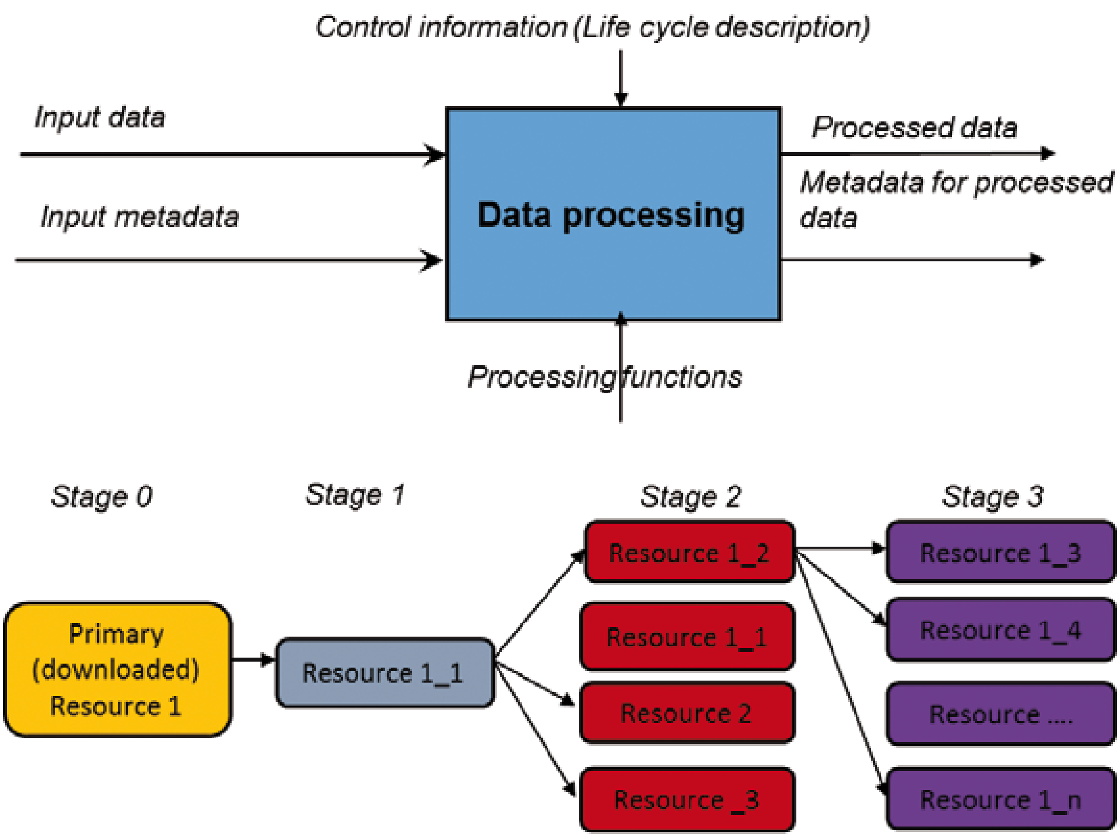


Fig. 1 - Pipelined data processing.

In addition, data management is required when data are downloaded, processed and used. For all of these processes data harmonization is necessary. Data harmonization is implemented with the help of the universal data base (UDB) developed as a result of data integration to allow data to be presented in a unified form to be further used for generation of information products (Fig. 1).

The UDB provides a unified access to data and metadata. A set of parameters of the environment may change with time. These changes should traced automatically and the UDB should be adapted in due time to ensure adequate data downloading. The UDB should have a data model, which makes it possible to deal with any data being integrated due to the flat data structure used for all types of data. The UDB should include a set of functions for preliminary data processing (library of processing functions) such as conversion to unified units of measurement, filtration (e.g. by specific parameters), calculation of derived characteristics, integration of data from various sources, data accumulation in time, data indexation, etc. These functions are included into the data processing algorithm and are reflected in the data life cycle. Implementation of the algorithm, fixed in the data life cycle, results in one or several derived tables of the data base. The UDB also manages the process of services updating.

Approaches and implementation of marine data harmonization in Russian Federation

Nikolay Mihaylov, All-Russian Research Institute of Hydrometeorological Information - World Data Center (Russian Federation), nodc@meteo.ru

Evgenii Vjazilov, All-Russian Research Institute of Hydrometeorological Information - World Data Center (Russian Federation), vjaz@meteo.ru

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

Alexander Voronsov, All-Russian Research Institute of Hydrometeorological Information - World Data Center (Russian Federation), vorv@meteo.ru

The marine data standardization and their integration and availability play a significant role for ocean research and practical marine activities. The data need to numerous scientific and design organizations, authorities and other maritime activities stakeholders. The data have a different levels of the processing (observations, forecasts, climatic summaries, analysis) and is haracterized by the large number of marine parameters (at least 800), essentially distributed geographically. It is required to provide the exchange and access to vast amounts of data that are placed in the distributed sources and have diverse types (databases, data files, GIS layers, geo-services and others), methods and formats of data presentation.

The data harmonization is a powerful mechanism for increasing the data accessibility taking into account the marked specificity and issues. Moreover, the concept of "data harmonization" rather complex and its precise (conventional) definition does not exist. But most often, a data harmonization means the possibility to combine data from heterogeneous sources into integrated and consistent information products, in a way that is of no concern to the end-user. When data are not harmonized, users have to spend a lot of time and other resources to search and conversion of data.

In Russia targeted data harmonization held in unified system of information on the world ocean (hereinafter, ESIMO). The system provides the information - communication infrastructure for integration of distributed and heterogeneous data supplied by multy-discipline marine systems, and acccess to the integrated data basing on "single window" principal.

Data harmonization solution is based on a series of components, such as a common parameters dictionary, a metadata and data model, exchange standards for data and services. These components are implemented in a Web-based environment.

Creation and support of the uniform dictionary of parameters of system allow leading all attributes used in information resources to uniform system of names.

The implementation of a data management plan to uplift historical data: long-term change detection in the Belgian Continental Shelf

Ruth Lagring, Royal Belgian Institute of Natural Sciences (Belgium), ruth.lagring@naturalsciences.be

Marielle Adam, Royal Belgian Institute of Natural Sciences (Belgium), madam@naturalsciences.be

Annelies Goffin, Vlaams Instituut voor de Zee (Belgium), annelies.goffin@vliz.be

Yvan Stojanov, Royal Belgian Institute of Natural Sciences (Belgium), ystojanov@naturalsciences.be

Lennert Tyberghein, Vlaams Instituut voor de Zee (Belgium), lennert.tyberghein@vliz.be

Thomas Vandenberghe, Royal Belgian Institute of Natural Sciences (Belgium), tvandenberghe@naturalsciences.be

The amount of marine data collected during research and monitoring programs is increasing continuously. Based on these data, long-term changes have been reported in all measured indicators defining an ecosystem status. Over the years not only the climate is changing, also the know-how has undergone an immense evolution. First of all, methods for data collection and analyses have improved, and secondly means for data handling and storage have been upgraded. As in the past a lot of data were stored locally, today more than ever there is a need for integrated quality-checked, intercalibrated and integrated datasets. The project "4 Decades of Belgian Marine MONitoring: uplifting historical data to today's needs" (4DEMON) has the challenge of integrating data on contamination, eutrophication and ocean acidification gathered in the past with recent data from the Belgian Continental Shelf (BCS). This both safeguards the historic data and uplifts its value as it can serve many new and cross-disciplinary research objectives.

To streamline the resuscitation process a data management plan was implemented. At the base is a workflow with all data management tasks, the latest processes and tools. This involves a secure online file sharing system (MDA), a Data Inventory and Tracking system (DITS), the import process in large integrated databases (IDOD and IMERS) and international dissemination via projects like OBIS and EMODNet by the National Oceanographic Data Centres (VLIZ and BMDC).

The intermediate results are very promising: thanks to the inventorisation of over 1000 datasources an immense amount of data could be resuscitated. As a result, the project could already indicate shifts in species, test standardization methods and relate trends to reduction policies. The adequate data management plan can be applied in many projects.

IOWMETA - A modular metadata information system enabling the extraction of ISO-compliant metadata from heterogeneous sources

Susanne Jürgensmann, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), susanne.juergensmann@io-warnemuende.de
Susanne Feistel, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), susanne.feistel@io-warnemuende.de
Steffen Bock, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), steffen.bock@io-warnemuende.de

The purpose of the metadata information system IOWMETA is to provide a comprehensive catalogue and a central infrastructure hub for all kinds of distributed research data stored at the Leibniz Institute of Baltic Sea Research Warnemünde (IOW). Metadata, describing heterogeneous data, can be standardized and thereby facilitate an international exchange of data and metadata. Comprehensive metadata and thorough documentation ensure long-term usability and maximum potential knowledge gain from already existing research datasets.

IOWMETA is based on the metadata standards of ISO 191xx and incorporates the open-source platform GeoNetwork, offering data access for scientists, federal agencies and the public. A set of more than 40 ISO-elements (including the mandatory ISO 19115 core) are put into practice as an ever-expanding relational database consisting of currently 100 tables. Complementary a software interface is in continuous development that reflects the underlying database schema in terms of its class structure and comprises basic import and export functionality. These two components build the technical core for subsequent modular extensions.

The interface facilitates the straightforward implementation of modules, mapping metadata from arbitrary data sources to ISO-compliant data structures and subsequently storing the results to IOWMETA. The following three examples serve as an illustration for the highly flexible extension capabilities of this modular approach.

The first module automatically extracts metadata from the oceanographic database of IOW, the so-called IOWDB. The IOWDB constitutes the most important long-term data source at IOW. The relational database contains not only research data but also associated metadata including information about scientists, platforms, methods and parameter descriptions. The scientific relevance of the data in IOWDB and its already well-structured content make it the ideal use case for automatic metadata extraction.

A maximum of descriptive information is extracted and transformed in order to meet the requirements of ISO 191xx. Additionally, the resulting metadata is enriched by a link back to the corresponding dataset stored in IOWDB and therefore provides a direct download option. This way, several ten thousands of datasets with the granularity of one dataset per CTD-cast are automatically described with standardized metadata.

Another module in development integrates metadata from a Zoobenthos database, containing taxonomy, abundances and biomasses. The extracted metadata i.e. referencing systems, data granularity and keywords differ from the previous example. However, the software interface allows an automatic integration of metadata into IOWMETA with little effort.

The third exemplary module uses as its data source a file system, rather than a relational database. A central data storage at IOW contains either raw, validated or processed data from scientific cruises. An XML file for indexing the content is placed in each folder of the file system, describing in a predetermined manner the data stored in the folder. This XML can automatically be retrieved and parsed. The module in IOWMETA converts the information from the XML to the ISO elements.

Based on the same interface a different kind of extension exports metadata from the IOWMETA database. The metadata are recursively extracted and transformed into an XML structure, compliant with ISO 19139. Furthermore, IOWMETA includes a version-control system, allowing the later reconstruction of previous changes in data and metadata.

IEO marine data discovery, representation and retrieval through metadata interoperability

Pablo Otero, Instituto Español de Oceanografía Madrid (Spain), pablo.otero@md.ieo.es
Elena Tel, Instituto Español de Oceanografía Madrid (Spain), elena.tel@md.ieo.es
Gonzalo Gonzalez-Nuevo, Instituto Español de Oceanografía Vigo (Spain), gonzalo.gonzalez@vi.ieo.es

Introduction

Access to marine data is vitally important for marine researchers, but also for a wide variety of professionals who use these data to tackle problems related to climate change, coastal engineering, fishing or aquaculture, among others. In addition, the demand of this kind of information by the general public is becoming more and more common (recreational navigation, nautical sports, tourism, etc.). Unfortunately, marine datasets are usually stored in specialized portals and they are often not indexed for internet search engines, and therefore will not appear in end-user search results. Moreover, finding them on these portals is usually complex for the non-specialized public. Data producers interested in targeting a wider and general public, should design user-friendly sites, with the use of the proper terminology, prioritizing end-user interests and letting their applications interact with the public through social network sites. However, from an organizational point of

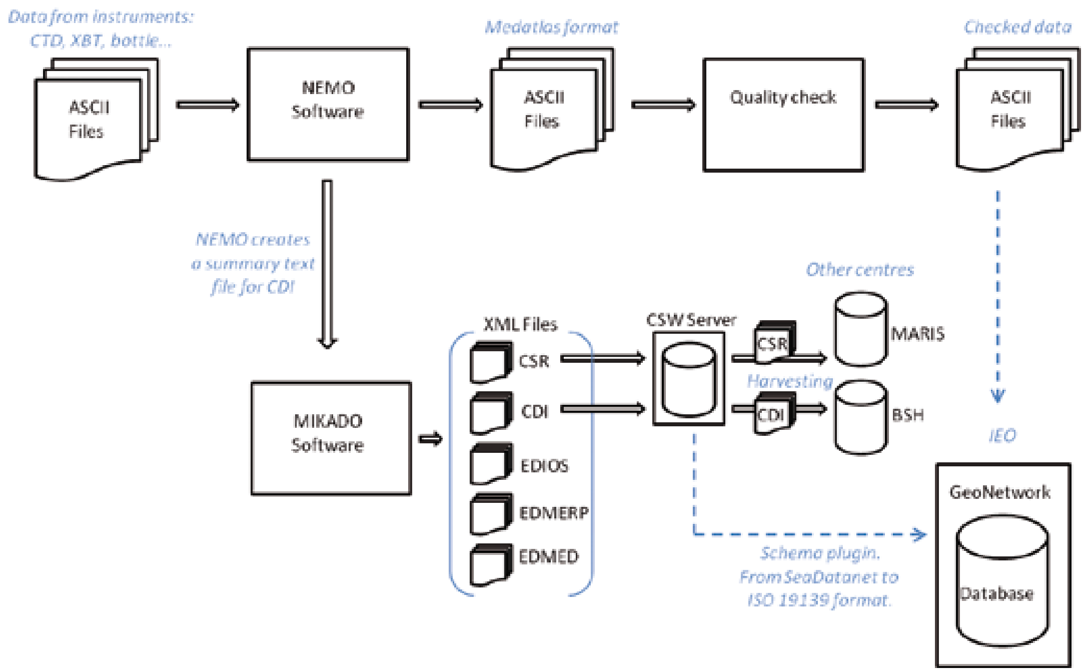


Fig. 1 - Schema of the marine data infrastructure and work-flow at the IEO.

view, achieving this goal implies to walk a long path where standardization and interoperability are main steps for managing the large and diverse data sets that are collected by the oceanographic surveys and the ocean observation systems.

In Spain, since 1964 the *Instituto Español de Oceanografía (IEO)* maintains the National Oceanographic Data Center (NODC), responsible for the compilation, storage and distribution of marine data. The integration of the institution into larger international frameworks includes the Global Ocean Observing System (GOOS) and its regional groups, International Oceanographic Data and Information Exchange (IOC/IODE), or European consortiums as SeaDataNet and EmodNET. They have contributed to move forward under the guidelines of standardization for geospatial data and information of projects, vessels and observatories. However, an additional effort must be done to accomplish the INSPIRE Directive (2007/2/EC) and to bring marine data closer to the Spanish end-user community.

Methods

Nowadays, marine data received at the IEO Datacenter are transcript in the auto-descriptive ASCII SDN-Medatlas format (1) by using the NEMO javatool (2). Data Quality Control procedures (3) are performed to detect missing mandatory information, format errors, duplicates and outliers. Following the agreed criteria, a quality flag is attached to each numerical value in order to preserve original data, provide them added-value and permit future re-validations. Each profile or dataset is accompanied of a XML-metadata file, which describes the dataset. These metadata are shared within the SeaDataNet infrastructure. To properly accomplish with ISO-19139 and satisfy the requirements under the Spanish transposition of INSPIRE Directive (2007/2/EC), a transformation of these metadata are performed. Transformed metadata are distributed through a customized GeoNetwork portal, focused on cataloguing data served by the IEO.

Conclusions

The aim of the IEO GeoNetwork Data Portal is to become in the starting place to find spatial national datasets for the Spanish community. The portal facilitates the discovery, multiple usage and dissemination of geospatial ocean-related information.

Acknowledges. In the last years, the IEO data center activity has been partially supported by SeaDataNet (FP6/2002-2006/026212), SeaDataNet-II (FP7/2007-2013/283607), EMODnet-Lot4-Chemistry (EU-DGMARE/2012/10)

References

SeaDataNet data file formats: ODV, MEDATLAS, NetCDF (deliverable D8.5) http://www.seadatanet.org/content/download/16251/106283/file/SDN2_D85_WP8_Datafile_formats.pdf.

SeaDataNet NEMO user manual. http://www.seadatanet.org/content/download/25784/176636 / file/sdn_Nemo_UserManual_V1.6.2.pdf.

SeaDataNet Quality Control Procedures v2.0. http://www.seadatanet.org/content/download /18414/119624/file/SeaDataNet_QC_procedures_V2_%28May_2010%29.pdf.

MEDIN – The Hub for UK Marine Data

Clare F Postlethwaite, Marine Environmental Data and Information Network (United Kingdom), cfpo@bodc.ac.uk

Hannah Williams, Marine Environmental Data and Information Network (United Kingdom),

Lesley Rickards, Marine Environmental Data and Information Network (United Kingdom), ljr@bodc.ac.uk

Gaynor Evans, Marine Environmental Data and Information Network (United Kingdom), gaev@bodc.ac.uk

Marine science and marine management decisions are underpinned by good quality data. Marine data are assets that can be reused many times to, for example: supplement new measurements; provide temporal or geographical context; monitor environmental change; or make decisions about where and how commercial or recreational activities should take place. Marine data are collected by a plethora of organisations for a range of reasons including academic research, environmental monitoring, defence and commercial activities. As long as any limitations with such data are fully understood, these data are potentially reusable. There are over 200 holders of marine data in the UK so it has traditionally been difficult for the marine community to make full use of the UK marine data assets. The Marine Environmental Data and Information Network (MEDIN) has been working with both public and private organisations to make it easier to share UK marine data and to ensure that data are well documented and securely archived. To do this, MEDIN has provided:

- A network of 7 accredited Data Archive Centres to ensure long term curation of a wide range of marine data.
- A single portal to search the data holdings of the Data Archive Centres, as well as those held my several other public and private sector organisations.
- A marine specific discovery metadata standard.
- A range of guidelines to identify the metadata needs of different data types.

MEDIN has significantly increased access to UK marine data by promoting best practice, influencing the release of terabytes of marine industry data, providing marine metadata standards as well as a single hub to search the data collected by hundreds of disparate organisations. However, there are still challenges ahead. In particular MEDIN are looking to make more data directly downloadable from the metadata - “2 clicks to data”; engage with more private sector partners; and improve the functionality of the MEDIN search portal.

A step toward data interoperability: NetCDF metadata comparative analysis in RITMARE Italian Project

Elena Partescano, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), epartescano@inogs.it
Alessandro Crise, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), acrise@inogs.it
Alessandra Giorgetti, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), agiorgetti@inogs.it
Lorenzo Corgnati, ISMAR Centro Nazionale delle Ricerche (Italy), lorenzo.corgnati@sp.ismar.cnr.it
Carlo Mantovani, ISMAR Centro Nazionale delle Ricerche (Italy), carlo.mantovani@cnr.it
Vega Forneris, ISAC Centro Nazionale delle Ricerche (Italy), v.forneris@isac.cnr.it
Cristina Tronconi, ISAC Centro Nazionale delle Ricerche (Italy), c.tronconi@isac.cnr.it
Massimiliano Drudi, Istituto Nazionale di Geofisica e Vulcanologia (Italy), massimiliano.drudi@ingv.it

Introduction

The Ritmare (La RIcerca ITaliana per il MARE) project 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) (from http://www.ritmare.it).

In detail, the Sub-Project 5 (SP5) aims to strengthen and credit the Italian observing system in Europe by integrating, as far as possible, various national scientific, industrial and institutional components (especially ISPRA and Civil Protection) and also creating infrastructure edge products, offering at the same time to the domestic industry the opportunity to create new products and expand into new markets. Additionally, the subproject defines requirements, methodologies and technologies that ensure data quality and significance, defined as the ability to represent and interpret ecosystem characteristics and changes, and to analyze its trends (from http://www.ritmare.it).

To guarantee free and open access to data and metadata acquired by the observation sites, SP5 challenge has been to develop a system integrating and giving access to data acquired by satellites, moorings, radars, models, gliders and floats, trying to identify a technological solution that guarantees the interoperability between data collected by heterogeneous platforms.

The data used in SP5 derive from different marine communities and the data format used is mainly NetCDF format, but the conventions used are different (mainly Ocenasites 1.2, CF-1.6 and CF-1.4).

This format guarantees, by integrating different standard protocols and interfaces such as OpenDAP, WMS, WCS, an easy access and sharing of array-oriented data.

As a first step, we collected differentThredds Servers used as end-points to share data: gliders, floats and moorings (http://insitu-ritmare.cineca.it/thredds/catalog/Ritmare/catalog.html) developed in collaboration with CINECA; satellites and HF and X band radar (http://ritmare.

artov.isac.cnr.it/thredds/catalog.html) made available by CNR-ISAC and CNR-ISMAR.

Then, we decided to investigate the different NetCDF formats and in particular to identify the common metadata used, to define a common set of metadata to guarantee an efficient data discovery.

In the case of radar data, the activities carried on in the RITMARE project led to the definition of data and metadata interoperable formats and QA/QC procedures to be adopted as European standards for coastal radar applications, thanks to the participation in the EuroGOOS HFR Task Team.

Comparative analysis

The first step was to collect information about each type of NetCDF (radar, satellite, glider, float, mooring and model). In particular, for every kind of NetCDF we collected information about: identity of the collating center, the data provider, the data type, the NetCDF version and convention, the project involved, the metadata standards, the list of standard services activated, the endpoint, availability, comment and suggestions. The second step was to analyse the set of metadata used, by all the communities taking in consideration diverse convention, mainly Ocenasites 1.2, CF-1.6 and CF-1.4.

As a result, we proposed 11 metadata as mandatory and common to all NetCDFs, answering questions such as what, where, when, how and who acquired data, to permit an efficient data discovery.

The last part of this work was to compare the main metadata catalogs used by SeaDataNet and EMODNet European marine community, the CDI (Common Data Index) with the metadata included in the NetCDFs used in Operative Oceanography (Coriolis, EuroGOOS and Argo). From this comparison we identified 10 common metadata, 6 fields used into CDI but absolutely absent in all kinds of NetCDF and 7 fields used in CDI and discontinuously in NetCDFs.

Another relevant element is that CDIs uses SeaDataNet convention (BODC vocabularies and EDMO catalogs), while the NetCDFs use Ocenasites 1.2, CF-1.6 and CF-1.4 conventions. This leads to the need of a format converter to compare the information coming from different metadata catalogs.

Conclusion

The aim of this work was to identify a technological solution that permits a data discovery using common metadata, present in data acquired by different platforms using different data conventions (Ocenasites 1.2, CF-1.6 and CF-1.4). The first step was to collect different end-points to share data: (http://insitu-ritmare.cineca.it/thredds/catalog/Ritmare/catalog.html) developed in collaboration with CINECA and (http://ritmare.artov.isac.cnr.it/thredds/catalog.html) developed by CNR-ISAC and CNR-ISMAR.

The second step was to analyze the metadata used in different marine conventions and identify a common set of metadata. After this, we compared the NetCDF metadata used in Ritmare project with the major marine metadata catalogs such as CDI.

The future step will be to find a technological solution, such as Geonetwork, that permits to link to THREDDS Data Server allowing an efficient data discovery directly on NetCDF files.

Acknowledgment. This research was supported by RITMARE project. We thank our colleagues from CNR-IREA, who provided expertise and assisted the research, and from CINECA for their technical support.

The CoCoNet solution for management and access heterogeneous marine datasets and metadata

Valentina Grande, ISMAR Centro Nazionale delle Ricerche (Italy), valentina.grande@bo.ismar.cnr.it
Artem Kruglov, Ukranian Scientific Center of Ecology of the Sea (Ukraine), antilles@i.ua
Federica Foglini, ISMAR Centro Nazionale delle Ricerche (Italy), federica.foglini@bo.ismar.cnr.it
Oleksandr Leposhkin, Ukranian Scientific Center of Ecology of the Sea (Ukraine), thejoker_2@mail.ru
Richard Lisovskyi, Ukranian Scientific Center of Ecology of the Sea (Ukraine), ricgardl@te.net.ua
Oleksandr Neprokin, Ukranian Scientific Center of Ecology of the Sea (Ukraine), o.neprokin@gmail.com

The CoCoNet project

The aim of the project was identify groups of putatively interconnected Marine Protected Areas (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 enhances 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. The coastal focus was widened to offshore and deep sea habitats, comprising them in MPAs Networks. These activities also individuated areas where Offshore Wind Farms (OWFs) might become established, avoiding too sensitive habitats but acting as stepping stones through MPAs. Two pilot project areas (one in the Mediterranean Sea and one in the Black Sea) tested in the field the assumptions of theoretical approaches. The project covered a high number of countries and involved researchers covering a vast array of subjects, developing a timely holistic approach and integrating the Mediterranean and Black Seas scientific communities.

The Work Package 9 named Data Management and Synthesis provided a common framework for data management and final synthesis of the outcomes.

The database

In the lifespan of the project the WP9 collected a big amount of data from different sources and disciplines, from biology to geology, oceanography, socioeconomics (real world). The data are organized and stored in 11 ESRI File geodatabases with ArcGIS (physical model), implemented as UML diagram with Microsoft Visio, (logical model), starting from the INSPIRE Directive (conceptual model). The 11 themes are: PreotectedSites, OffshoreWindFarms, HabitatsAndBiotopes, Biodiversity, Threats, Socioeconomics, Geology, Oceanography, MaritimeUnits, SeaRegions, Elevation. Other 3 File geodatabases have been developed for the pilot project areas: HabitatMapping, Connectivity and a more detailed version of the geodatabase Oceaography.

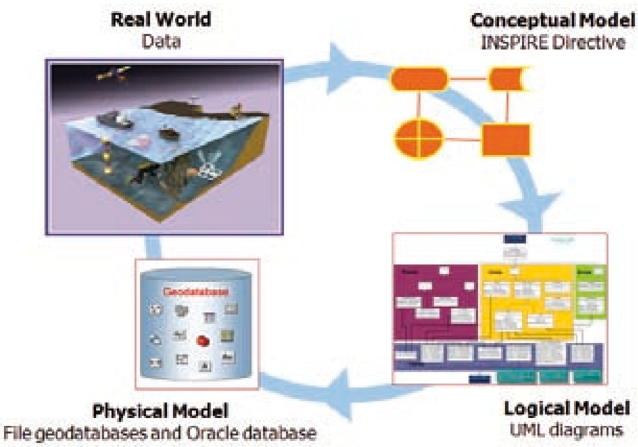


Fig. 1 - CoCoNet workflow, from the real word to a physical model.

The 14 File geodatabases converge in a unique Oracle database and are accessible thought a WebGIS portal.

The WebGIS platform

We implemented two CoCoNet WebGIS platforms: 1) at basin scale for the Mediterranean and Black Seas ; 2) at mesoscale for the pilot project areas. The platforms have been developed though the software ArcGIS Server 10 and the Moka CMS. The Moka CMS GIS (Content Management System GIS) is the core of the GIS infrastructure. The CMS is a tool for creating GIS application using cartographic object organized in a catalogue. The CoCoNet WebGIS platforms are free accessible from a webpage at the link: http://coconetgis.ismar.cnr.it/.

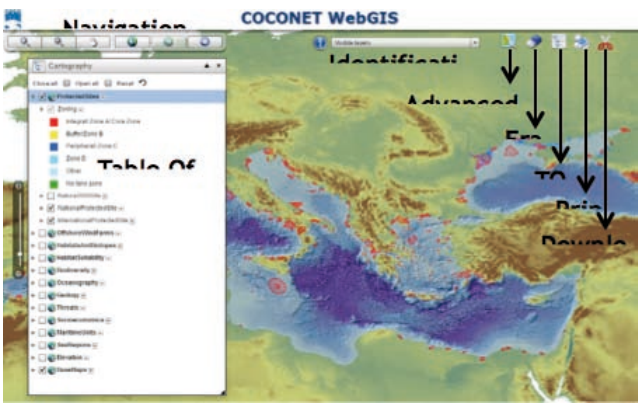


Fig. 2 - CoCoNet WebGIS platform at basin scale and its functions.

The platforms allow to browse the data with the navigation tools and the TOC. The user can identify the single object in the map with the identification tools, query the contents in the tables with the advanced search and print a map. At least, the user can download the layers according to the data policy of the project or use the 14 geodatabase as map service and feature service (WMS and WFS OGC services) on his desktop.

The Metadata portal

Each layer in the WebGIS platforms is described by a metadata file. We created the metadata with Mikado software according to the SeaDataNet vocabularies and published them thought the SeaDataNet portal (http://www.seadatanet.org/Metadata).

A metadata file for each layer, created with ArcCatalog according to the ISO 19139, is stored in the CoCoNet metadata portal (http://gp.sea.gov.ua:8082/geoportal/catalog/main/home.page). Here, more detailed information about the datasets provided by each partner is available in specific metadata files.

DASSH: where research and data meet

Matt Arnold, The Marine Biological Association (United Kingdom), matarn@mba.ac.uk
Charly Griffiths, The Marine Biological Association (United Kingdom), chgr@mba.ac.uk
Mike Hall, The Marine Biological Association (United Kingdom), mikhal@mba.ac.uk
Esther Hughes, The Marine Biological Association (United Kingdom), esthug@mba.ac.uk
Dan Lear, The Marine Biological Association (United Kingdom), dble@mba.ac.uk

Looking Back

Based at the historic Marine Biological Association (MBA) in Plymouth, the Data Archive for Marine Species and Habitats (DASSH) sits at the hub of marine science in the UK. In 2014 the MBA celebrated its 130th year, and in such an old organisation there are inevitably some very old datasets, the oldest dating back to 1813.

Fig. 1 - The MBA c. 1890 (MBA archive).

Originally these datasets came from a variety of sources. Some from academic research, but many from individual naturalists such as George Montagu who described many species in the Salcombe area of Devon in the late 18th and early 19th century.

The Here and Now

Today, the data held at the MBA derive from a wider range of sources; from amateur marine biologists, thematic recording schemes, statutory monitoring, academic research, and industry and commercial organisations; they range in scale from individual species records to large interdisciplinary studies.

The role of DASSH is to resolve and publish these disparate datasets for marine science professionals to find, access and utilise. Ultimately improving our understanding of the ocean and its inhabitants, and to make planning and management decisions based on the best-available scientific evidence.

DASSH publishes data at a wide range of geographic scales. At the national level DASSH provides data to the UK National Biodiversity Network (NBN) and forms the biological node of the Marine Environmental Data and Information Network (MEDIN); also assisting their work in data standards. At the European and International level DASSH data feed into portals such as EMODnet Biology, EurOBIS and GBIF, and participates in EMBRC and ENVRI+.

Moving Forward

DASSH is currently working on a number of tools to automate data flow, discovery and reuse. The “DASSH autoDOI” tool will automatically assign Digital Object Identifiers (DOIs) to datasets flowing through DASSH, while the new Drupal-based metadata editor for the MEDIN network will improve the user experience, and ensure high-quality metadata can be produced for the UK marine data community in a timely manner. Through our work on ENVRI+ we are involved in creating standards for non-spatial marine data, understanding the flow of data from peer-reviewed journals to EU data infrastructures, and the semantic interoperability of marine biodiversity data.

The Talk

This talk will be split into two parts. Part one will focus on how DASSH processes datasets from highly heterogeneous sources and disseminates these data to a number of national and European data outlets. Part two will focus on the tools currently in development at DASSH to assist with data management, namely the new MEDIN metadata editor and autoDOI tool.

Initiating the Global Data Assembly Centres for Marine Biogeochemistry

Benjamin Pfeil, Bjerknes Climate Data Centre, University of Bergen and Bjerknes Centre for Climate Research (Norway), Benjamin.Pfeil@uib.no

Maciej Telszewski, International Ocean Carbon Coordination Project of the Scientific Committee on Oceanic Research (SCOR), and Intergovernmental Oceanographic Commission (IOC) of UNESCO, Institute of Oceanology of Polish Academy of Sciences, m.telszewski@ioccp.org

Kevin O'Brien, University of Washington/JISAO (United States of America), kob@uw.edu

Toste Tanhua, Geomar, Helmholtz Centre for Ocean Research (Germany), ttanhua@geomar.de

Are Olsen, University of Bergen and Bjerknes Centre for Climate Research (Norway), are.olsen@uib.no

Eugene Burger, National Oceanic and Atmospheric Administration (United States of America), Eugene.Burger@noaa.gov

Alex Kozyr, Carbon Dioxide Information Analysis Center, Environmental Sciences Division Oak Ridge National Laboratory (United States of America), kozyra@ornl.gov

For the past decades the international ocean biogeochemistry community has mainly used and depended upon one global data center, the Carbon Dioxide Information Analysis Center ocean trace gases section (CDIAC-Oceans) at the U.S. Department of Energy's Oak Ridge National Laboratory, USA. CDIAC-Oceans provides data management support for ocean carbon measurements from Repeat Section cruises, VOS/SOOP lines, time series and moorings data, has accommodated most community requests for data archival and data access and has also actively engaged with the science community, supporting large synthesis projects like SOCAT, the LDEO Database, GLODAP, CARINA, PACIFICA and GLODAPv2. The funding support for the ocean trace gases section of CDIAC has been endangered several times in the past and puts in jeopardy the uninterrupted data management that the ocean biogeochemical data community has come to rely upon as well as the trust and recognition from the scientific community that CDIAC-Oceans has built though decades of interactions. The loss of CDIAC-Oceans will likely have a negative impact on ocean carbon data submissions and reduction in value added products.

The uncertainty of funding for CDIAC highlights the vulnerability of a system that relies too heavily on individual data managers or institutions. At the same time, it provides an opportunity to review the requirements for modern data access and data management systems that have evolved significantly during the last decades and which currently are not being met through the CDIAC infrastructure. Operational data management systems that (a) provide automated data ingestion, (b) conform to modern standards for data and metadata, (c) utilize standard vocabularies, (d) have easy-to-use data access tools, and (e) provide stable data citations are driven not only by user requirements, but also by funding and government agencies as they promote open access to data. In the discussion of CDIAC funding and the vulnerability of ocean biogeochemistry data, we see a strong opportunity to implement a data management infrastructure that can thrive in the modern world of integrated science data.

A modern data management infrastructure needs to be established in which existing data centers (e.g. CDIAC, CCHDO, BCO-DMO, PANGAEA) and data from various other networks (e.g. OceanSites, Argo) can be integrated through interoperable discovery and access services. This is essential for providing access to data, while at the same time ensuring that credit for data creation and data synthesis products is appropriately assigned. We propose to mimic the successful data management approach implemented for the Argo profiling float network (http://www.argodatamgt.org). The Argo network addresses national funding agency requirements of having data housed in specific locales by setting up two Global Data Assembly Centers (GDACs), one in the US and one in Europe. Data holdings are mirrored between the data centers and can be accessed through either one. This redundancy makes access to the data collection, by nature, more resilient.

We suggest establishing a system of Global Data Assembly Centers for ocean biogeochemistry (e.g. GDAC-OBGC) where two initial GDACs are established, each with specific roles and responsibilities. The two GDACS will be complementary systems that will leverage the unique capabilities of each, to provide a complete solution for data ingestion, data quality control, data access, data citation and data archival.

A strong focus will be on interoperable access of standards compliant carbonate system data and metadata, irrespective of where they are archived. In addition, it is paramount that support for automated data ingestion, both for real time and delayed mode sources, be integrated into the data management workflow. This is crucial to being able to keep pace with the higher volume of data now being generated by autonomous platforms. First order quality control checks built into the automated ingestion streams can further reduce the quality control burden. By providing interoperable access, and adhering to standards and conventions, this framework will make future data synthesis products and activities much more efficient than with the current non-centralized data management system.

Another important emphasis of the GDAC will be an external review process by ensuring that (a) data are being quality assured and controlled according to community agreed standards, (b) direct feedback is given to the data source, (c) duplicates are being identified and resulting issues are resolved, (d) metadata are complete according to community agreed best practices or existing standards, (e) data and metadata are available through interoperable services, (f) reports are made to IODE and JCOMM Committees on data management status and activities, (g) data citation practices as outlined by the Research Data Alliance (RDA) and DataCite are incorporated, (h) data requests and searches from users can be reproduced and (i) there is clear tracking of the complete data lifecycle for each dataset. The last three bullet points are often overlooked but are increasingly becoming more important to ensure that PIs get credit for data they create and that users/reviewers can reproduce the exact data requests for data that is referenced in scientific publications.

The implementation of the above framework will facilitate continuation of the data synthesis and assessment products such as GLODAP, SOCAT and create a foundation for additional data products, including the integration of data such as time series data and coastal data. In addition, the implementation of such a framework will support compatible efforts internationally, providing a cohesive process toward more uniform data management strategies within the ocean biogeochemistry community. In the long term, such efforts will provide a significant cost savings by reducing data management overhead as well reducing the data management burden on individual scientists.

ENVRIplus: creating a coherent, interdisciplinary and interoperable cluster of Environmental Research Infrastructures across Europe

Helen Glaves, British Geological Survey (United Kingdom) hmg@bgs.ac.uk

The ENVRIplus project aims to provide common solutions to the shared challenges encountered by European Environmental and Earth System Research Infrastructures (RIs) in their efforts to deliver new services for science and society. To reach this overall goal, ENVRIplus brings together the RIs, related projects and networks together with specialist technical partners to build common solutions for the shared problems encountered in the construction and implementation of the European Research Infrastructures (RIs).

Fig. 1 - European Research Infrastructures participating in ENVRIplus.

Collaboration within the ENVRIplus project and also with other related stakeholders will enable multidisciplinary Earth system science across the traditional scientific fields, which is vital in order to address today's global challenges. This cooperation will avoid the fragmentation and duplication of effort, making the products of the individual Research Infrastructures easier to use in combination whilst also improving their innovation potential and the cost/benefit ratio of the Research Infrastructure operations.

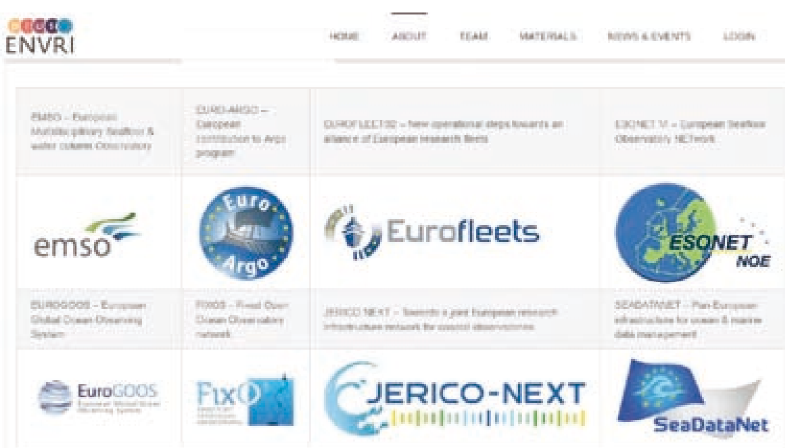


Fig. 2 - Marine RIs contributing to ENVRIplus.

ENVRIplus is driven by 3 overarching goals: 1) promoting cross-fertilization between Research Infrastructures; 2) implementing innovative concepts and technologies across RIs, and 3) facilitating research and innovation in the field of environment for an increasing number of users outside the Research Infrastructures.

All of the Earth system science domains (Atmosphere, Marine, Biosphere and Solid Earth) represented in the ENVRIplus consortium (Fig. 1) are working work together in order to capitalize on the progress made in the individual disciplines and strengthen intero-perability amongst RIs and domains. Many of the key European marine RIs are participating in ENVRIplus including SeaDataNet, Eurofleets, EuroArgo etc. (Fig. 2) will directly benefit from this coordinated approach to developing common solutions for shared issues within and between the RIs.

ENVRIplus is funded by the European Commission through the H2020 programme (grant number 654182). The project began 1st May 2015 and will run for 48 months.

Project 56: Data Management

Taco de Bruin, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Taco.de.Bruin@nioz.nl

Leandro Ponsoni, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Leandro.Ponsoni@nioz.nl

Jennifer Welsh, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Jennifer.Welsh@nioz.nl

Janine Nauw, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Janine.Nauw@nioz.nl

Wil Sistermans, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Wil.Sistermans@nioz.nl

Ronald de Koster, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Ronald.de.Koster@nioz.nl

Jan Nieuwenhuis, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Jan.Nieuwenhuis@nioz.nl

Thomas Richter, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Thomas.Richter@nioz.nl

Marten Tacoma, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Marten.Tacoma@nioz.nl

Richard Brink, NIOZ Royal Netherlands Institute for Sea Research (The Netherlands), Richard.Brink@nioz.nl

The National Ocean and Coastal Research Programme (in Dutch: ZKO) was a NWO Earth and Life Sciences (ALW) funded initiative focusing on the scientific analysis of five societal challenges related to a sustainable use of the sea and coastal zones. These five main challenges are safety, economic yield, nature, spatial planning & development and water quality. According to the NWO website, the ZKO program was ‘set up to strengthen the cohesion and collaboration within Dutch marine research’. The programme encouraged ‘research into current issues and policy-related topics important for the sustainable management of the Dutch coast’.

The ZKO programme was divided into 4 different themes or regions, with the first theme subdivided into three parts:

- Carrying capacity (Wadden Sea area)
  - Carrying Capacity: Line 1 - Policy-relevant Research
  - Carrying Capacity: Line 2 - Monitoring
  - Carrying Capacity: Line 3 - Hypothesis-driven Research
- Oceans
- North Sea
- Transnational Wadden Sea research

A total of 56 projects was funded during 2007-2015, ranging from research on governance of the Wadden Sea to research on trace elements and isotopes in the Atlantic Ocean.

Project 56 was the data management project for the entire ZKO programme. Its aim was to archive all relevant data from ZKO funded projects and to make the results, data and publications, publicly available, following the ZKO Data Policy. The data management project was carried out by the NIOZ Data Management Group.

The emphasis of this project was on providing an overview and quick access to the results of all ZKO research projects. This was done by developing a dedicated data portal, in which the metadata were based on SeaDataNet vocabularies throughout. This enables an easy connection to the SeaDataNet infrastructure in the future.

The presentation will introduce the various research projects of the ZKO programme and its many, many different datatypes. It will explain how the data portal was organized and is embedded in national and international data access infrastructures. Finally, the presentation will address the level of cooperation by the project scientists in submitting the data and the role of both the ZKO Data Policy as well as the funding agency in this matter.

Raising the bar on data recovery and discovery for the Southern Ocean Observing System

Roger Proctor, Integrated Marine Observing System (Australia), roger.proctor@utas.edu.au
Pip Bricher, Southern Ocean Observing System (Australia), data@soos.aq
Joana Beja, British Oceanographic Data Centre (United Kingdom), joja@bodc.ac.uk
Alex Kozyr, Carbon Dioxide Information Analysis Centre (United States of America), kozyra@ornl.gov
Anton Van de Putte, Royal Belgian Institute for Natural Sciences (Belgium), antonartica@gmail.com
James Cusick, Australian Antarctic Division (Australia), james.cusick@aad.gov.au
Steve Diggs, Scripps Institution of Oceanography (United States of America), sdiggs@ucsd.edu

The Southern Ocean is the most undersampled and important ocean on the planet, linking the major ocean basins and influencing the global climate. The first steps in developing a Southern Ocean Observing System are to bring together disparate multi-disciplinary datasets to improve discoverability of legacy datasets and hence facilitate their reuse, and to lay the foundation for efficient management of data from future Southern Ocean studies. Our initial focus is on assembling the registry of mooring sites and associated datasets and developing a web-based data discovery tool. This tool will utilise the Global Change Master Directory catalogue developed and maintained by NASA. Initial findings indicate a long tail of data from large research projects which has little visibility and are at risk of being lost forever. We will indicate the scale of the problem and outline our approach to bringing these data the surface.

The European Marine Observing Network and the development of an Integrated European Ocean Observing System. An EuroGOOS Perspective

Patrick Gorringe, EuroGOOS AISBL (Belgium), patrick.gorringe@eurogoos.eu
Vicente Fernández, EuroGOOS AISBL (Belgium), Vicente.fernandez@eurogoos.eu
Glenn Nolan, EuroGOOS AISBL (Belgium), glenn.nolan@eurogoos.eu

The ocean benefits many sectors of society, being the biggest reservoir of heat, water, carbon and oxygen and playing a fundamental role regulating the earth's climate. We rely on the oceans for food, transport, energy and recreation. Therefore, a sustained marine observation network is crucial to further our understanding of the oceanic environment and to supply scientific data to meet society's need.

Marine data and observations in Europe, collected primarily by state governmental agencies, is offered via five Regional Operational Oceanographic Systems (ROOS) within the context of EuroGOOS (http://www.eurogos.eu), an International Non-Profit Association of national governmental agencies and research organizations (40 members from 19 member states) committed to European-scale operational oceanography within the context of the Intergovernmental Global Ocean Observing System (GOOS). Strong cooperation within these regions, enabling the involvement of additional partners and countries, forms the basis of EuroGOOS work. Ocean data collected from different type of sensors (e.g. moored buoys, tide gauges, Ferrybox systems, High Frequency radars, gliders and profiling floats) is accessible to scientist and other end users through data portals and initiatives such as the European Marine Observations and Data Network (EMODnet) (www.emodnet.eu), SeaDataNet (http://www.seadatanet.org/) and the Copernicus Marine Service Copernicus (www.copernicus.eu).

Although a relatively mature European ocean observing capability already exists and its well-coordinated at European level, some gaps have been identified, for example the demand for ecosystem products and services, or the case that biogeochemical observations are still relatively sparse particularly in coastal and shelf seas. Assessing gaps based on the capacity of the observing system to answer key societal challenges e.g. site suitability for aquaculture and ocean energy, oil spill response and contextual oceanographic products for fisheries and ecosystems is still required. In this respect, an important effort is being carried out at European level aiming to stablish and consolidate an Integrated and sustained European Ocean Observing System (EOOS).

SESSION 3

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

Copernicus Marine Environment Monitoring Service In Situ TAC: an In situ operational data provision system for operational oceanography

Loïc Petit de la Villéon, Institut français de recherche pour l'exploitation de la mer (France), petit@ifremer.fr

Sylvie Pouliquen, Institut français de recherche pour l'exploitation de la mer (France), Sylvie.Pouliquen@ifremer.fr

Henning Wehde, Institute of Marine Research (Norway), henningw@imr.no

Joaquin Tintore, Balearic Islands Coastal Observing and Forecasting System (Spain), jtintore@socib.es

Thierry Carval, Institut français de recherche pour l'exploitation de la mer (France), Thierry.Carval@ifremer.fr

Lid Sjur Ringheim, Institute of Marine Research (Norway), sjur.ringheim.lid@imr.no

Susanne Tamm, Bundesamt für Seeschiffahrt und Hydrographie (Germany), susanne.tamm@bsh.de

Stephane Tarot, Institut français de recherche pour l'exploitation de la mer (France), Stephane.Tarot@ifremer.fr

Veselka Marinova, Institute of Oceanology Bulgarian Academy of Science (Bulgaria), marinova@io-bas.bg

Leonidas Perivoliotis, Hellenic Centre for Marine Research (Greece), lperiv@hcmr.gr

Marta de Alfonso Alonso-Muñoyerro, Puertos del Estado (Spain), mar@puertos.es

Thomas Hammarklint, Sveriges Meteorologiska och Hydrologiska Institut (Sweden), thomas.hammarklint@smhi.se

Fernando Manzano Muñoz, Puertos del Estado (Spain), fmanzano@puertos.es

Charles Troupin, Balearic Islands Coastal Observing and Forecasting System (Spain), ctroupin@socib.es

Kevin Balem, Institut français de recherche pour l'exploitation de la mer (France), Kevin.Balem@ifremer.fr

Corentin Guyot, Institut français de recherche pour l'exploitation de la mer (France), Corentin.Guyot@ifremer.fr

The CMEMS (Copernicus Marine Environment Monitoring Service) context

Copernicus, previously known as GMES (Global Monitoring for Environment and Security), is the European Programme for the establishment of a European capacity for Earth Observation and Monitoring. Copernicus includes 3 components: space, in-situ and services.

The services component encompasses a Marine Monitoring service among others (Emergency Management, Atmosphere Monitoring, Security and Climate Change). The focal point of the presentation will be laid on the InSitu Thematic Assembly Centre (InSituTAC) as part of the Services component only.

The In Situ component of the Copernicus marine service

The In Situ TAC - Thematic Assembling Centre- is a consistent and distributed provision system of ocean In Situ observations for MFCs –Marine Forecasting Centres. It manages real-time and delayed mode data.

The In Situ TAC is a distributed system organized in 7 region Distribution Units (DUs) see Fig.1.

The operation of the regional DUs is based on the regional EuroGOOS ROOS (Regional Oceanographic Operational System, more on http://eurogoos.eu). The Global component is an integrated service on JCOMM networks (Argo, Drifters, OceanSites, Gliders...) that complement the European observing systems in the global Ocean.

The focused ocean variables are (a): temperature, salinity (CMEMS V1), surface currents (CMEMS V2), (b): wave data (CMEMS V3) and (c): biogeochemical data (CMEMS V4). Focal point on the evolvement of the InSituTAC data provision lays on (a) for the first two versions (2015/2016), on (b) for the third version (2017) and c. for the fourth version (2018). Ancillary data are also distributed when available (such as sea level, met-ocean parameters), with no additional quality control made and in Situ TAC level.

Key features

The commitments for Copernicus In Situ TAC are to:

- Provide real time In Situ data in an operational mode within a few hours from acquisition by the observing system operators;
- Provide regularly delayed mode datasets (REP Products –Reprocessed products-) that have been passed through high level assessment procedures to detect data that are not coherent with the others);
- Provide all the above datasets in a fully consistent manner (NetCDF OceanSITES version 1.2, format with homegeneous and documented quality control.). from central and regional portals;
- Monitor the operational service by providing Key Performance Indicators(KPI° including continuously monitoring of the availability of distriution services to users
- Monitor the quality of the products distributed and identify when needed cross cutting activities to improve the products;
- Properly document the products made available in Product Users manuals (PUMS) and Quality Information Documents (QUIDs);
- Report on operations status on a regular basis (monthly operation reports, quaterly and yearly reports);
- Contribute to cross-cutting activities led by Mercator-Ocean on communication and traning, Ocean State reporting , harmonization across the CMEMS services.

The presentation will detail the In Situ TAC organization and the data integration system that has been developed to provide efficient data delivery services and the In Situ TAC data schema between the different in situ TAC actors as well as the collaborations with EMODNet-Physics, SeaDataNet network and EuroGOOS ROOS.

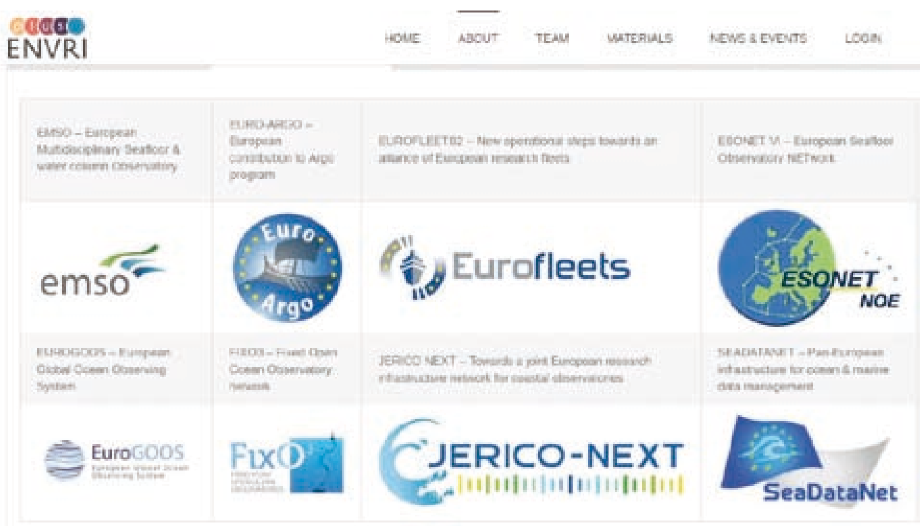
The presentation will also detail the technical solutions that have been used to address the key features described above.

A common European database for underway data from FerryBoxes

Gisbert Breitbach, Helmholtz Zentrum Geesthacht (Germany), gisbert.breitbach@hzg.de
Willi Petersen, Helmholtz Zentrum Geesthacht (Germany), wilhelm.petersen@hzg.de
Susanne Reinke, Helmholtz Zentrum Geesthacht (Germany), susanne.reinke@hzg.de
Patrick Gorringe, EUROGOOS (Belgium), patrick.gorringe@eurogoos.eu
Antonio Novellino, ETT People and Technology (Italy), antonio.novellino@ettsolutions.com

Since the beginning of FerryBox activities in the Baltic Sea in the '80s and on an European level with the FP5 project "FerryBox" (2003-2005) there has been a continous increasing of FerryBox lines as well as ongoing activities on harmonization of operation. However, progress in visibility and accessibility of FB data has been lacking even if there exists long-term FB data sets in the Baltic and North Sea that would be of greatest interest for scientists and the general user of marine data and information.

FerryBoxes on ships operating on fixed routes are a valuable tool for monitoring purposes. Within the European JERICO-NEXT project the idea arose to develop an European database to gather data from all available FBs and to increase the visibility and accessibility. At the Institute of Coastal Research at HZG such a database had already been established within the coastal observatory COSYNA (www.cosyna.de) which fullfills most of the requirements. This database will be expanded to enable all available European FB routes to use this database. The talk will present the data model of the database and the possibilities for visualising the data. The concept of importing FB data will be shown as well as first results from different European regions. The connection of this database with the European marine data management activities of EMODnet and CMEMS and links with the EuroGOOS ROOSs will also be demonstrated.



When diving animals help us to observe the oceans: the MEOP data portal

Fabien Roquet, Stockholm University (Sweden), roquet@misu.su.se
Lars Boehme, SMRU St-Andrews University (United Kingdom), lb284@st-andrews.ac.uk
Marthan Bester, University of Pretoria (South Africa), mnbester@zoology.up.ac.za
Horst Bornemann, Alfred Wegener Institute (Germany), horst.bornemann@awi.de
Sophie Brasseur, Wageningen University and Research (Netherlands), sophie.brasseur@wur.nl
Jean-Benoit Charrassin, Le Laboratoire LOCEAN, Sorbonne University (France), jbclod@locean-ipsl.upmc.fr
Dan Costa, University of California, Santa Cruz (United States of America), costa@ucsc.edu
Mike Fedak, SMRU St-Andrews University (United Kingdom), maf3@st-andrews.ac.uk
Christophe Guinet, CEBC, Université de La Rochelle/CNRS (France), christophe.guinet@cebc.cnrs.fr
Ailsa Hall, SMRU, St-Andrews University (United Kingdom), ajh7@st-andrews.ac.uk
Robert Harcourt, Macquarie University (Australia), robert.harcourt@mq.edu.au
Mark Hindell, IMAS, University of Tasmania (Australia), mark.hindell@utas.edu.au
Kit M. Kovacs, Norwegian Polar Institute (Norway), kit.kovacs@npolar.no
Christian Lydersen, Norwegian Polar Institute (Norway), christian.lydersen@npolar.no
Clive McMahon, IMAS, University of Tasmania (Australia), clive.mcmahon@utas.edu.au
Baptiste Picard, CEBC, Université de La Rochelle/CNRS (France), picard.baptiste@gmail.com
Gilles Reverdin, LOCEAN, Sorbonne University (France), gilles.reverdin@locean-ipsl.upmc.fr
Cécile Vincent, CEBC, Université de La Rochelle/CNRS (France), cvincent@univ-lr.fr

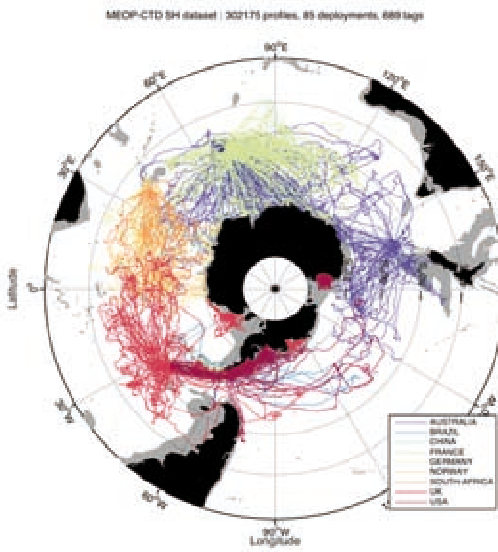
Seals help gather information on some of the harshest environments on the planet, through the use of miniaturized ocean sensors glued on their fur. The resulting data – gathered from remote, icy seas over the last decade – are now freely available to scientists around the world from the data portal http://www.meop.net.

The Polar oceans are changing rapidly as a result of global warming. Ice caps in Antarctica and Greenland are melting, releasing large quantities of freshwater into surface waters. The winter sea ice cover is receding in the Arctic and in large areas of the Southern Ocean, which promotes further warming. Southern winds are intensifying for reasons that are not fully understood. To understand the changing marine environment, it is necessary to have a comprehensive network of oceanographic measurements. Yet, until recently, the harsh climate and remoteness of these areas make them extremely difficult to observe. Diving marine animals equipped with sensors are now increasingly filling in the gaps.

When diving animals help us to observe the oceans

Since 2004, hundreds of diving marine animals, mainly Antarctic and Arctic seals, were fitted with a new generation of Argos tags developed by the Sea Mammal Research Unit of the University of St. Andrews in Scotland (Fig. 1). These tags can be used to investigate simultaneously the at-sea ecology (displacement, behaviour, dives, foraging success...) of these animals while collecting valuable oceanographic data (Boehme et al. 2009). Some of these species are travelling thousands of kilometres continuously diving to great depths (590 ± 200 m, with maxima around 2000m). The overall objective of most marine animal studies is to assess how their foraging behavior responds to oceanographic changes and how it affects their ability to aquire the resources they need to survive. But in the last decade, these animals have become an essential source of temperature and salinity profiles, especially for the polar oceans. For example, elephant seals and Weddell seals have contributed 98 % of the existing temperature and salinity profiles within the Southern Ocean pack ice. The sensors are non-invasive (attached to the animal's fur, they naturally fall off when the animal moults) and the only devices of their kind that can be attached to animals.

MEOP: an international data portal for ocean data collected by marine animals



The international consortium MEOP (Marine mammals Exploring the Ocean Pole-to-pole), originally formed during the International Polar Year in 2008-2009, aims to coordinate at the global scale animal tag deployments, oceanographic data processing and data distribution. The MEOP consortium includes participants from 12 countries (Australia, Brazil, Canada, China, United Kingdom, United States, France, Germany, Greenland, Norway, South Africa and Sweden). The MEOP consortium is associated with GOOS (Global Ocean Observing System), POGO (Partnership for Observation of the Global Oceans), and SOOS (Southern Ocean Observing System). At the European level, the European Animal-Borne Instrument (ABI) EuroGOOS Task Team is about to be launched to facilitate and promote the use of animal-borne instruments. Over 300,000 oceanographic profiles (i.e. representing 1/3 of the total number of Argo profiles) collected by marine biologists have already been made freely available to the international community through the MEOP data portal (Fig. 2). This so-called MEOP-CTD database is a significant contribution to the observation of the world ocean that can be used to conduct regional Polar studies.

The MEOP-CTD database of animal-derived temperature and salinity profiles

An increasing number of studies now show the importance of these remote and inaccessible parts of the ocean, which are so difficult to observe. For example, more than 90% of extra heat in the Earth system is now stored in the oceans and the Southern Ocean in particular is a key site for understanding the uptake of heat and carbon. MEOP provides several thousand oceanographic profiles per year helping us to close gaps in our understanding of the climate system. Instrumented animals complement efficiently other existing observing systems, such as Argo buoys, providing data in sea-ice covered areas and on high-latitude continental shelves. Recent published work (Roquet *et al.,* 2013; Roquet *et al.,* 2014) has shown how important such observations are in predicting ice cover and mixed layer depth in large parts of the oceans where the observations were made. The inclusion of these data should improve significantly the quality of the projections provided by ocean-climate models. All these data are now available into a format file (Argo standard format) easily usable by oceanographers and accessible via the MEOP portal where it can be freely and easily downloaded by users (national data centers, researchers...) with a guaranteed quality level. This database is updated on an annual basis, and it has already been integrated into major oceanographic data centres including NODC, BODC and Coriolis.

References

Boehme, L. *et al.*, 2009. *Technical Note: Animal-borne CTD-Satellite Relay Data Loggers for real-time oceanographic data collection*. Ocean Science, 5:685-695.

Boehme, L. *et al.*, 2013. *Estimates of the Southern Ocean General Circulation Improved by Animal-Borne Instruments*. Geoph. Res. Letts., 40:1-5. doi: 10.1002/2013GL058304

Boehme, L. *et al.*, 2014. *A Southern Indian Ocean database of hydrographic profiles obtained with instrumented elephant seals*. Nature Scientific Data, 1:140028, doi: 10.1038/sdata.2014.28

Expanding the Ocean Biogeographic Information System (OBIS) beyond species occurrences, by including associated environmental data - experiences from the OBIS-ENV-DATA project

Leen Vandepitte, Flanders Marine Institute (Belgium), leen.vandepitte@vliz.be
Daphnis De Pooter, Flanders Marine Institute (Belgium), daphnis.depooter@vliz.be
Klaas Deneudt, Flanders Marine Institute (Belgium), klaas.deneudt@vliz.be
Bart Vanhoorne, Flanders Marine Institute (Belgium), bart.vanhoorne@vliz.be
Pieter Provoost, UNESCO-IOC Project Office for IODE (Belgium), p.provoost@unesco.org
Ward Appeltans, UNESCO-IOC Project Office for IODE (Belgium), w.appeltans@unesco.org
Nicolas Bailly, Hellenic Centre for Marine Research (Greece), nbailly@hcmr.gr
Menashè Eliezer, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), menashe.eliezer@gmail.com
Ei Fujioka, Duke University (United States of America), efujioka@duke.edu
Philip Goldstein, Univ. of Colorado Museum of Natural History (United States of America), philip.goldstein@colorado.edu
Alessandra Giorgetti, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), agiorgetti@ogs.trieste.it
Mirtha Lewis, Centro Nacional Patagónico (Argentina), mirtha@cenpat-conicet.gob.ar
Marina Lipizer, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), mlipizer@ogs.trieste.it
Kevin Mackay, National Institute of Water and Atmospheric Research (New Zealand), kevin.mackay@niwa.co.nz
Gwenaelle Moncoiffe, British Oceanographic Data Centre (United Kingdom), gmon@bodc.ac.uk.
Stamatina Nikolopoulou, Hellenic Centre for Marine Research (Greece), stamatinan@gmail.com
Shannon Rauch, Woods Hole Oceanographic Institution (United States of America), smrauch1@gmail.com
Andres Roubicek, CSIRO Oceans and Atmosphere (Australia), Andres.Roubicek@csiro.au
Carlos Torres, Universidad Autónoma de Baja California (Mexico), ctorres@uabc.edu.mx
Anton Van de Putte, Royal Belgian Institute for Natural Sciences (Belgium), antonarctica@gmail.com
Matteo Vinci, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), mvinci@ogs.trieste.it
Nina Wambiji, Kenya Marine and Fisheries Research Institute (Kenya), nwambiji@gmail.com
Francisco Hernandez, Flanders Marine Institute (Belgium), francisco.hernandez@vliz.be

The Ocean Biogeographic Information System (OBIS) aims to integrate smaller, isolated datasets into a larger, more comprehensive picture of life in our oceans. Therefore, OBIS provides a gateway to many datasets containing information on where and when marine species have been observed. The datasets within OBIS are contributed by a network of hundreds of institutes, projects and individuals, all with the common goal to gain more scientific knowledge and to make these data and knowledge easily available to the public. Until recently, OBIS had solely focused on biogeographic data, in the form of presence of marine species in space and time.

Data collected for biological studies however often include more than just the biological parameters such as presence or abundance. Data collected at the same time can include physical and chemical measurements which can provide insights into the environmental conditions the species live in. Details on the nature of the sampling methods, equipment used and effort can also be of major importance.

In early 2015, the IOC Committee on International Oceanographic Data and Information Exchange (IODE) recommended the establishment of OBIS-ENV-DATA as a pilot project of IODE, under the OBIS umbrella. This project will run for 2 years and involves 11 institutions from 10 countries in North-America, South-America, Europe, Africa and Australia. The project will investigate how 'combined datasets' – datasets containing biological parameters in combination with additional physical and chemical measurements and observations of the habitat – can be kept together. The project will also demonstrate how institutes holding combined datasets can collaborate on the joint management and exchange of such data, and to show the benefit of this approach for marine sciences, biological analysis and modelling. In addition, the project will seek solutions to maximize data sharing with regional and global data systems (e.g., OceanDataPortal or the World Ocean Database), optimize interoperability and minimize the risks of data duplication.

A first workshop was organized in October 2015, bringing together data providers that manage combined datasets. Several project partners contributed pilot datasets to the workshop. The goal was to run through the different data conversion steps to come to a suitable format, discuss the problems encountered and write a preliminary manual that can be used as a guideline for further data exchange. Prior to the workshop, four options were proposed. All pilot datasets were transformed into one or several of these formats, to illustrate and explore potential issues. During the workshop, two additional options were identified, bringing a total of 6 options to the discussion table.

Workshop participants mostly struggled with the fact that the Darwin Core Archive standard (DwC-A) maintains a star schema, where a core file links to different extension files. The general perception was that a simple link between the Occurrence and MeasurementOrFacts extensions would simplify the format and workload. The chosen option includes a customized or extended MeasurementOrFacts Extension (eMoF), allowing a simple link to the Occurrence Extension (Fig. 1). In addition, it includes a hierarchy in events, thereby dealing with a lot of previously identified problems in the management of such data. As the original DwC star schema is still maintained – it only includes an additional link between the two extensions – there is no conflict or incompatibility with IPT, the preferred data exchange mechanism to transfer data to OBIS.

This new OBIS data standard allows for the management of sampling methodology, animal tracking and telemetry data. It also makes it possible to include environmental measurements such as nutrient concentrations, sediment characteristics, climate variables or other abiotic parameters measured during sampling, to characterize the environment where the biogeographic data was collected. The new standard builds on the DwC-A standard and on practices adopted by the Global Biodiversity Information Facility (GBIF). It consists of a DwC Event Core in combination with a DwC Occurrence Extension and a proposed enhancement modification to the DwC MeasurementOrFacts Extension. This new structure enables the linkage of measurements

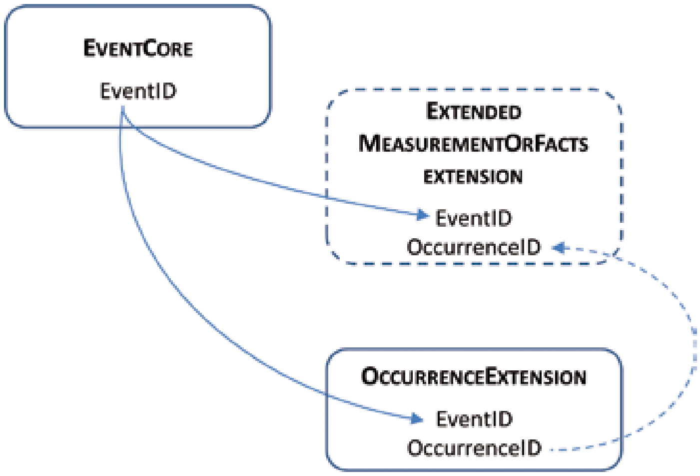


Fig. 1 - Schematic representation of the chosen option. The characteristic IPT star scheme is mainteined. An additional linkage for the biological data is added, more specifically between the Occurrence extension and the newly adopted Extended MeasurementOrFacts extension (eMOF).

or facts - quantitative or qualitative parameters - to both sampling events and species occurrences, and includes additional fields for parameter standardization. The effectiveness of this new format is demonstrated by the pilot datasets available at http://ipt.iobis.org/obis-env/.

The workshop report will be published as a scientific paper to promote the project and its potential for future data management and data exchange. OBIS data providers will be encouraged to submit the environmental data associated with the occurrence data in this newly defined format and OBIS will make these available through their portal and other regional and global repositories such as SeaDataNet and EMODnet.

High resolution and automated flow cytometry data management

Soumaya Lahbib, Mediterranean Institute of Oceanography (France), soumaya.lahbib@mio.osupytheas.fr

Mathilde Dugenne, Mediterranean Institute of Oceanography (France), mathilde.dugenne@mio.osupytheas.fr

Maurice Libes, Observatoire de Sciences de l'Univers (France), maurice.libes@osupytheas.fr

Cherif Sammari, Institut National des Sciences et Technologies de la Mer (Tunisia), cherif.sammari@instm.rnrt.tn

Malika Belhassen, Institut National des Sciences et Technologies de la Mer (Tunisia), belhassen.malika@instm.rnrt.tn

Melilotus Thyssen, Mediterranean Institute of Oceanography (France), melilotus.thyssen@mio.osupytheas.fr

Gérald Grégori, Mediterranean Institute of Oceanography (France), gerald.gregori@mio.osupytheas.fr

For the first time in the North Western Maditerranean sea, an automated flow cytometer (CytoSense), was coupled to a pCO2 and pH sensor mounted after a FerryBox system on board of an opportunity vessel The C/F CARTHAGE of the Compagnie Tunisienne de Navigation (CTN). This vessel crosses the Mediterranean sea four times a week between Tunis (Tunisia) and Marseilles (France) and between Tunis and Genoa (Italy).

This installation was possible thanks to the collaboration between the Tunisian National Institute of Marine science and Technologies (INSTM) and the Mediterranean Institute of Oceanography (MIO) within the frame of the Continuous High Resolution Observation of the MEditerranean sea project. A*MIDEX CHROME aims to study the physicochemical and biogeochemical context of surface waters, integrating high-resolution space and time scales (mesoscale and weekly).

To achieve this goal, datasets needed to be managed and well presented for a better use by scientists. FerryBox datasets of temperature, salinity, pH, chlorophyll-a, fluorescence and dissolved oxygen are managed by the INSTM. While CytoSense dataset, managed by the MIO, is composed of pictures taken by an image in flow device for the recognition of cells up to 20 μm (part of the microplankton group) and several parameters linked to functional groups (pico-nano-microplankton, cryptophytes, cyanobacteria) based on cells pulse shapes and specific optical properties.

For one project, data files acquired from the instrument are batch processed, converted and validated though CytoBase Input Processor, which is a standalone software built on R programm by Mathilde Dugenne. Then, outputs are composed of : one data file (for all measurements) and Picture table able to couple between analysis and the taken picture. Data integration into MySQL database is processed automaticlly using Extract Transform and Load tool (ETL). Finally, data retrieval are displayed through dynamic charts on a web-based interface called "CYTOBASE" [https://en-chrome.mio.univ-amu.fr/?page_id=938]. Results of this work provided to CHROME project scientists' an easy tool for data storage, smart query and decision support that would be deployed in the marine field.

However, CYTOBASE is not interoperable yet with International infrastructure like "SeaDataNet". In fact, several pan European infrastructure manage physico-chemical databases (Coriolis, SeaDataNet), but flow cytometry data don't benefit yet from a proper management and quality control according to international standards among UE projects.

This work presents the first data management methodology and the dedicated tools for this kind of dataset with a possibility to link with international infrastructure once a new flow cytometry vocabulary will be defined.

DINEOF daily cloud-free SST for the Eastern Mediterranean and Black Sea

Andreas Nikolaidis, Oceanography Center of the University of Cyprus (Cyprus), andreas.nikolaidis.cut.cy@gmail.com

George Zodiatis, Oceanography Center of the University of Cyprus (Cyprus), oceanosgeos@gmail.com

Stavros Stylianou, Oceanography Center of the University of Cyprus (Cyprus), stylian@gmail.com

George Nikolaidis, Oceanography Center of the University of Cyprus (Cyprus), nikolaidis_g@hotmail.com

Elena Zhuk, Marine Hydrophysical Institute RAS (Russian Federation), alenixx@gmail.com

Evangelos Akylas, Cyprus University of Technology Department of Civil Engineering and Geoinformatics (Cyprus), evangelos.akylas@gmail.com

DINEOF (Data INterpolating Empirical Orthogonal Functions) is an EOF-based technique to reconstruct missing data in satellite images. Sea surface temperature (SST) hourly cloudy data collected by EUMETSAT organization, have been used in order to present a fully reconstructed set. DINEOF method is described with details at Alvera-Azcárate *et al*. (2009), Beckers *et al*. (2003, 2006), Beckers and Rixen (2003), Nikolaidis *et al*. (2014a, b) and can be downloaded separately from the GeoHydrodynamics and Environment Research Internet page of the University of Liège (GHER). DINEOF is applied and updated daily (Fig. 1) with the latest NAR SST level 3 data. The data from the EUMETSAT Satellite Application Facility on Ocean & Sea Ice used in this study are accessible through the SAF's homepage http://www.osi-saf.org. This project was inspired by a similar GHER project for the Western Mediterranean. A major difference to currently existing projects, is the direct free products availability over the Internet in both image and binary (netCDF) format, at the https://emed-bsea-sst.github.io/Data/ web page.

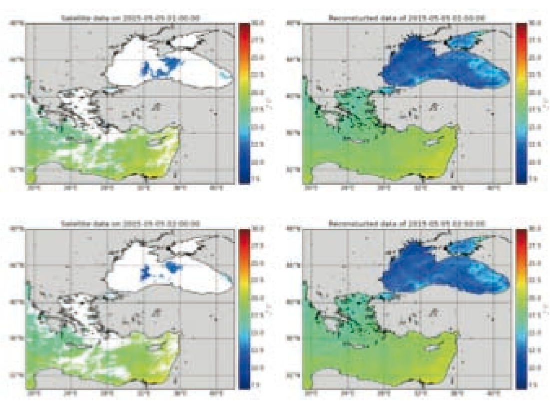


Fig. 1 - Sample of the reconstruction of two hour SST data.

References

ALVERA-AZCÁRATE, A., BARTH, A., SIRJACOBS, D., AND BECKERS, J., 2009, 07. *Enhancing temporal correlations in EOF expansions for the reconstruction of missing data using DINEOF.* Ocean Sci. Discuss. Ocean cience Discussions, 6(2), 1547-1568. doi:10.5194/osd-6-1547-2009

ALVERA-AZCÁRATE A., BARTH A.,SIRJACOBS D., LENARTZ F., BECKERS J.-M., 2011. *Data Interpolating Empirical Orthogonal Functions (DINEOF): a tool for geophysical data analyses*. Medit. Mar. Sci., 5-11.

DINEOF home page http://modb.oce.ulg.ac.be/mediawiki/index.php/DINEOF

Daily DINEOF SST reconstruction in the Western Mediterranean. (n.d.). Retrieved May 12, 2016, from http://gher-diva.phys.ulg.ac.be/DINEOF/dineof.html

NIKOLAIDIS, A., GEORGIOU, G., HADJIMITSIS, D., AND AKYLAS, E., 2014, 01. *Filling in missing sea-surface temperature satellite data over the Eastern Mediterranean Sea using the DINEOF algorithm.* Open Geosciences, 6(1). doi:10.2478/s13533-012-0148-1

NIKOLAIDIS, A., STYLIANOU, S., GEORGIOU, G., HADJIMITSIS, D., AND AKYLAS, E., 2014, 08. *New toolbox in ArcGIS for the reconstruction of missing satellite data using DINEOF algorithm: A case study of reconstruction of Chlorophyll-a gaps over the Mediterranean Sea*. Second International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2014). doi:10.1117/12.2069691

RIXEN M., BECKERS J. M., 2003. *EOF Calculations and Data Filling from Incomplete Oceanographic Datasets.* Journal of Atmospheric and Oceanic Technology, Vol. 20(12), pp. 1839-1856.

Network Common Data Form (NetCDF) http://www.unidata.ucar.edu/software/netcdf/.

The challenge of providing metadata for a 200 year long global mean sea level dataset

Andrew Matthews, Permanent Service for Mean Sea Level (United Kingdom), antt@noc.ac.uk
Lesley Rickards, British Oceanographic Data Centre (United Kingdom), ljr@bodc.ac.uk
Elizabeth Bradshaw, British Oceanographic Data Centre (United Kingdom), elizb@bodc.ac.uk
Kathy Gordon, Permanent Service for Mean Sea Level (United Kingdom), kmg@noc.ac.uk
Angela Hibbert, Permanent Service for Mean Sea Level (United Kingdom), anhi@noc.ac.uk
Svetlana Jevrejeva, Permanent Service for Mean Sea Level (United Kingdom), sveta@noc.ac.uk
Simon Williams, Permanent Service for Mean Sea Level (United Kingdom), sdwil@noc.ac.uk
Philip Woodworth, Permanent Service for Mean Sea Level (United Kingdom), plw@noc.ac.uk

The Permanent Service for Mean Sea Level Dataset

The Permanent Service for Mean Sea Level (PSMSL) provides an internationally recognised focus for knowledge and expertise on sea level science, and constructs and maintains a global data set of long term sea level change information from tide gauges. The PSMSL data bank holds over 67,500 station-years of monthly and annual mean sea level data from over 2300 tide gauge stations, some dating back to the start of the 19th century. Data from each site are quality controlled and, wherever possible, reduced to a common datum, whose stability is monitored through a network of geodetic benchmarks. All data are freely available from the PSMSL website, and are used in a wide range of disciplines, including oceanography, geology, geodesy and climate change studies. The spatial distribution and length of each series in the PSMSL dataset is shown in Figure 1.

The PSMSL was established as a "Permanent Service" of the International Council of Scientific Unions (ICSU) 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. Today, the PSMSL operates under the auspices of the International Council for Science (the current name of ICSU), and is a regular member of the ICSU World Data System.

Metadata Challenges

The sea level record at a location is the result of many years of effort. Over this period many changes in the method of observation will have been made. Several sensors of different types may have been used, perhaps at different locations, and the organisation responsible for taking the measurements may have changed on one or more occasions.

Throughout the observational period the stability of the tide gauge site must be monitored to prevent artificial trends being introduced into the series, usually through annual geodetic levelling. Additionally, larger scale land motion can result from factors including earthquakes, volcanic activity, glacial isostatic adjustment and groundwater extraction.

Before including the data in the PSMSL dataset, the high frequency data must be reduced to monthly mean data in a way that removes any tidal cycle. The sampling rate of high frequency

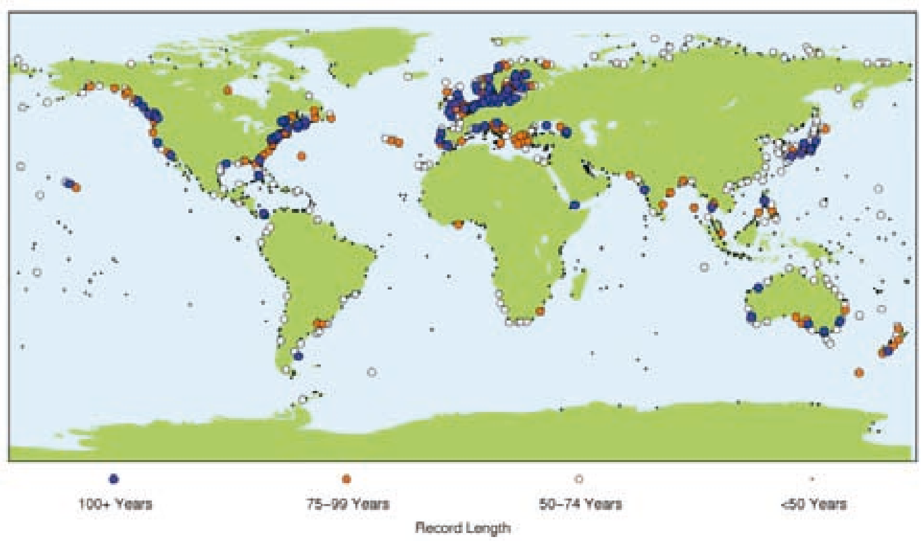


Fig. 1 - Length of Mean Sea Level records in the PSMSL database.

data may have changed, and in some older series observations may only have been made in daylight hours.

As a result, providing a coherent set of metadata that summarises the entire history of sea level measurement at a site, including all the caveats that users of our data should be aware of, can be highly challenging, particularly when historical records can be incomplete, vague, or even contradictory.

As may be expected from an organisation with an over 80 year history, change occurs slowly, and as a result PSMSL has only recently begun to make efforts to provide metadata meeting internationally approved standards. Currently, large amounts of metadata are still presented as blocks of plain text rather than structured information. Other information is not yet publically accessible, in some cases only existing in PSMSL's extensive paper archive.

Here we will describe the PSMSL dataset and our efforts to provide well structured and standardized metadata. Data will soon be distributed in CF-compliant netCDF format. Furthermore, we have been investigating the feasibility of using various OGC standards, including Observations and Measurements (O&M) and SensorML. We will illustrate some of the challenges faced in trying to adopt standards designed to describe 21st century technologies to represent data collected using methods that have been evolving since the 19th century.

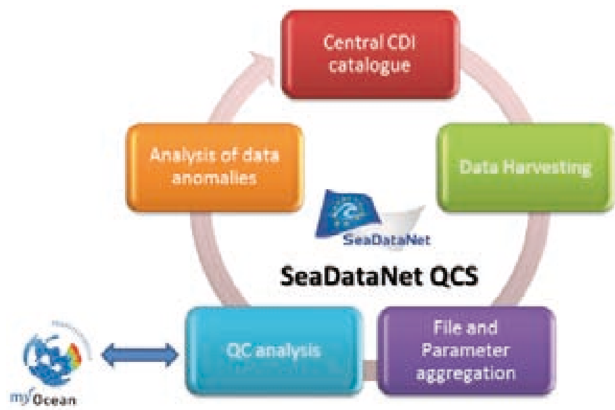
This presentation will not offer a complete solution to these problems, merely some steps towards the latest evolution of an ongoing endeavour to accurately describe the changing height of the ocean. Whatever solutions are adopted will need to meet the requirements of the International Association of Geodesy's Global Geodetic Observing System (GGOS) and the Intergovernmental Oceanographic Commission's Global Sea Level Observing System (GLOSS), and also the needs of other users of our data.

SeaDataNet II data products: the North Atlantic Ocean Region

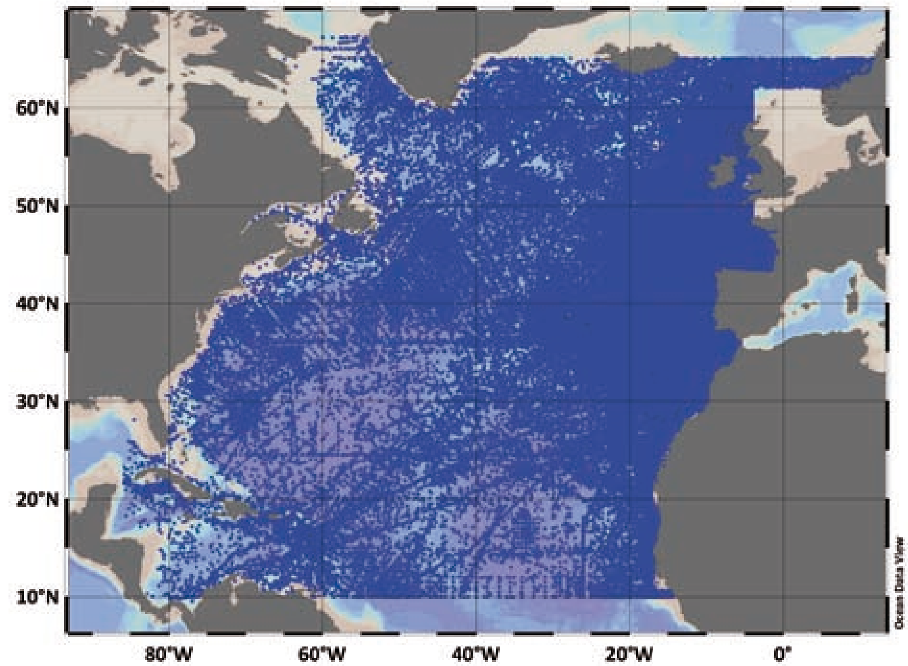
Christine Coatanoan, Institut français de recherche pour l'exploitation de la mer (France), christine.coatanoan@ifremer.fr
Simona Simoncelli, Istituto Nazionale di Geofisica e Vulcanologia (Italy), simona.simoncelli@ingv.it
Reiner Schlitzer, Alfred Wegener Institute (Germany), reiner.schlitzer@awi.de
Michèle Fichaut, Institut français de recherche pour l'exploitation de la mer (France), michele.fichaut@ifremer.fr
Dick Schaap, Marine Information Service (Netherlands), dick@maris.nl
Sissy Iona, Hellenic Centre for Marina Research (Greece), sissy@hnodc.hcmr.gr

During the SeaDataNet II (SDN) EU-project, a quality control (QC) strategy has been implemented and continuously reviewed aiming at improving the quality of the global dataset and creating the best products. This strategy has been developed in collaboration with MyOcean In-Situ Thematic Assemble Centre (INS-TAC) at regional levels to serve operational and scientific oceanography communities. Temperature and salinity historical data collections were created by sea basins, as aggregated datasets and climatology products, and covering the time period 1900-2013. Those ODV qualified dataset collections are available through the SDN web catalog at http://sextant.ifremer.fr/en/web/seadatanet/.

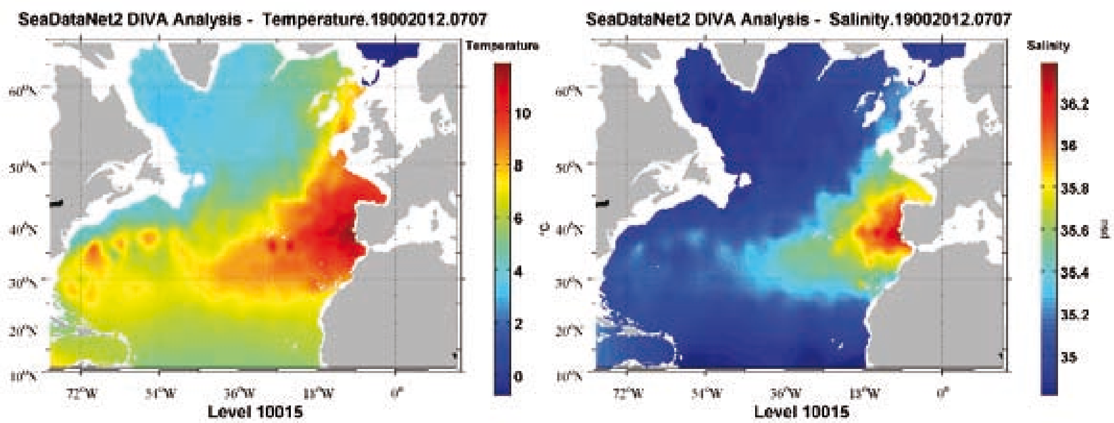
A specific procedure has been implemented to assure and certify the best quality for the datasets (Fig. 1). After the dataset harvesting from the central CDI catalogue, QC has been performed at regional levels in a coordinate way, using the ODV software as a common and basic QC analysis tool. Those datasets have also been scrutinized by the MyOcean regional coordinators, which have sent feedbacks to the SDN regional partners.



- This loop allowed to highlight doubtful data and to organize the data anomalies in lists that have been sent to each concerned data originator together with a guideline to explain the expected corrections. This implemented QC strategy involved the National Oceanographic Data Centers (NODC), on the base of those lists, to check and eventually correct the original data and then to resubmit the corrected data in the SDN dataflow. The QC procedure has also been designed to be iterative in order to facilitate the update and improvement of SDN database content.



Detailed descriptions of the SDN data sets for the North Atlantic Ocean are given. The general description of the dataset, the data quality assessment procedure and results are presented. During SDN, several releases have been produced and the insertion of new data has showed a large increase of the data collection for the North Atlantic Area (Fig. 2). Regarding the number of stations, only a small number of data have been detected as bad.



The final dataset aimed to study in details the water masses circulation in this area (Fig. 3) but integration of new data should improve the quality of the product. Future releases should have to more sustain the QC strategy and encourage NODCs to provide new data and take into account the data quality assessment outcomes.

New Mediterranean Sea climatologies

Simona Simoncelli, Istituto Nazionale di Geofisica e Vulcanologia (Italy), simona.simoncelli@ingv.it
Alessandro Grandi, Istituto Nazionale di Geofisica e Vulcanologia (Italy), alessandro.grandi@ingv.it
Sissy Iona, Hellenic Centre for Marina Research (Greece), sissy@hnodc.hcmr.gr

New temperature and salinity monthly climatologies have been produced for the Mediterranean Sea within the framework of SeaDataNet2 EU project. The climatologies are based on the V1.1 historical data collection of all available temperature and salinity in situ profiles (http://dx.doi.org/10.12770/cd552057-b604-4004-b838-a4f73cc98fcf) spanning the time period 1900-2013 (Simoncelli *et al*., 2016).

The Mediterranean climatology is defined between 9.25°W-36.5°E of longitude and 30-46°N of latitude with an horizontal resolution of 1/8 of a degree on 33 IODE vertical standard levels. DIVA software (4.6.9 version) has been used (Troupin et al, 2012) for both the analysis and reference field computation. The salinity reference field has been computed through annual semi-normed analysis considering all available observations, while monthly temperature reference fields have been considered due to the large temperature seasonal variability.

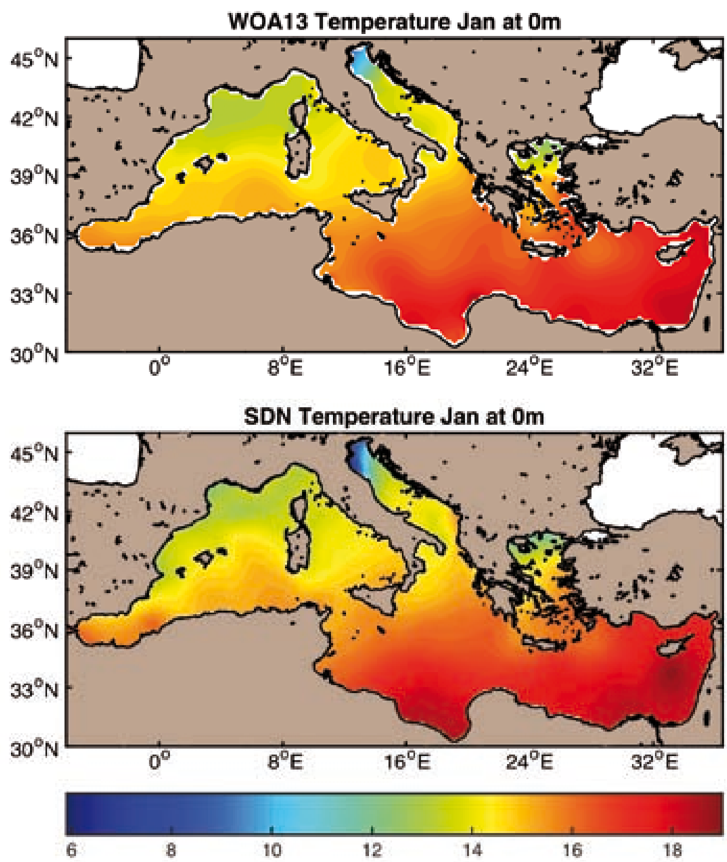


Fig. 1 - January temperature maps at the surface from WOA13 (top) and the corresponding SeaDataNet2 (SDN) field (bottom).

Consistency Analysis

The quality of the Mediterranean Sea climatology has been analyzed considering the World Ocean Atlas (WOA13) temperature (Locarnini *et al.*, 2013) and salinity (Zweng *et al.*, 2013) climatologies as a reference. WOA13 climatology has 1/4 of horizontal resolution and it is defined over 57 vertical levels from the surface up to 1500m, thus the consistency analysis has been performed on 24 coincident levels. We first checked the consistency between SDN and WOA13 by visual inspection. Then we used statistical indexes like BIAS and RMSE to quantify the differences between SDN and WOA13 fields (Fig. 2).

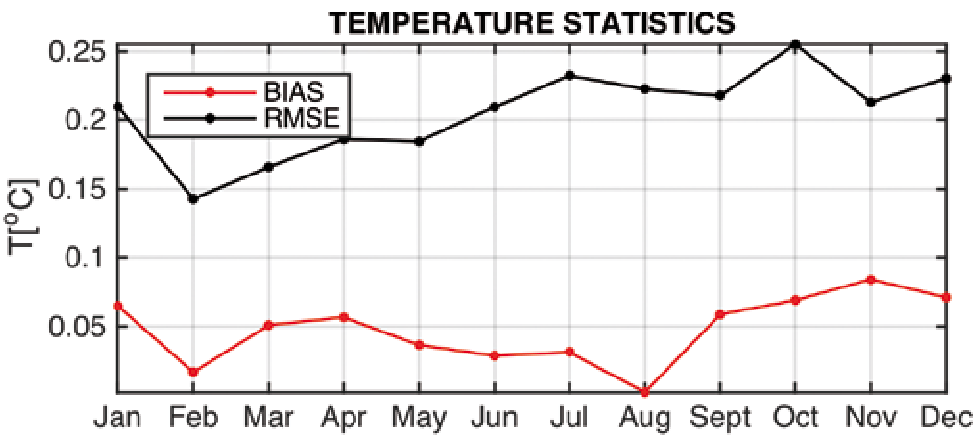


Fig. 2 - Basin monthly averages of BIAS and RMSE computed between temperature SeaDataNet2 (SDN) climatology and WOA13.

Conclusions and Developments

The new SeaDataNet Mediterranean Sea climatologies are in good agreement with the WOA13 but presents higher horizontal resolution and it covers the entire water column. The quality of the presented climatology allows its use for many applications.

New climatologies of the recent era, which considers the available in situ observations starting from the mid nineties, and a new consistency analysis will be presented considering Copernicus Marine Environment Monitoring Service products.

References

Simoncelli S., Coatanoan C., Back O., Sagen H., Scory S., Myroshnychenko V., Schaap D., Schlitzer R., Iona S., Fichaut M., 2016. *The SeaDataNet data products: regional temperature and salinity historical data collections.* EGU 2016 - European Geosciences Union General Assembly 2016. 17-22 April 2016, Austria.

Troupin, C, Barth, A, Sirjacobs, D, Ouberdous, M, Brankart, J.-M, Brasseur, P, Rixen, M, Alvera Azcarate, A, Belounis, M, Capet, A, Lenartz, F, Toussaint, M.-E, and Beckers, J.-M., 2012. *Generation of analysis and consistent error fields using the Data Interpolating Variational Analysis (Diva).* Ocean Modelling, 52-53, 90-101. doi:10.1016/j.ocemod.2012.05.002.

Development of the Impulsive Noise Register System[1]

Neil Holdsworth, International Council for the Exploration of the Sea (Denmark), neil.holdsworth@ices.dk
Carlos Pinto, International Council for the Exploration of the Sea (Denmark), carlos.pinto@ices.dk

Introduction

Underwater noise, sound that has the potiental to cause negative impacts on marine life[2] is one of the key descriptors of marine ecosystem health under the Marine Strategy Framework Directive (MSFD). Noise is further categorized into ambient (background noise without distinguishable sources) and impulsive (sound comprising one or more short duration pulses, from sources such as airgun arrays, impact pile driving, powerful military search sonars and explosions) – in this system we focus solely on impulsive noise. The reporting and monitoring of impulsive underwater noise events is a relatively new requirement in national and regional monitoring programmes. The establishment of an impulsive noise register system for OSPAR and HELCOM, to be used for assessment purposes of the North East Atlantic and Baltic regions respectively.

In essence the end users are the regional assessment leads working within their expert groups to deliver indicator based assessments on impulsive noise to the OSPAR Intermediate Asessment 2017, and the 2nd Holistic assessment (HOLAS II) in HELCOM. Their requirement was for a system that would comprise a web interface where noise events would be uploaded, quality controlled, displayed via a map interface and downloaded via a number of services. In addition, the system should be capable of calculating indicators of impulsive noise into a standard data product that can be stored, downloaded or queried from the web map interface.

The system relies on national registers of licenced activities, as such it was not suitable to make a continuous harvesting of data between the national registers and the Impulsive Noise Register as the national registers are fluid – the activities change in between when the licence is granted and the actitivity is carried out. Whereas the purpose of the Impulsive Noise Register is to show an inventory in a reporting year of actual activity that can be used directly to produce an indicator (pulse block days per assessment unit).

Reporting mechanism

The first challenge of developing the system was to define a standard reporting format. The starting point was the TG Noise format – a tabular description of required fields transposed into Excel. ICES defined a data model from this description and developed an Excel template which was then underpined with an XML schema representing the noise format. A macro button in the Excel template then exports (and validates) the filled spreadsheet into the XML file following the XML Noise Schema.

1 ICES, 2016 http://underwaternoise.ices.dk .

2 TGNOISE final report, Feb 2012, http://ec.europa.eu/environment/marine/pdf/MSFD_reportTSG_Noise.pdf.

Data submitters have therefore have 2 options to provide validated data into the International noise register; either by filling in the Excel form template and using the macro button, or creating directly from their national database/system the XML files using the XML NOISE schema. The validated files are uploaded via the 'upload service' on the web interface (Fig. 1).

The spatial reporting units are heterogenous in size and shape. Data submitters can report according to licence areas (polygons), ICES sub-square (the same unit as the indicator) or by points (lat/lon, WGS84). This creates a challenge when displaying the events on a map, and calculating a indicator based on a standard spatial assessment unit.

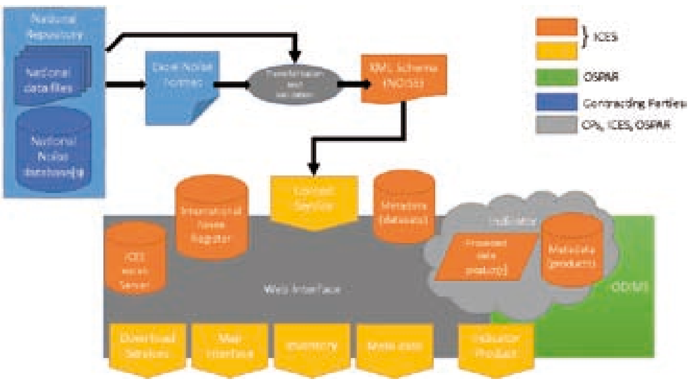


Fig. 1 - Data workflow for international noise register and operational indicator.

Output and visualization

The output calculation is simple, but based on some arbitrary decisions on the spatial extent of a noise event. All the reporting is converted into ICES sub-squares. In the case of reported polygons events, any overlapping ICES sub-square that intersects the reported polygon is included as a pulse block day event.

ICES has developed a map interface (currently only available to HELCOM/OSPAR for Beta testing) for the visualization of the noise base data, the calculated indicators and the reference layer reporting units. Users can query the Impulsive noise base data by year, noise value codes, source events, and mitigation systems.

Also available are the indicators (Pulse Block Days per OSPAR Region [see figure 2], Pulse Block Days per ICES sub-square or pulse block days by value code) and the reporting units (ICES sub-rectangle, UK Licence Blocks, German Naval Polygon). All these layers/data are available for download as shapefiles, images, and in some cases as excel files.

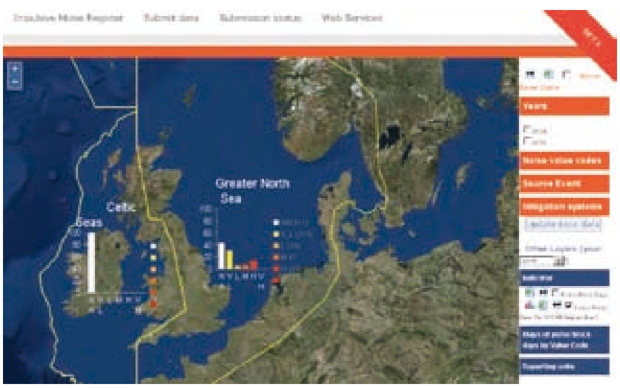


Fig. 2 - Map facility showing the indicator of Pulse Block days per OSPAR Region.

EMODnet Physics: past, present and future

Patrick Gorringe, EuroGOOS (Belgium), patrick.gorringe@eurogoos.eu
Giuseppe Manzella, ETT People and Technology (Italy), giuseppe.manzella@ettsolutions.com
Dick Schaap, Marine Information Service (The Netherlands), dick@marils.nl
Sylvie Pouliquen, Institut français de recherche pour l'exploitation de la mer (France), Sylvie.Pouliquen@ifremer.fr
Lesley Rickards, British Oceanographic Data Centre (United Kingdom), ljr@bodc.ac.uk
Antonio Novellino, ETT People and Technology (Italy), antonio.novellino@ettsolutions.com

Access to marine data is of vital importance for marine research and a key issue for various studies, from climate change prediction to off shore engineering. Giving access to and harmonising marine data from different sources will help industry, public authorities and researchers find the data and make more effective use of them to develop new products, services and improve our understanding of how the seas behave.

EMODnet Physics provides a combined array of services and functionalities (facility for viewing and downloading, dashboard reporting and machine-to-machine communication services) to obtain, free of charge data, meta-data and data products on the physical conditions of European sea basins and oceans. Moreover, the system provides Web Services and Web catalogues in order to exchange data and products according to the most recent standards. EMODnet Physics is built on and is working in collaboration with EuroGOOS and its ROOSs, CMEMS and the SeaDataNet network of NODCs. By means of joint activities with its three pillars and with the most relevant organizations and associations within the sector, EMODnet Physics is undergoing significant improvements and expansion. The portal is providing access to data and products of: wave height and period; temperature and salinity of the water column; wind speed and direction; horizontal velocity of the water column; light attenuation; sea ice coverage and sea level trends.

EMODnet Physics is continuously increasing the number and type of platforms in the system by unlocking and providing high quality data from a growing network. Nowadays the system integrates information from more than 11.000 platforms, among which 2915 moorings, 2728 drifting buoys and around 1200 ARGO floats.

EMODnet Physics has also been updated with two ready-to-use data products: the Ice (Copernicus CMEMS - SEAICE_GLO_SEAICE_L4_NRT_OBSERVATIONS_011_001) and Sea Level Trends (produced through the Permanent Service for Mean Sea Level - PSMSL).

EMODnet Physics actively collaborates with EU wide projects and initiatives such as H2020 Jerico-NEXT, AtlantOS, INSTAC and SeaDataNet in order to explore synergies and build strong links to foster data collection, integration and dissemination, building on existing infrastructures and by this avoiding duplication of efforts.

EMODnet Physics will continue to further extend the number and type of data and platforms feeding the system; improve the capability of the system producing data and products that could match the market needs of the current and potential new end and intermediate users; to connect with other initiatives at European and global scale as to stimulate international exchange of oceanographic data and products and by this encourage the development of a coordinated network.

EMODnet Chemistry: biogeochemical data management for long-term research and support to EU policies

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

Alessandra Giorgetti, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), agiorgetti@inogs.it

Marina Lipizer, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), mlipizer@inogs.it

Alberto Brosich, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), abrosich@inogs.it

The EMODnet initiative is developed through a stepwise process that foresees different phases.

The positive outcomes from the EMODnet Chemistry Pilot phase led to a second ongoing phase where the partnership, the parameters requested and the geographical coverage have been enlarged. As in the Pilot phase, the aim is to make available and reusable a large amount of fragmented and inaccessible data, hosted by European research institutes and environmental agencies. Furthermore, there is an increased focus to provide tools useful to address the requirements of the Marine Strategy Framework Directive –MSFD, developing visualization and data access services.

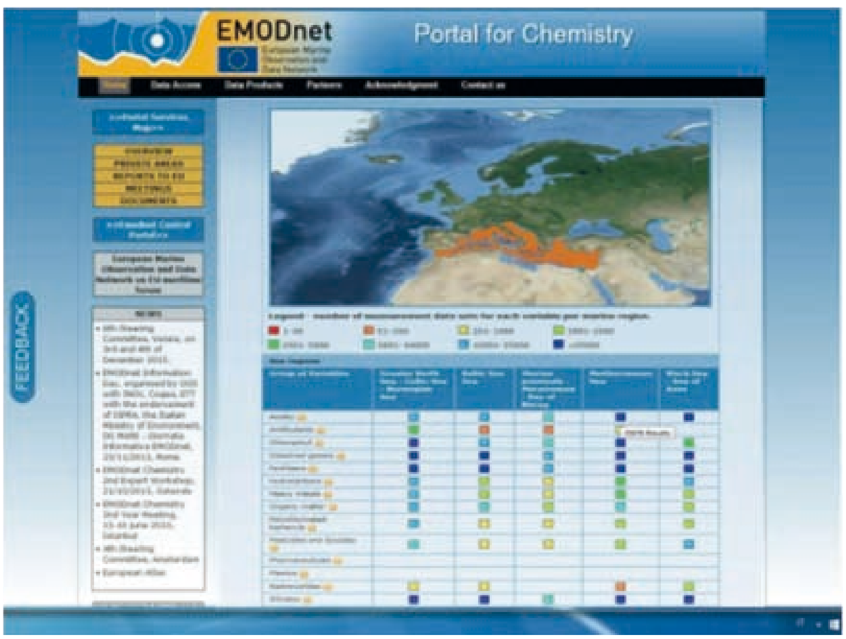


Fig. 1 - EMODnet Chemistry data portal: quick search of chemicals by region (here showing acidity data in the Mediterranean sea).

The technical set-up is based on the principle of adopting and adapting the SeaDataNet infrastructure. SeaDataNet is developing a decentralized infrastructure following EU-INSPIRE implementation rules; it can be considered a European *de facto* standard with more than 100 nodes connected; it is part of the Ocean Data Interoperability Platform (ODIP project) that links the experience of European data management infrastructures with American and Australian ones. The management relies on a distributed network of National Oceanographic Data Centres. They contribute to data harvesting and enrichment with the relevant metadata. Data are processed into interoperable formats (using agreed standards ISO XML, ODV) with the use of common vocabularies and standardized quality control procedures.

Data quality control and data aggregation are key points for EMODnet Chemistry as the consortium is dealing with heterogeneous data coming from different sources. A data validation loop has been agreed within the EMODnet Chemistry community and is routinely performed. Data quality control is first done at national and then at regional level (in the Atlantic, in the Baltic, in the North, in the Mediterranean and in the Black Sea).

Beside quality control, aggregations of parameters and units conversions are performed routinely thanks to the development of a dedicated vocabulary (P35) embedded in the ODV software. This combination of tools provides an effective help to the regional leaders to homogenize the great variability of harvested information in categories relevant for a standard environmental reporting. A panel of experts has been established to discuss and validate a standard and agreed vocabulary.

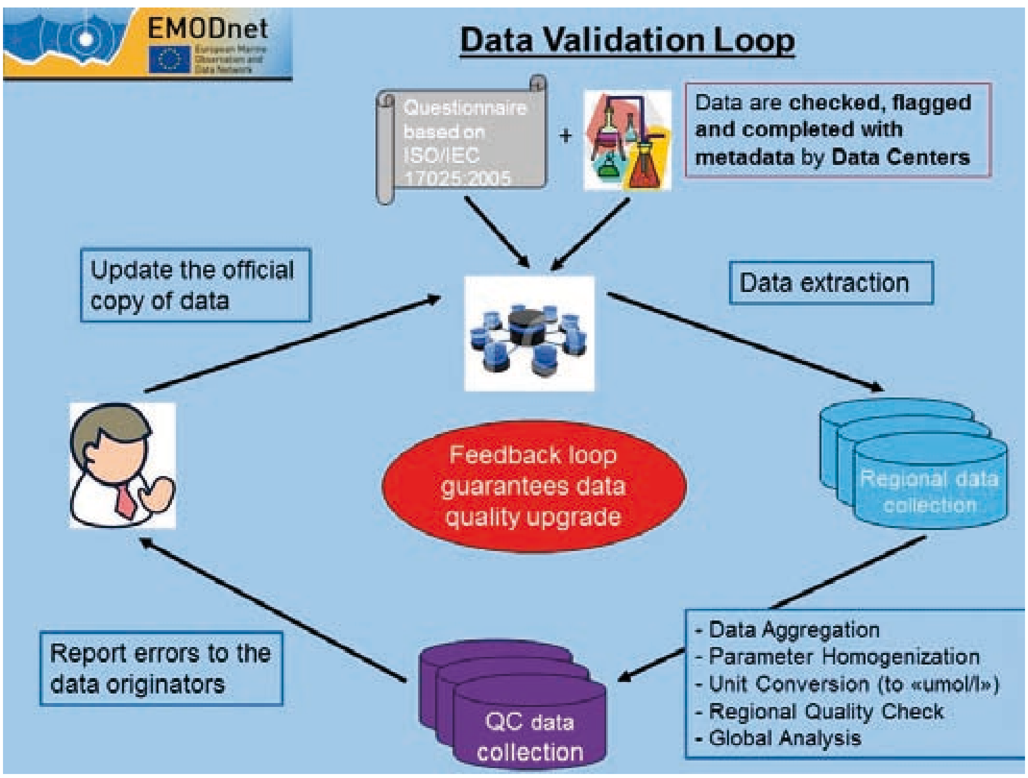


Fig. 2 - Date validation loop from data collection to validation and updating.

The validated and aggregated regional datasets are used to develop data products useful for analysis of long- term variability in biogeochemical parameters and to address MSFD requirements. EMODnet Chemistry provides interpolated maps of nutrients, generated as 10 years seasonal running windows, useful to detect temporal and spatial variability, and services for the visualization of time series and profiles of several chemical parameters. All visualization services are developed following OGC standards as WMS and WPS.

In order to test new strategies for data storage, reanalysis and to upgrade the infrastructure performances, EMODnet Chemistry has chosen the Cloud environment offered by Cineca (the Consortium of Italian Universities and research institutes) to store both regional aggregated datasets and analysis, with their visualization services.

EMODNet Bathymetry - building and providing a high resolution digital bathymetry for European seas

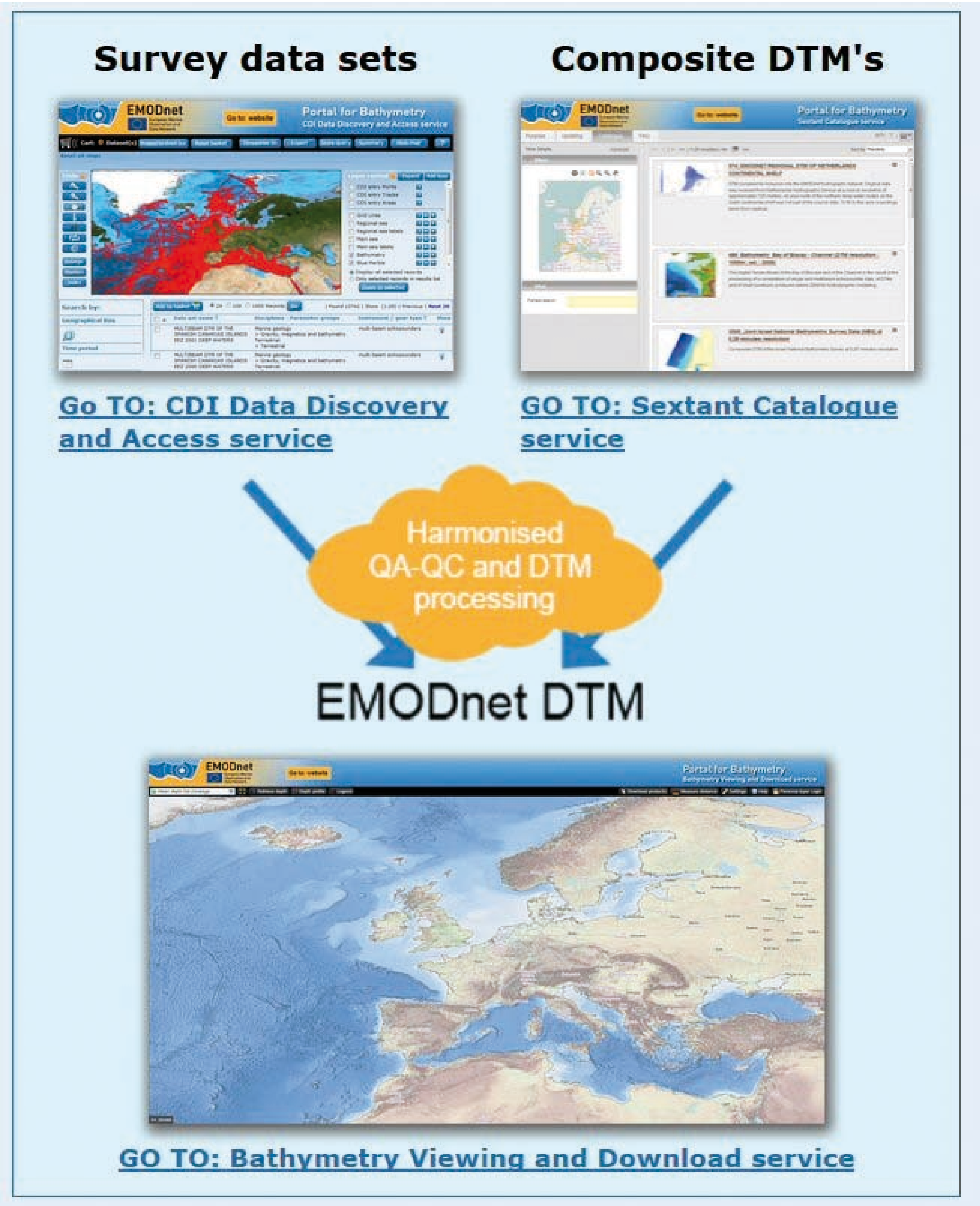
Dick Schaap, Marine Information Service (The Netherlands), dick@maris.nl
On behalf of EMODnet Bathymetry consortium

Access to marine data is a key issue for the **EU Marine Strategy Framework Directive** and the **EU Marine Knowledge 2020** and **Blue Growth agendas** and includes the **European Marine Observation and Data Network (EMODnet)** initiative. EMODnet aims at assembling European marine data, data products and metadata from diverse sources in a uniform way. The EMODnet data infrastructure is developed through a stepwise approach in three major phases. Currently EMODnet is entering its 3rd phase with operational portals providing access to marine data for bathymetry, geology, physics, chemistry, biology, seabed habitats and human activities, complemented by checkpoint projects, analysing the fitness for purpose of data provision.

The EMODnet Bathymetry project develops and publishes Digital Terrain Models (DTM) for the European seas. These are produced from survey and aggregated data sets that are indexed with metadata by adopting from SeaDataNet the Common Data Index (CDI) data discovery and access service and the Sextant data products catalogue service. SeaDataNet is a network of major oceanographic data centres around the European seas that manage, operate and further develop a pan-European infrastructure for marine and ocean data management. SeaDataNet is also setting and governing marine data standards, and exploring and establishing interoperability solutions to connect to other e-infrastructures on the basis of standards such as ISO and OGC. The SeaDataNet portal provides users a number of interrelated meta directories, an extensive range of controlled vocabularies, and the various SeaDataNet standards and tools. SeaDataNet at present gives overview and access to more than 1.8 million data sets for physical oceanography, chemistry, geology, geophysics, bathymetry and biology from more than 100 connected data centres from 34 countries riparian to European seas.

The latest EMODnet Bathymetry DTM has a resolution of 1/8 arcminute * 1/8 arcminute and covers all European sea regions. Use is made of available and gathered surveys and already more than 13.000 surveys have been indexed by 27 European data providers from 15 countries and originating from more than 120 organisations. Also use is made of composite DTMs as generated and maintained by several data providers for their areas of interest. Already 44 composite DTMs are included in the Sextant data products catalogue. For areas without coverage use is made of the latest global DTM of GEBCO who is partner in the EMODnet Bathymetry project. In return GEBCO integrates the EMODnet DTM to achieve an enriched and better result. The catalogue services and the generated EMODnet can be queried and browsed at the dedicated EMODnet

Bathymetry portal which also provides a versatile DTM viewing service with many relevant map layers and functions for retrieving. Activities are underway for further refinement following user feedback. The EMODnet DTM is publicly available for downloading in various formats.



The presentation will highlight key details of the EMODnet Bathymetry project, the latest release of the EMODNet DTM for all European seas, its portal and its versatile viewer, and will look forward to the next phase.

Towards better data access: the development of a new data download tool for marine biological data

Simon Claus, Flanders Marine Institute (Belgium), simon.claus@vliz.be
Leen Vandepitte, Flanders Marine Institute (Belgium), leen.vandepitte@vliz.be
Filip Waumans, Flanders Marine Institute (Belgium), filip.waumans@vliz.be
Bruno Vitorino Pino, Flanders Marine Institute (Belgium), bruno.pino@vliz.be
Francisco Hernandez, Flanders Marine Institute (Belgium), francisco.hernandez@vliz.be

A new tool to easily select and download marine biological data has been developed within the framework of the European Marine Observation and Data Network, EMODnet. A detailed user analysis and functional analysis of the current EMODnet Biology dataportal has been performed. Based on these analyses, a new concept that differentiates more between data – unprocessed raw observations or measurements – and data products derived from the data has been formulated. The data component focuses on its easy and intuitive downloading, while the focus of the data products component focuses on a good visualization, a better overview of existing data and a quick understanding of the data.

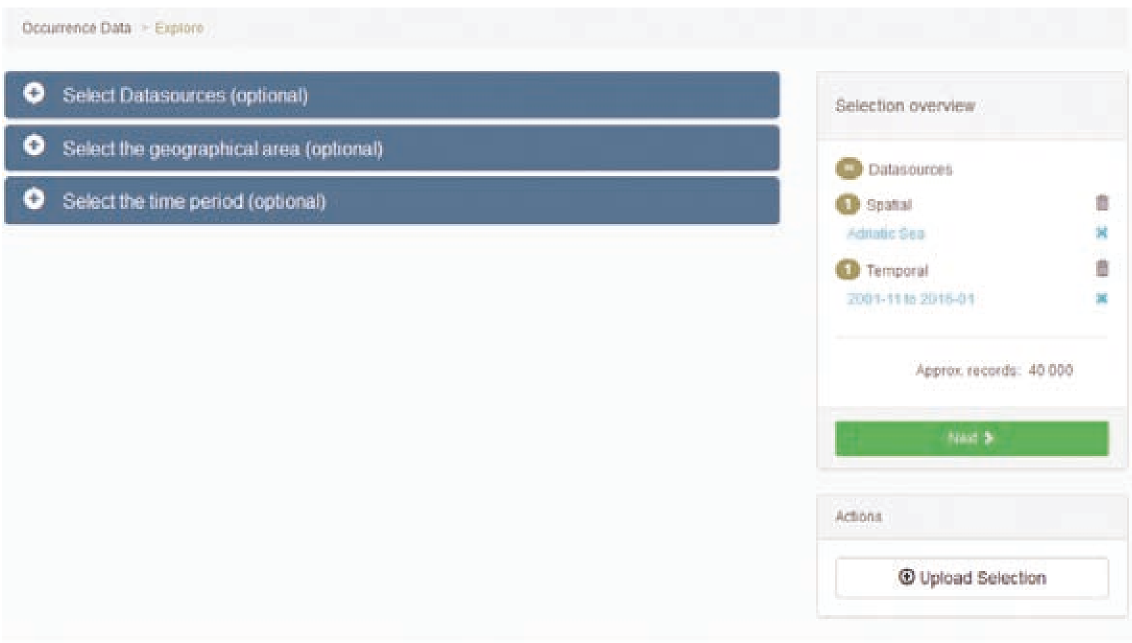


Fig. 1 - Screenshot of the EMODnet biology download toolbox, for fast selection and downloading of European marine biological data.

The data download component guides the user through a step-wise workflow where one can select datasets, perform predefined geographic and temporal selections, can add specific taxonomic or functional filters, and select for data with a certain quality and precision. At any time the user will have a clear overview of the specific performed filters and queries through a selection overview. The final step of the data download toolbox allows the actual downloading of the selected data. We ask for limited user information, such as the purpose of data download to monitor better the usage.

The data is stored in two linked Darwincore dataschemes – the Occurrence data scheme and the Measurement of Fact datascheme. The Occurrence datascheme contains the information of the geographic location, time and taxonomic information of the sample - addition of a "MeasurementsOrFacts" datascheme "MeasurementsOrFacts" allows EMODnet to capture biological measurements and related abiotic data such as e.g. length- and weight information of taxa, stomach content data or the sediment-composition at the time of sampling, which was previously not possible. the link between the Occurrence records and the Measurement or Facts is made through the EventID and OccurrenceID.

The selected data can be downloaded as csv (comma separated value), the selection can be saved a json file (JavaScript Object Notation), the selection can be shared through a unique webURL and the data can be accessed through a webservice (WFS).

The EMODnet Biology downlaod toolbox is freely available at http://www.emodnet-biology.eu/toolbox/en/toolbox/occurrence/explore.

EMODnet essential data needs and gaps: a comparative review of the Atlantic, Black Sea and Medsea Checkpoints

Jacques Populus, Institut français de recherche pour l'exploitation de la mer (France), jacques.populus@ifremer.fr

Atanas Palasov, Institute of Oceanology Bulgarian Academy of Science (Bulgaria), palazov@io-bas.bg

Nadia Pinardi, Istituto Nazionale di Geofisica e Vulcanologia (Italy), n.pinardi@sincem.unibo.it

Eric Moussat, Institut français de recherche pour l'exploitation de la mer (France), eric.moussat@ifremer.fr,

Frédérique Blanc, Collecte Localisation Satellites Group (France), fblanc@cls.fr

Costs of ocean observation are huge and – in Europe - are largely carried by Member States. The resulting data is stored in a myriad of national, regional and international databases. EU Programmes and Directives have been taken to collect observations, to harmonize data and services and to assemble data both for assessing health of the marine environment and for a sustainable Blue Growth.

Now, the overarching aim of the Marine Knowledge philosophy (Marine Knowledge 2020) is to "collect once and use many times".

However do the existing monitoring systems fit the user needs? If no, what is-it due to?

To answer these questions, the European Marine Observation and Data network (EMODnet)

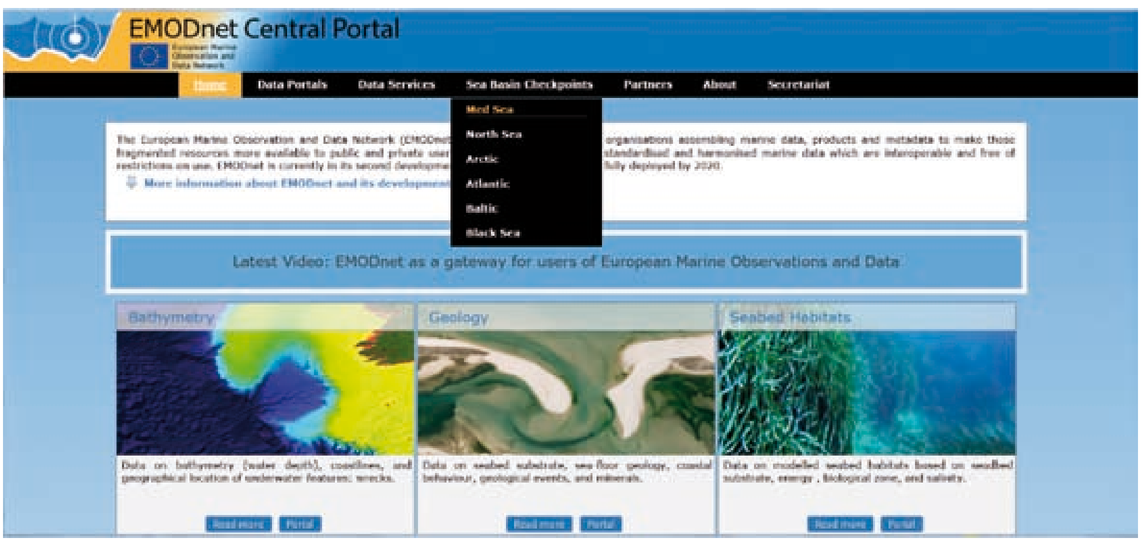


Fig. 1 - EMODnet Central portal (http://www.emodnet.eu/): Sea Basin checkpoints.

checkpoints have been committed with the assessment of marine monitoring systems with a view to supporting sustainable Blue Growth in European Sea Basins.

To do so, a series of end-user applications ("challenges") of paramount importance for:

- the Blue Economy sector (offshore industries, fisheries);
- the marine environment variability and change (eutrophication, river inputs and ocean climate change impacts);
- emergency management (oil spills);
- and preservation of natural resources and biodiversity (Marine Protected Areas)

have been selected by the EU Directorate-General for Maritime Affairs and Fisheries (also known as DG MARE).

The purpose of these challenges is to develop innovative products from existing data sources and to assess the fitness for use of these data for these applications.

The results are designed for:

- Institutional stakeholders for decision making on observation and monitoring systems
- Data providers and producers to know how their data collected once for a given purpose could fit other user needs
- End-users interested in a regional status and possible uses of existing monitoring data

The approach initiated by the Medsea checkpoint and adopted by the Atlantic and Black Sea checkpoints involves the development of information and indicators based on a common reference framework : the Geographic Information standards for Data Product Specifications, Data Quality and Metadata (ISO 19131, ISO 19157 and ISO 19115 respectively) and Common Vocabularies : environmental matrices and SeaDataNet lists of Parameter Groups (PO3), of Discovery Parameters (P02) and of BODC Parameter Usage (P01) ensuring consistency of the gap analyses of input data sources between the basins.

The fitness for use of the input datasets are assessed using 2 categories of criteria to determine how these datasets fits the user requirements which drive them to select a data source rather than another one and to show performance and gaps of the present monitoring systems :

- data appropriateness i.e. the extent to which data ("what") fits the user needs in term of "completeness", "consistency" and "accuracy"
- data availability i.e. the extent to which data can be discovered and obtained by users ("how") in term of "visibility", "accessibility" and "performance"

The initial tasks of the challenges have been of three types:

- setting up a list of characteristics (parameters and objects of interest for the challenge) based on partners knowledge and conceptual representation of their challenges;
- listing the monitoring systems used by the checkpoints: data sources are informed in a "many to many" relationship with the characteristics.
- giving an insight into the quality of the data sources (in terms of data availability and appropriateness in relation to each challenge).

The comparison between the three basins is made on the basis of simple statistics which have been produced thanks to the harmonization of the quality concepts and metadata as well as the adoption of common vocabularies.

This work was performed inter alia on three basins, the Black Sea, the Mediterranean and the Atlantic. The results revealed discrepancies in the way basins are covered by existing monitoring systems and why for, but also by the way people approach a challenge from a basin to the next (due to environmental differences, different contexts or cultures etc.).

As a conclusion, most salient data gaps were already identified. They will be further investigated as the project draws on by performing a full download of each data set needed by the challenges, by describing data quality and by thoroughly reporting it in two successive “Data adequacy reports”. The final aim of this endeavour is to provide the Commission with a detailed overview of data gaps and redundancies enabling a more focused data acquisition policy to open existing monitoring systems to a wider community of users.

Streamlining the data ingestion process in the EMODnet context

Serge Scory, Royal Belgian Institute of Natural Sciences (Belgium), Serge.Scory@naturalsciences.be

Michèle Fichaut, Institut français de recherche pour l'exploitation de la mer (France), Michele.Fichaut@ifremer.fr

Sissy Iona, Hellenic Centre for Marine Research (Greece), sissy@hnodc.hcmr.gr

Antonio Novellino, ETT People and Technology (Italy), antonio.novellino@ettsolutions.com

Dick Schaap, Marine Information Service (The Netherlands), Dick@maris.nl

Thanks to EMODNET, the European marine and observation data network, since 2009, marine data have now their "one-stop-shop". They are easy to find, freely available, interoperable and reliable. Bringing together more than 100 organisations that observe and measure the oceans, process data according to international standards, EMODNET is one of the most successful initiatives of the Integrated Maritime Policy. Quality controlled and accurate data of the maritime environment are critical to underpin the sustainable growth of the blue economy.

Many data however, collected by public authorities, researchers and private operators of coastal or offshore facilities still do not arrive to the national or regional repositories and are thus unavailable to scientists, engineers and other potential users. In order to facilitate and streamline the process whereby marine data from whatever source is delivered on a voluntary basis for safekeeping to data repositories from where it can be freely disseminated the European Commission granted recently a 3-year contract aiming at setting up a "data ingestion mechanism".

The activities started recently (May 2016) and are undertaken by a European network of 44 organisations (governmental departments, marine research institutes and SME's) from 29 coastal countries. Geographically the network has nodes in the countries around all European marine basins and it covers also all EMODnet data themes.

The challenges and expectations are great: the major services must be made operational within the first year of the contract!

Our communication will present the technical principles of the Data Ingestion Portal, the associated work flows, the initiatives that will be taken to encourage the mobilization of new marine data resources and the foreseen provision for the sustainability of the portal.

POSTERS

Classification of geological data as a key to optimizing the preparation of CDI and ODV in the Geo-Seas project

Georgij Konshin, University of Latvia (Latvia), georgij@lu.lv
Valdis Seglinsh, University of Latvia (Latvia), valdis.seglins@lu.lv

In this report consider some topical issues of implementation the project GeoSeas in order to optimize the preparation of records CDI and ODV. The main reason that prompted us to prepare this report is that the existing guidelines and instructions for all their quality have one drawback: the examples tend to relate to the field of oceanography and is not always understood by geologists. In other words, over the Geo Seas dominates the experience of its predecessor SeeDataNet. Therefore it is reasonable to explain the problems we encountered in simple examples and we hope that it will be useful for geological data description and probably is necessary to determine main questions solved during the project progress.

Bottom sediments - main pilot object

The bottom sediments (BS) are the substantial object in the Latvian marine geology archives. As a result of geological survey of the Baltic Sea a significant number of geological samples or cores were collected. Their capacity usually ranges from 0.0 to 20-30 cm, rarely reaching several meters. Most of bottom sediment samples were selected through sediment grabs, rarely with a vibrocorer or with drilling. In the first case, the samples do not have the form of core and layered descriptions. However, in other cases, samples were collected as cores and have a simplify sediment (lithologic) description. Samples of bottom sediment were taken at 3,459 stations. The acquired samples were further analyzed in the laboratory using different analyses such as grain size analysis, chemical analysis, spectral analysis, etc. The results of each of these analyses have stored in a separate table of local database Geobank.mdb referring to the same metadata of the field sampling.

Data set definition

We considered that the separate data set should relate to one station and should coincide with same data set in a local centre where we can download them. The separate dataset must include three groups of metadata The first group includes the metadata that relate to the CDI XML format. This metadata define data set identification and give answers to the following basic questions e.g. Where? When? What? How? Who? Where to find data? Station / Cruise data? Other relevant information? Other relevant services? The second group includes the metadata that relate to the conditions of sampling. The third group includes the metadata that relate to the results of analysis of samples.

The data set name

Consists of a combination of three elements: abbreviation of the geological object type (e.g., BS - Bottom sediments); abbreviations of analyze: grn (granulometric), ch (chemical), sp (spectral), min (mineralogical) ptr (petrographical) or it combinations (e.g., grnsp); station's number throughout local database (e.g., 1003). So get like BSgrn1003. The names of Dataset_ID for the identification data set consist of the last two elements, for example, grn1003.

Creating the Test database for approval the methods of CDI records and ODV files preparation

To create a pilot database must be clear about the basic principles and mechanism of action of Geo-Seas as an information retrieval system. As can be seen, an important feature of GeoSeas is the use of such specific means as CDI and ODV. It should be emphasized that the use of ODV as a data transport format is one of the characteristics of GeoSeas, what distinguishes this system from databases EU-Seased (Eurocore, Euromarsin. Euroseismic), where access to samples and other data sets is determined by negotiation between the user and the repository where the sample is stored..

Data set classification

It is clear that in determining the composition of test database is necessary to consider how it presents geological data A. Sampling factors/conditions: A1. At the station was taken only one sample usually by sediment grab. It not looks like a core. A2. At the station were taken a few samples usually by vibrocorer or drilling. Samples were taken at different depths below the bottom of the sea. It looks like a core, for which a core logs / profile was made (lithology). B. Laboratory studies. B1. The sample was studied for one kind of analyze (e.g. spectral analyze). That is Simple data set. B2. The Sample was studied for several kind of analyze (e.g. spectral and granulometric). That is Complex data set. Simple data set are divided into 6 groups: BSgrn, BSch, BSmin, BSpt, BSsp with the data of granulometric, chemical, mineralogical, petrographical, spectral analysis, respectively. Complex data set are divided into 9 groups: BSgrnsp, BSgrnch, BSgrnmin, BSspptr, BSgrnchsp, BSgrnptrsp, BSgrnminch, BSrgnminsp, BSgrnspchmin with the data of two or three or even four kinds of analysis of the six possible. Based on the proposed classification, we created a test database consisting of 4 data sets which cover geological data sets different in content and structure. 5.1. A1&B1 (e.g., BSsp). 5.2. A1&B2(e.g., BSgrnsp). 5.3. A2&B1(e.g., BSch). 5.4. A2&B2(e.g., BSgrnsp).

Data sets	Sampling factors	Laboratory studies	Data name
1	A1 one sample usually by sediment grab	B1 The sample was studied for one kind of analyze	BSsp
2	A1 one sample usually by sediment grab	B2 The Sample was studied for several kind of analyze	BSgrnsp
3	A2 few samples usually by vibrocorer	B1 The sample was studied for one kind of analyze	BSch
4	A2 few samples usually by vibrocorer	B2 The Sample was studied for several kind of analyze	BSgrnsp

ODV main principles

One of the principles of putting geological data in ODV format is, it is more convenient to separate in-situ measurements and ex-situ measurements (such as laboratory analyses) into separate ODV files. So the solution would be to have only one CDI file, describing the *field* measurement, and two ODV files: one describing the lithology in the field and the other one containing the parameters measured in the lab. These two files have the same structure but different content. These files will be zipped and this will be the data file that is listed in coupling.txt.

ODV classification

Data sets	Sampling factors	Laboratory studies	ODV type
BSgrn12	A1	B1	**ODV1** on the station was selected one sample which studied using a single analysis
BSgrnsp506	A1	B2	**ODV2** on the station was selected one sample which studied using two analysis
BSch993, BSsp3 BSmin25, ptr1341	A2	B3	**ODV3** on the station was selected several samples which studied using one analysis

In conclusion it should be noted that the proposed classification is deliberately allowed to approach the method of preparation CDI and ODV during the Geo-Seas project.

Towards an integrated EU data system within AtlantOS

Valérie Harscoat, Institut français de recherche pour l'exploitation de la mer (France), Valerie.Harscoat@ifremer.fr

Sylvie Pouliquen, Institut français de recherche pour l'exploitation de la mer (France), Sylvie.Pouliquen@ifremer.fr)

AtlantOS WP7 partners, Institut français de recherche pour l'exploitation de la mer (France), atlantos_wp7_coordination@ifremer.fr

The H2020 AtlantOS project started in June 2015 and aims to optimise and enhance the Integrated Atlantic Ocean Observing Systems (IAOOS). One goal is to ensure that data from different and diverse in-situ observing networks are readily accessible and useable to the wider community, international ocean science community and other stakeholders in this field. To achieve that, the strategy is to move towards an integrated data system within AtlantOS that harmonises work flows, data processing and distribution across the in-situ observing network systems, and integrates in-situ observations in existing European and international data infrastructures (Copernicus marine service, SeaDataNet NODCs, EMODnet, OBIS, GEOSS) so called Integrators.

The targeted integrated system will deal with data management challenges for efficient and reliable data service to users:

- Quality control commons for heterogeneous and nearly real time data
- Standardisation of mandatory metadata for efficient data exchange
- Interoperability of network and integrator data management systems

Presently the situation is that the data acquired by the different in situ observing networks contributing to the AtlantOS project are processed and distributed using different methodologies and means. Depending on the network data management organization, the data are either processed following recommendations elaborated y the network teams and accessible through a unique portal (FTP or Web), or are processed by individual scientific researchers and made available through National Data Centres or directly at institution level. Some datasets are available through Integrators, such as Copernicus or EMODnet, but connected through ad-hoc links.

To facilitate the access to the Atlantic observations and avoid "mixing pears with apples", it has been necessary to agree on (1) the EOVs list and definition across the Networks, (2) a minimum set of common vocabularies for metadata and data description to be used by all the Networks, and (3) a minimum level of Near Real Time Quality Control Procedures for selected EOVs. Then a data exchange backbone has been defined and is being setting up to facilitate discovery, viewing and downloading by the users. Some tools will be recommended to help Network plugging their data on this backbone and facilitate integration in the Integrators. Finally, existing services to the users for data discovery, viewing and downloading will be enhanced to ease access to existing observations.

An initial working phase relying on existing international standards and protocols, involving data providers, both Networks and Integrators, and dealing with data harmonisation and integration objectives, has led to agreements and recommendations on:

- a list of EOVs across the Networks
- a minimum set of metadata common vocabularies to be used by all networks
- a minimum level of Near Real Time Quality Control Procedures for selected EOVs (T, S , Current, O2, Chl, Nitrate, Sea Level, Carbon)
- basic services (discovery, viewing and downloading) to distribute the data

The setup phase has started, both on Networks and Integrators sides, to adapt the existing systems in order to move toward this integrated EU data system within AtlantOS.

W.A.V.E.S – Web Accessible Visualization and Extraction System for GLODAPv2 Database

Alex Kozyr, Oak Ridge National Laboratory (United States of America), kozyra@ornl.gov
Misha Krassovski, Oak Ridge National Laboratory (United States of America), krassovskimb@ornl.gov

The Global Ocean Data Analysis Project (GLODAP) is a cooperative effort of investigators funded for ocean synthesis and modeling projects by the U.S. National Oceanic and Atmospheric Administration (NOAA), Department of Energy (DOE), and National Science Foundation (NSF).

Cruises conducted as part of the WOCE, JGOFS, and NOAA Ocean-Atmosphere Carbon Exchange Study (OACES) over the decade of the 1990s generated oceanographic data of unparalleled quality and quantity. GLODAPv2 is a uniformly calibrated open-ocean data product containing inorganic carbon and carbon-relevant variables. This new product includes data from approximately one million individual seawater samples collected from over 700 cruises during the period 1972-2013. Extensive quality control and subsequent calibration were carried out for salinity, oxygen, nutrient, carbon dioxide, total alkalinity, pH, and chlorofluorocarbon data. The Carbon Dioxide Information and Analysis Center (CDIAC), serving as the primary DOE disseminator for climate data and information, developed database and web accessible systems that permit users worldwide to query and retrieve data from the GLODAPv2 collection. This presentation will showcase this new system, discuss technologies used to build the GLODAPv2 resource, and describe integration with a metadata search engine provided by CDIAC as well.

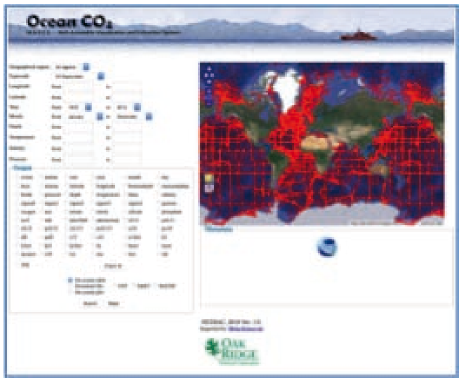


Fig. 1 - GLODAPv2 WAVES Interface.

Perform a search of the database by:

1. Geographical region
2. Expocode
3. Coordinates
4. Date
5. Depth, Temperature, Salinity, Pressure
6. Measurement parameter(s)

Output in the following formats:

On-screen table
Downloadable file (CSV, TabSV, NetCDF)
On-screen plot.

The ICOS OTC Data Lifecycle Plan

Camilla Stegen Landa, University of Bergen and Bjerknes Centre for Climate Research (Norway), Camilla.Landa@uib.no
Steve Jones, University of Exeter (United Kingdom), s.d.jones@exeter.ac.uk
Benjamin Pfeil, University of Bergen and Bjerknes Centre for Climate Research (Norway), Benjamin.Pfeil@uib.no
Truls Johannessen, University of Bergen and Bjerknes Centre for Climate Research (Norway), Truls.Johannessen@uib.no

Introduction

The Integrated Carbon Observation System (ICOS) is a Pan-European research infrastructure aiming to provide the long-term observations required to understand the present state and predict future behavior of the global carbon cycle and greenhouse gas emissions. To facilitate the research that will lead to this understanding, ICOS will offer a high precision, long-term observing network of stations across Europe and adjacent regions, which measures greenhouse gas fluxes from ecosystems and oceans, and greenhouse gas concentrations in the atmosphere. The specific tasks of collecting and processing data measured at national network stations are divided among the Ocean Thematic Centre (OTC), the Ecosystem Thematic Centre (ETC) and the Atmosphere Thematic Centre (ATC). All quality controlled data will be made available through the Carbon Portal (CP).

The distributed, interdisciplinary and complex nature of this infrastructure enforces that a data lifecycle plan is needed in order to be able to make the data accessible for current and future climate change research. The data lifecycle plan is an essential part of data management plans for European Infrastructure projects, and becomes increasingly important for projects dealing with large data volumes and quality control performed by various entities. The data lifecycle plan documents in detail all changes applied to a dataset from its collection at the sensor until its final publication, and allows a transparent tracking of all quality control procedures applied. Responsibilities should also be assigned in the plan.

Data definitions, transfers and responsibilities have been discussed within ICOS for several years and the data lifecycle plan is aligned with recommendations of the *Common Operations of Environmental Research Infrastructures Reference Model (ENVRI RM)*. ENVRI RM is a common ontological framework and standard for the description and characterization of environmental research infrastructures. The ENVRI RM identifies a set of common functionalities in research infrastructures and provides solutions to common problems. The concepts and frameworks of the ENVRI RM were used to examine the requirements and optimize the design of the ICOS data infrastructure.

Method

Our focus is on the data lifecycle plan for the ICOS OTC. The OTC is hosted by Norway and the United Kingdom and is responsible for coordinating the marine network of ICOS. The marine network consists of instrumented Ships of Opportunity/Voluntary Observation Ships (SOOP/VOS), fixed ocean time series stations and repeat hydrography (Fig. 1). SOOP/VOS and fixed time series stations are equipped with a suite of automated instrumentation to measure atmospheric and surface ocean pCO2, sea surface temperature, salinity and related variables. During repeat hydrographic cruises discrete samples of total alkalinity, total carbon, pH, nutrients and related variables are taken. Approximately 50 different lines and stations operated from 12 countries are planned for the ocean network.

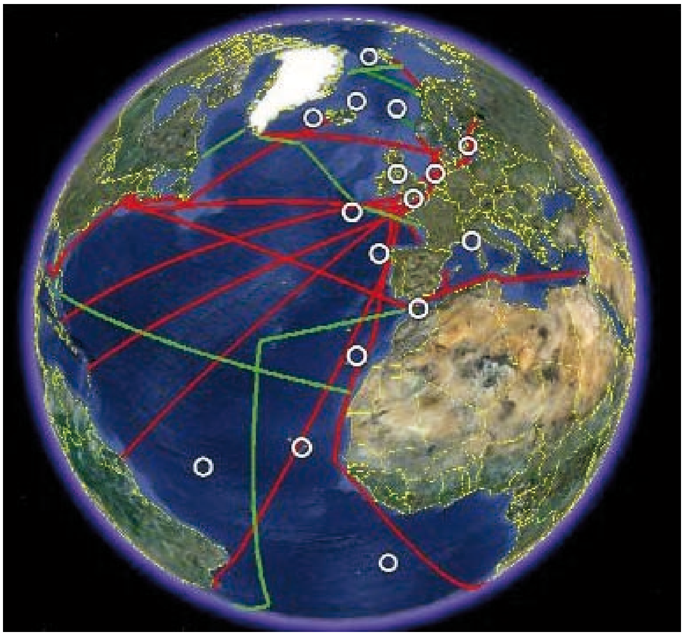


Fig. 1 - The suggested network of stations for the ocean network: Fixed Ocean Stations (cycles), Ships of Opportunity/Voluntary Observation Ships (red lines) and Repeat Sections (green lines).

The data lifecycle within OTC is of complex nature with many steps and transmissions of data and metadata between the OTC, the CP and the principal investigators. However, the data lifecycle can be boiled down to data acquisition, data curation and data access. Within these basic pillars are many steps which include different levels of quality checks and assurance, version control, data archiving, assigning of persistent digital identifiers, and data publications. The OTC will aim to follow international procedures and best practices for the above steps.

An automated system for data submission and quality control will be established for the automated data streams from the marine network of ICOS. This system will be based upon the automated data ingestion system as developed within the marine biogeochemistry community for the Surface Ocean CO2 Atlas (SOCAT) project (www.socat.info), which contains amongst others user friendly quality control tools, automated range checkers and visualization options. These features will be optimized to OTC's needs. The datasets submitted to the automated ingestion system for SOCAT are quality controlled by the principal investigators, and in many cases made public, prior to submission. OTC plan to extend and optimize this infrastructure and speed up the process of data release. All tool-kits developed will be made available to the international marine community with the aim that raw data is streamed directly from the instruments into automation systems.

All changes applied to a dataset will be logged and archived using a version control system so that older versions can be restored if needed. In addition, different data levels have been defined, starting from raw data and ending with the published dataset (or data product) which is the final data level. The data levels in between differ in the amount of processing and quality control applied. Users can get access to data from different levels depending on their needs, which are often a counterbalance between the data quality needed and the freshness of data. It will be possible for users to get access to near real-time data, final quality controlled ICOS datasets, and data products where ICOS data are included. Persistent digital identifiers (DOIs) will be assigned for data citation to ensure that data sources, providers and funding sources are being acknowledged.

Standardized vocabularies from the *Natural Environmental Research Council* (NERC) vocabulary server will be implemented for the OTC data. This vocabulary server was developed by the NERC DataGrid program as part of an effort to support uniform data and metadata discovery and access. Standardized vocabularies will ease the use of the data by the modeling community and exchange of data with different communities, and ensure interoperability with existing systems (e.g. SeaDataNet and EmodNet).

Integrating data management services in marine sciences

Lisa Paglialonga, GEOMAR Helmholtz Centre for Ocean Research (Germany), lpaglialonga@geomar.de
Hela Mehrtens, GEOMAR Helmholtz Centre for Ocean Research (Germany), hmehrtens@geomar.de
Pina Springer, GEOMAR Helmholtz Centre for Ocean Research (Germany), pspringer@geomar.de
Carsten Schirnick, GEOMAR Helmholtz Centre for Ocean Research (Germany), cschirnick@geomar.de
Claas Faber, GEOMAR Helmholtz Centre for Ocean Research (Germany), cfaber@geomar.de

Organisation

The joint GEOMAR Data Management Team is a cooperation of Helmholtz Centre for Ocean Research Kiel (GEOMAR) and several large-scale research projects including collaboration with other marine research institutions. The coalition has established in Kiel a single sustainable data management infrastructure (KDMI) for marine research and supports the entire lifecycle of data description, data storage and data archiving. The platform is used to prepare data for peer-reviewed publications, exchange data within a project and for publication in longterm data archives. It is continuously improved by extensions developed in close cooperation with scientists and data centres.

Data management services

KDMI provides a web-based portal including OSIS (Ocean Science Information System) https://portal.geomar.de/kdmi which allows scientists to upload and share data files in the context of cruises, expeditions or models with allowance for any file format and content structure (Fig. 1). Access control for a file is primarily based on the community context it was uploaded within, but may be further restricted by the file's owner e.g. for a limited time of exclusive research (e.g. moratorium). Metadata of what, when, where and who are always public for institutional use and document the file's content. Visualisation of metadata is provided by KML formatted files which contain the appropriate links in order to retrieve and display e.g. the events of a cruise and availability of files or publications on-the-fly in tools such as GoogleEarth.

Fig. 1 - Ocean Science Information System (OSIS).

A new and now widely appreciated service has been established realizing data management plan time schedules. It allows setup and surveillance of deliverables based on expected or collected research data. OSIS now also includes a numerical model database which provides an overview of existing model runs linked to their output results and to publications. Files (e.g. configuration files) can be uploaded as well and aid scientists in documenting their procedures and their outcome enriched by a minimal set of common metadata. Model output referenced in peer-reviewed papers is made publicly available via GEOMAR's OPeNDAP Server https://data.geomar.de.

Another data management service is the media archive (ProxSys, commercial software) (Fig.2). It allows to store, describe and search images and videos acquired during research dives of remotely operated and autonomous underwater vehicles equipped with up to four high-resolution cameras (stills or movies). An annotation system is currently developed (Fig. 3) which uses webservices to access images and videos and provides them for all projects members. This system avoids extensive network traffic and facilitates cooperation and reduces obsolete interchange.

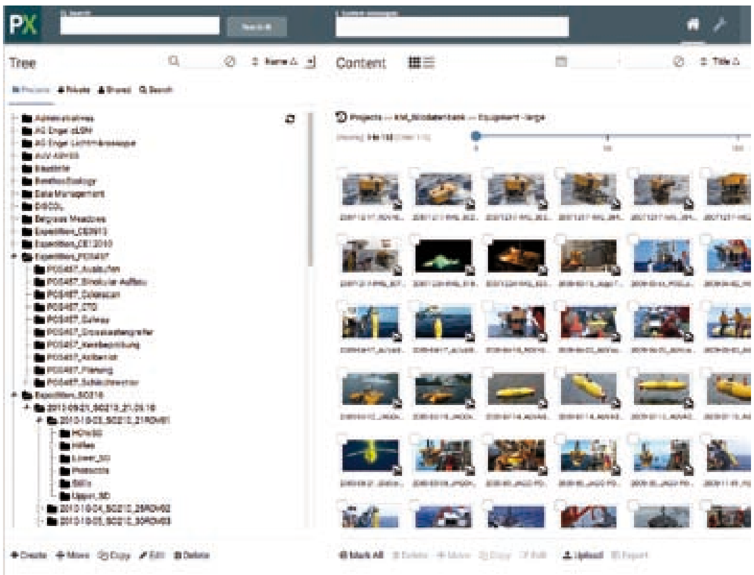


Fig. 2 - Media storage (ProxSys).

Fig. 3 - Annotation System.

We also offer version control systems (Subversion and Git), which are not only used for code development and documentation but also for data analysis and exchange. The TRAC Wiki and ticket system linked to Subversion and Git in the near future is used for documentation as well as improved collaboration and dissemination.

Integration

The data management portal (https://portal.geomar.de) is the starting point for all provided services. It allows to maintain public and internal webpages where collaboration tools like wiki, blog, calendar and document exchange can be used. User administration for all data management services is realised via the portal software and thus independent from institute's identity management systems while still allowing guarded access to data services for internal and external users.

An institutional repository for full text print publications is actively maintained by the GEOMAR library and also used for project publications (OceanRep: http://oceanrep.geomar.de/). The repository is linked with the data management portal allowing linkage between authors, their data and publications. Linkage between data and print publication is available in both systems: publications in the repository provide links to data and vice versa. The repository is also used to auto-generate publication lists on-the-fly and to present them on personal or a projects' homepages. An automated data link routine adds icons when additional external data are available related to a publication.

The data management team assists as data curators when data are to be published in World Data Centres (e.g. WDC-MARE, http://www.pangaea.de) in order to warrant long-term archival and access to the data (Fig.4). This cooperation with a world data center will make the data globally searchable while links to the data producers will ensure citability and provide points of contact for the scientific community. PANGAEA is harvested by several international portals, e.g. GEOSS (http://www.geoportal.org).

Fig. 4 - PANGAEA dataset.

The concept of large-scale projects joining a single approach to establish one data management system has proved to be very successful. We experience propagation especially at GEOMAR and Kiel University to use the services known from the collaboration for new expeditions and research projects. We also observe an ongoing need to train new and young scientists and to include data output of new instrumentations. One of the essential data management experiences is the need for on-site personal contact and training of researchers during the entire data lifecycle. The high efficiency of data publications related to peer-reviewed journal publications still relies on on-site data stewardship conducted by a dedicated permanent data management team.

Global Bathymetric Data Sets - General Bathymetric Chart of the Oceans (GEBCO)

Pauline Weatherall, British Oceanographic Data Centre (United Kingdom), paw@bodc.ac.uk

Karen M. Marks, Laboratory for Satellite Altimetry, National Oceanic and Atmospheric Administration (United States of America), karen.marks@noaa.gov

Martin Jakobsson, Department of Geological Sciences, Stockholm University (Sweden), martin.jakobsson@geo.su.se

Lesley Rickards, British Oceanographic Data Centre (United Kingdom), ljr@bodc.ac.uk

On behalf of GEBCO

The GEBCO community (www.gebco.net) consists of an international group of experts in seafloor mapping who develop and make available a range of data sets and data products with the aim of providing the most authoritative publicly-available bathymetric data sets for the world's oceans. GEBCO operates under the joint auspices of the International Hydrographic Organization (IHO) and Intergovernmental Oceanographic Commission (IOC) of UNESCO.

Global gridded bathymetric data set – the GEBCO_2014 grid

GEBCO's latest bathymetric product is the GEBCO_2014 grid. This is a global terrain model at 30 arc-second intervals and was released in December 2014.

The grid is largely based on a database of quality-controlled ship-track soundings with interpolation between soundings guided by satellite-derived gravity data. Where they improve on this model, data sets generated by other methods have been included.

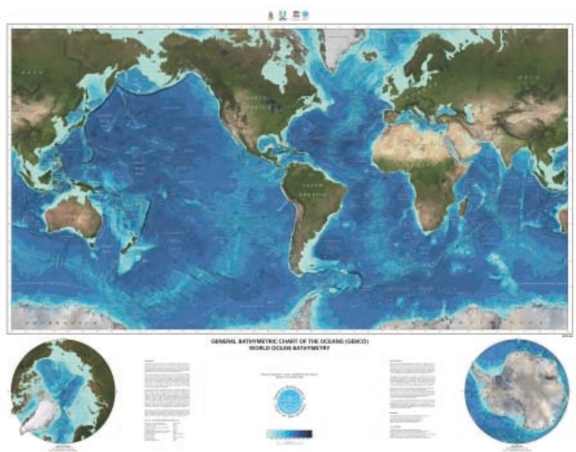


Fig. 1 - GEBCO world map.

Recognising the importance of regional mapping expertise to help improve its global bathymetric model, GEBCO is building collaborations with regional mapping efforts to help encourage and coordinate the incorporation of their compilations into GEBCO.

The GEBCO_2014 grid includes contributions from many regional mapping projects and data contributors such as the International Bathymetric Charts of the Arctic Ocean (IBCAO) and Southern Ocean (IBCSO); the Baltic Sea Bathymetry Database; EMODnet 2013 for European waters; Australian Bathymetry and Topography Grid, June 2009; Japan Coast Guard Grid for part of the North Western Pacific Ocean region and data from the Global Multi-Resolution Topography

(GMRT) synthesis, provided by the Lamont-Doherty Earth Observatory of Columbia University, USA.

The GEBCO_2014 grid is accompanied by a Source Identifier (SID) Grid showing which cells are based on soundings or pre-gridded data sets and which are interpolated.

GEBCO's grids can be download from the internet in netCDF (CF Compliant), Esri ASCII raster or data GeoTiff formats : www.gebco.net/data_and_products/gridded_bathymetry_data/.

Gazetteer of Undersea Feature Names

The GEBCO Sub-Committe on Undersea Feature Names (SCUFN) maintains and makes available a gazetteer giving the name, generic type, geographic location and extent of features on the seafloor.

The data set contains over 3,900 features and is available to view, search and download via the internet in a number of formats including comma separated text, spreadsheet and shapefile: www.gebco.net/data_and_products/undersea_feature_names/.

Via its web pages, SCUFN provides information on undersea feature terms and definitions. It also provides details on how to propose names for newly-discovered features.

Web services

GEBCO make available a Web Map Service (WMS) providing imagery based on the GEBCO 30 arc-second grid. The WMS includes layers showing global bathymetry in the form of a shaded relief image and as a 'flat' map colour-coded for depth. It also includes a layer showing the SID grid coverage - i.e. identifying those grid cells in the GEBCO_2014 grid constrained by measured data or pre-gridded data sets, or based on interpolation.

The WMS can be accessed from GEBCO's web site:

www.gebco.net/data_and_products/gebco_web_services/web_map_service/

IHO-IOC GEBCO Cook Book

The Cook Book is a technical reference manual on how to build bathymetric grids. It includes information on: data gathering and cleaning; producing bathymetric grids plus information on some of the software packages available to do this work.

www.gebco.net/data_and_products/gebco_cook_book/

GEBCO world map

This is a shaded-relief colour map image showing the bathymetry of the world's oceans based on GEBCO's global gridded bathymetric data set.

www.gebco.net/data_and_products/printable_maps/gebco_world_map/.

XBT Data Management and Quality Control in Japan (II) Improving Database by Historical XBT System

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

We have started to reassemble historical expendable bathythermograph (XBT) data in order to improve an ocean subsurface database in the North Pacific. In the 1980s and early 1990s temperature observed by XBT were reported at the standard depths or inflection depths because temperature profile was recorded on strip chart (right panel of Fig. 1) so that temperature and depth were digitized by hand (middle panel of Fig. 1). We therefore discovered and collected

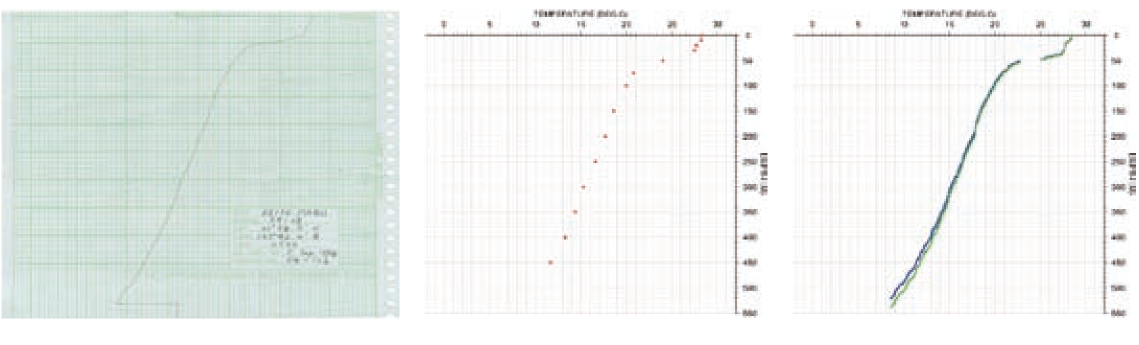


Fig. 1 - An example of XBT cast by R/V Keifu-Maru, JMA, in May 1984. Recording profile on strip chart (left), reading temperature at standard depths (center; stored in existing database), and traced profile in this study (right; bule line shows the observed depth derived from manufacture's fall rate equation, and green line shows the same but by Hanawa *et al*. (1995)).

about 3,300 of existing XBT strip charts of T-4 probe manufactured by Tsurumi-Seiki Co. Ltd. (TSK) in Japan Meteorological Agency and Japan Hydrographic and Oceanographic Department, and traced a recorded temperature profile at a higher resolution for every chart. The traced data are approximated by a function of elapse time and temperature so that we can convert to temperature profiles using fall rate equations which are provided by XBT manufacture, Hanawa et al. (1995) or others (left panel of Fig. 1). These temperature profiles can also compare with other instruments such as CTD or Argo floats at the same or neighborhood time and position in order to estimate their systematic errors or uncertainties. The related information such as probe type and manufacture, fall rate equation, type of recorder or converter, launch height on shipboard are also included in database as metadata. The improved database will be used to assess the climate change and the sea level rise in the North Pacific. Furthermore we also discovered about 700 of XBT strip charts by Fuji, antarctic research vesel, in the Southern Ocean (Figs. 2 and 3). These profiles obtained by T-5 probe by TSK, and they are not stored in Japan Oceanographic Data Center. Some traced profiles include errors such as spike or measurement failuer caused by noise, broken wire, bottom grounding or others, therefore automatic and expert quality control procedures develoed by International Quality-controlled Ocean Database (IQuOD) project will be adopted and imported to existing database.

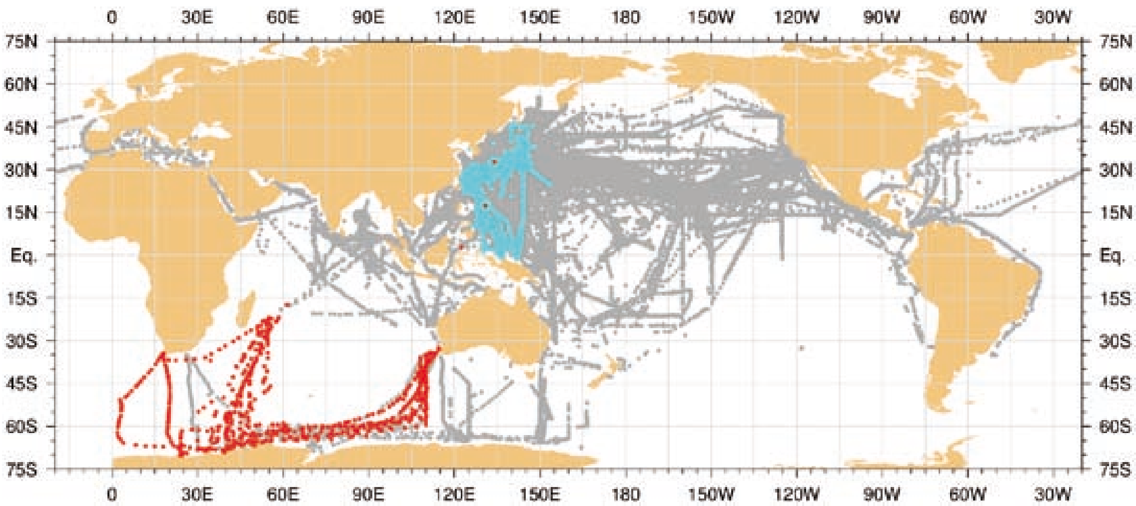


Fig. 2 - Station map of all XBT casts archived in Japan Oceanographic Data Center (gray), and traced stations measured by research/survey vessels of Japan Meteorological Agency and Japan Hydrographic and Oceanographic Department (cyan), and new discovered XBT casts by Fuji (red) in this study.

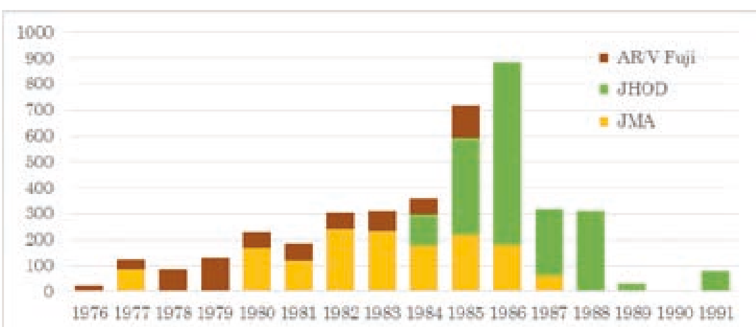


Fig. 3 - Annual change of traced historical XBT profiles in this study.

MarineMammals.be: improving marine mammal stranding, by catch and observation data collection

Thomas Vandenberghe, Royal Institute of Natural Sciences (Belgium), tvandenberghe@naturalsciences.be
Steven Degraer, Royal Institute of Natural Sciences (Belgium), sdegraer@naturalsciences.be
Thierry Jauniaux, Université de Liège (Belgium), t.jauniaux@ulg.ac.be
Francis Kerckhof, Royal Institute of Natural Sciences (Belgium), fkerckhof@naturalsciences.be
Bob Rumes, Royal Institute of Natural Sciences (Belgium), brumes@naturalsciences.be
Serge Scory, Royal Institute of Natural Sciences (Belgium), sscory@naturalsciences.be
Jan Haelters, Royal Institute of Natural Sciences (Belgium), jhaelters@naturalsciences.be
Elise Toussaint, Royal Institute of Natural Sciences (Belgium), etoussaint@naturalsciences.be

Introduction

The marine mammals' website (www.marinemammals.be) is the result of a long-standing collaboration between the Royal Belgian Institute of Natural Sciences (RBINS) and the University of Liège (ULg) and provides access to data of 50 years of marine mammal strandings in Belgium and Northern France. The site consists of two main parts, i.e. 'records of strandings and observations' and the Belgian Marine Mammal Biobank, a tissue bank (BMMB or just Biobank). While stranding and observation records can be publicly consulted, the Biobank is focused on the scientific community. Overall, the website aids in collecting data on strandings, necropsies, ad hoc observations and dedicated surveys of marine mammals (pinnipeds and cetaceans) in Belgian waters. It allows to gather successive observations of an individual animal, includes the probable cause of death of stranded animals, and facilitates the data flow to (international) fora dealing with marine mammals, nature conservation or fisheries.

Ad hoc stranding and observation data are entered into the database in a standardised way by RBINS staff. In order to support and refine the cause of death as assessed on the beach, necropsy data are included in the database. The database includes the tissue samples taken, together with their way of preservation, and their location; currently the tissue bank contains over 20000 samples. The target public of the BMMB is a wide range of scientists: biologists, veterinarians, toxicologists, microbiologists, pathologists... It holds samples from the liver, blubber, lung... but also stomach contents, lesions (pneumonia, encephalitis, ...) and pathogens (parasites, virus, bacteria,…). They originate from animals stranded or bycaught in the Belgian maritime zone, but also include selected tissues of marine mammals stranded in Northern France, the Netherlands and Ireland. In the near future it will be possible to request samples for specific research objectives with an account.

Observations and strandings

Care has been taken to remove subjectivity in categorical variables, notably the observation type and the parameters. Observation type is a general indicator how the animal presented itself to the initial observer. The following values are available: Bycatch, Caught, Caught Inland, Found dead in Harbour, Died during transport/rehab, Released from captivity, Sighted, Found dead on beach, Found alive on beach/harbour, Found dead at Sea.

Additionally there are parameters that provide information on the observation and the specimen. Observation parameters provide details on wind and sea state. Specimen parameters contain circumstantial parameters, measurements, external examination observations and cause of death. Circumstantial parameters describe what happened with the specimen before, during and after the intervention. Possible values are:

Before intervention: alive, died at sea, died on beach, killed on beach, died, circumstances unknown, animal back to sea on its own.

During intervention: intervention unknown, no intervention, escaped while trying to catch, released alive, died during intervention/rehab same day, taken to rehab, euthanized.

Collection: live animal sampled, nothing collected, fully collected for necropsy, sampled, then released for destruction, sampled, then left at location, necropsy at location, released for destruction, carcass disappeared.

Observation creation and modification form

For an observation, 5 levels of detail are available: observation (time, location, textual details, and observation parameters), specimen (species, sex, number, circumstantial parameters and measurements...), external examination, cause of death and sources and media.

The cause of death can be given by the probability (definitively, probably, possible, not, unknown) of 5 possible causes of death (Natural, Bycatch, Ship strike, Predation and other).

Necropsy form

For the dead animals that were collected for necropsy, or were necropsied at the stranding site, the necropsy data are recorded. They include parameters such as weight, age, nutritional status, parasite burden, and, if possible, a cause of death. An overview of microscopic and macroscopic lesions and tissue samples taken is added.

Dealing with (historical) data and making it accessible: Data Inventory and Tracking System (DITS) applied in the scope of the *"4 decades of Belgian marine monitoring"* project (4DEMON)

Marielle Adam, Royal Belgian Institute of Natural Sciences (Belgium), madam@naturalsciences.be
Ruth Lagring, Royal Belgian Institute of Natural Sciences (Belgium), ruth.lagring@naturalsciences.be
Yvan Stojanov, Royal Belgian Institute of Natural Sciences (Belgium), ystojanov@naturalsciences.be
Francis Strobbe, Royal Belgian Institute of Natural Sciences (Belgium), fstrobbe@naturalsciences.be
Thomas Vandenberghe, Royal Belgian Institute of Natural Sciences (Belgium), tvandenberghe@naturalsciences.be

Long-term datasets are important to study global trends in the environment and to define a reference status of an ecosystem. Results of analyses on such datasets allow policy makers to define proper strategies to reach regulatory objectives. Access to these data is also a crucial issue for the scientific community.

Belgium has a long oceanographic data collection history. In that context, the Belgian Marine Data Center, with its continuously growing experience in data rescuing and management, worked out a tool to identify the marine datasets and retrieve as much data as possible.

"4 decades of Belgian marine monitoring: Uplifting historical data to today's needs" (4DEMON)

The tool was developed in the scope of the project *"4 decades of Belgian marine monitoring: Uplifting historical data to today's needs"* (4DEMON). This project aims to build quality-checked, intercalibrated and integrated data on contamination, eutrophication and ocean acidification in the Belgian Continental Shelf (BCS) since 1970. Historical data is being compiled and integrated with more recent data resulting in robust long-term datasets that are made publically available. This will not only safeguard the historic Belgian marine data, but also uplift its value as it can serve many new and cross-disciplinary research objectives and help policy makers to take adequate management actions.

DITS in practice: from data inventory to database and data dissemination

Inherent to these kind of projects on historical data, there are several issues, such as complexity and heterogeneity of the data, potential missing metadata, numerous data formats and data sources. To face these matters, a well-coordinated system to track and to disseminate data needed therefore to be set up.

Accordingly, the "Data Inventory and Tracking System" (DITS) was developed for inventory of (historic) files or datasources like technical reports, logbooks and scientific publications compiled during the project. The submitted datasources, which can be fully described with all required metadata (like project, data type and data originator), can be combined into 'ingestion datasets' or logical groups of datasources. Original and processing files can be uploaded or refered to and each step of the process can be tracked (Fig. 1). Furthermore, missing information like sampling date or location can be indicated. The tool was designed by and for datamanagers to process data that need further handling, like import in central databases. Data submitters on the other hand can follow-up the status of their data and make updates via the online interface where, all records can be explored and exported.

While DITS serves as a shared tool for data management and traceability, it also allows to centralise, integrate and valorise data compiled in the frame of specific research projects and monitoring programs, forming an important reference for data managers, policy makers and scientists.

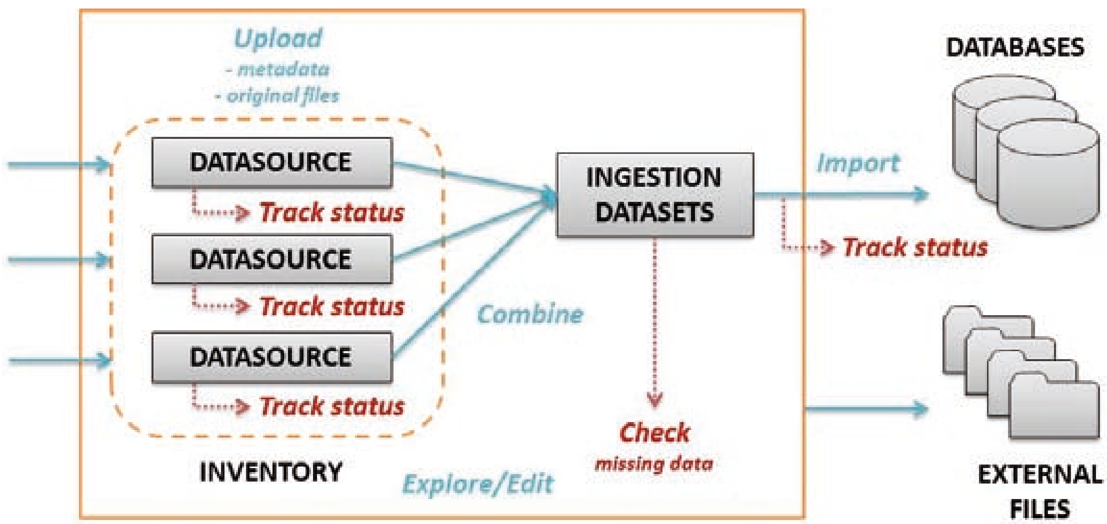


Fig. 1 - Data flow.

Analysis of ocean in situ observations and marine data products: Web-visualization and Services for EMODnet Chemistry

Giorgio Santinelli, Deltares (Netherlands), giorgio.santinelli@deltares.nl
Gerrit Hendriksen, Deltares (Netherlands), gerrit.hendriksen@deltares.nl
Alexander Barth, University of Liege (Belgium), a.barth@ulg.ac.be

Introduction

For the EMODnet chemistry dynamic visualisations and plotting, observations of chemical species are made available via state of the art techniques using globally accepted web services (OGC services). This has been done by using functions and scripts, reading so-called enriched ODV files and inserting data directly into a relational geo-geodatabase in the cloud. The main table is the one with observations which contains the main data and meta-data associated with the enriched ODV files. A particular implementation in data loading is used in order to improve on-the-fly computational speed.

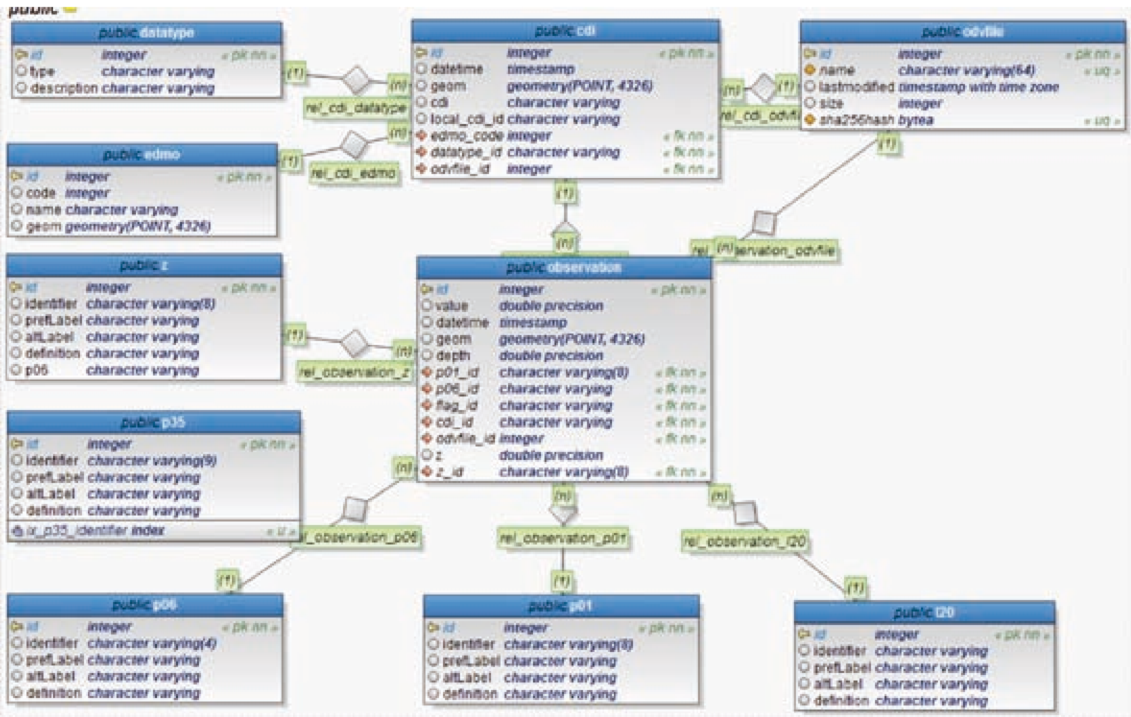


Fig. 1 - ODV database schema.

EMODnet database and Web Services

Data from Baltic Sea, North Sea, Mediterrean, Black Sea and part of the Atlantic region has been entered into the geodatabase, and consequently being instantly available from the OceanBrowser EMODnet portal. Furthermore, Deltares has developed an application that provides additional visualisation services for the aggregated and validated data collections. The visualisations are produced by making use of part of the OpenEarthTool stack (http://www.openearth.eu), by the integration of Web Feature Services (OGC) and by the implementation of Web Processing Services. The goal is the generation of server-side plots of timeseries, profiles, timeprofiles and maps of selected parameters from data sets of selected stations.

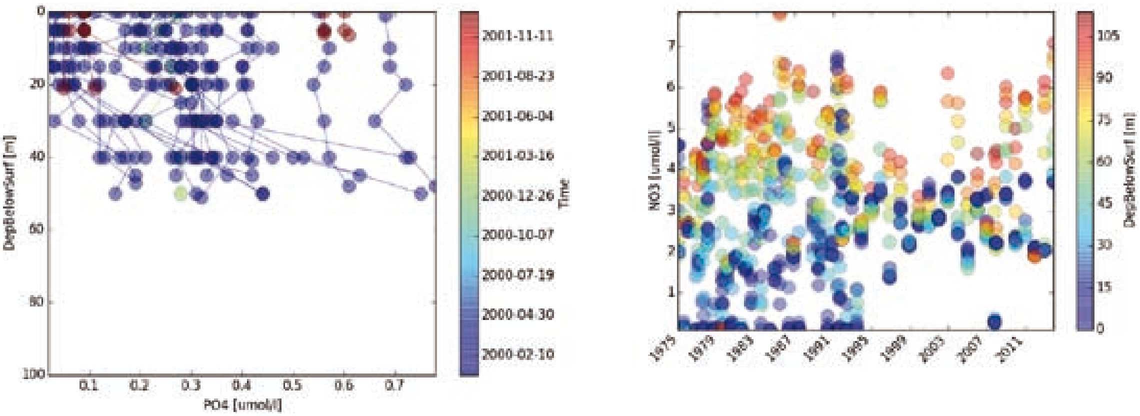


Fig. 2 - Dynamic visualization: profiles and timeseries.

Regional data collections are retrieved using EMODnet Chemistry cloud relational geo-database. The spatial resolution in time and the intensity of data availability for selected parameters is shown using Web Service requests via the OceanBrowser EMODnet Web portal. Selections in Oceanbrowser also result as a list of hypertext links for further data shopping and download.

OceanBrowser

The sparsity of observations poses a challenge common to various ocean science disciplines. Even for physical parameters where the spatial and temporal coverage is higher, current observational networks undersample a broad spectrum of scales. The situation is generally more severe for chemical and biological parameters because related sensors are less widely deployed.

OceanBrowser is a Web-interface to visualize gridded data sets in NetCDF, explore horizontal and vertical sections, scalar and vector fields. It is used in SeaDataNet and EMODnel Chemistry.

The analysis tool DIVA (Data-Interpolating Variational Analysis) is designed to generate gridded fields from in situ observations. DIVA has been applied to various physical (temperature and salinity), chemical (concentration of nitrate, nitrite and phosphate) and biological parameters (abundance of a species) in the context of different European projects. The in situ observation visualization service allows one to display vertical profiles and time series and it is built upon OGC

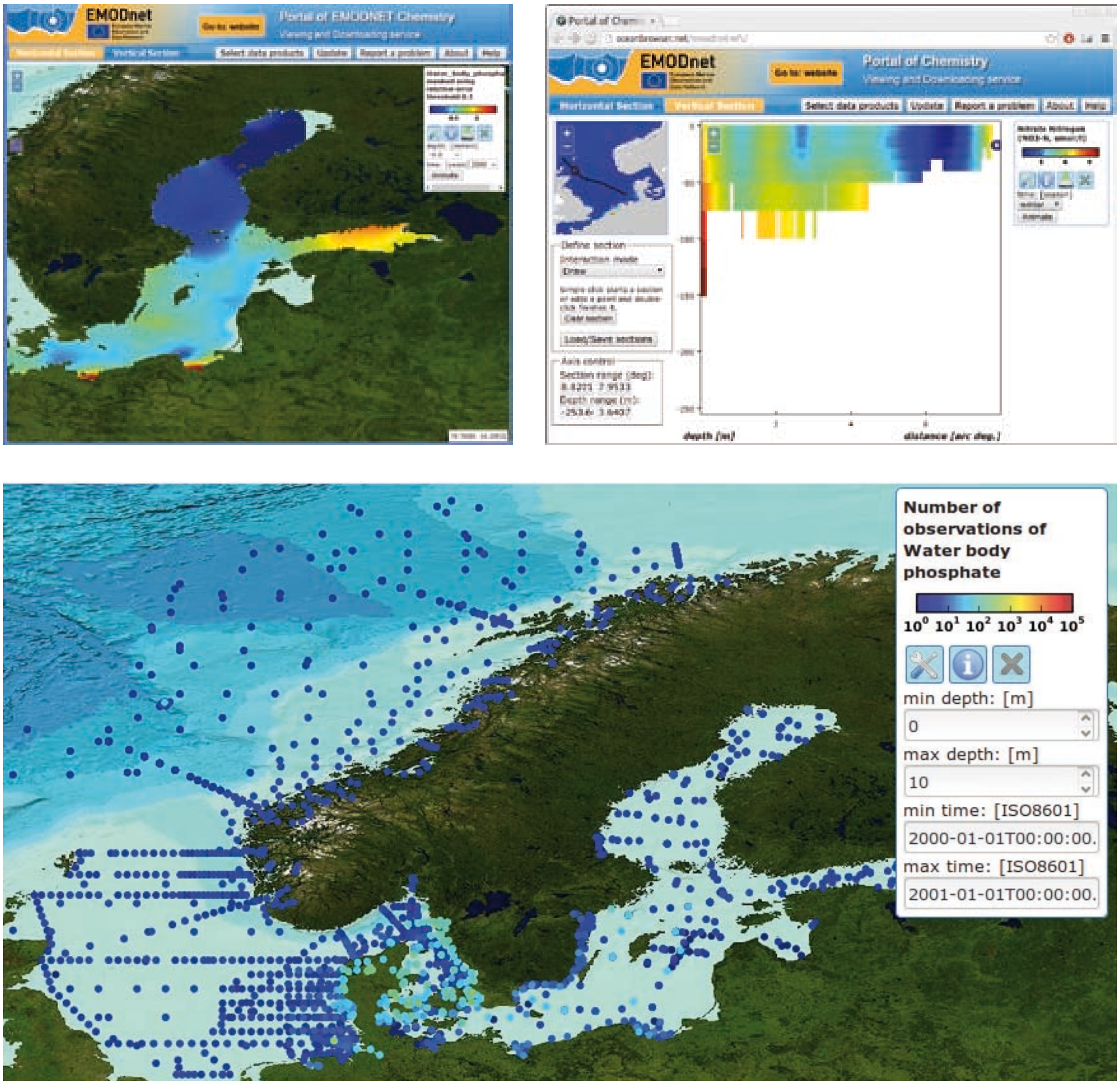


Fig. 3 - Snapshots fromOceanBrowser http://ec.oceanbrowser.net/emodnet/.

standards (the Web Feature Service and Web Processing Services) and following recommendation from the INSPIRE directive. In those projects it is used to visualize gridded data sets generated by the tool DIVA.

The ugly duckling that turned into a beautiful swan (ICES Stock Assessment Database)

Carlos Pinto, International Council for the Exploration of the Sea (Denmark), carlos@ices.dk
Inigo Martinez, International Council for the Exploration of the Sea (Denmark), inigo@ices.dk
Barbara Schoute, North Western Water Advisory Council (Ireland), schoute@bim.ie
Neil Holdsworth, International Council for the Exploration of the Sea (Denmark), NeilH@ices.dk
Cristina Morgado, International Council for the Exploration of the Sea (Denmark), Cristina.morgado@ices.dk

Introduction

ICES provides scientific analytical assessments and advice on fish stocks to a range of management partners, including the EU Commission and NEAFC. The stock assessment results, and their presentation are a key data output, and ICES has a stock assessment system (http://stock-assessment-graphs.ices.dk) that underpins its advice. This system consists of a database, and a web tool to produce charts, data tables and web services. This system ensures a transparent and consistent version controlled presentation of assessment results, it also allows end users to gain a better overview of the stocks assessed by ICES through one interface. The system started off as an Excel spreadsheet with macros for charting, and has become an online content management system for stock assessment results that form the evidence base of the ICES advice.

Historical development

In 2013 ICES developed a new Stock Assessment System that was able to produce standard charts of the annual fish stock assessments results, as the existing Excel based tool was prone to data and calculation errors, was difficult to share amongst multiple users and had no online presentation capability. The chartspresent the main assessment results including biological reference points and (where available) yield per recruit data. The system includes a relational database and a web based interface where assessment experts can upload assessment results (password protected), and adjust settings to be able to produce the charts that are included in the assessment advice sheets[1].

The system was very well received by the community of experts and end users, and added value in disseminating the results data which led to a change in the scope of the system, from producing charts to becoming the Stocks Assessment Database. This change in role led to substantial changes in the system.

What could be seen as a light change was in reality a concept change in the system, and led to iterative debugging and correction of the system and the database. The data uploaded prior to 2013 was not quality assured and there were few controls on data input. In the new system, the historic comparison charts showed clearly where old base data were not correctly documented (supported by metadata like units, survey or area). Overall, several issues arose due to the lack

1 See Cod species example http://tinyurl.com/npu72s3.

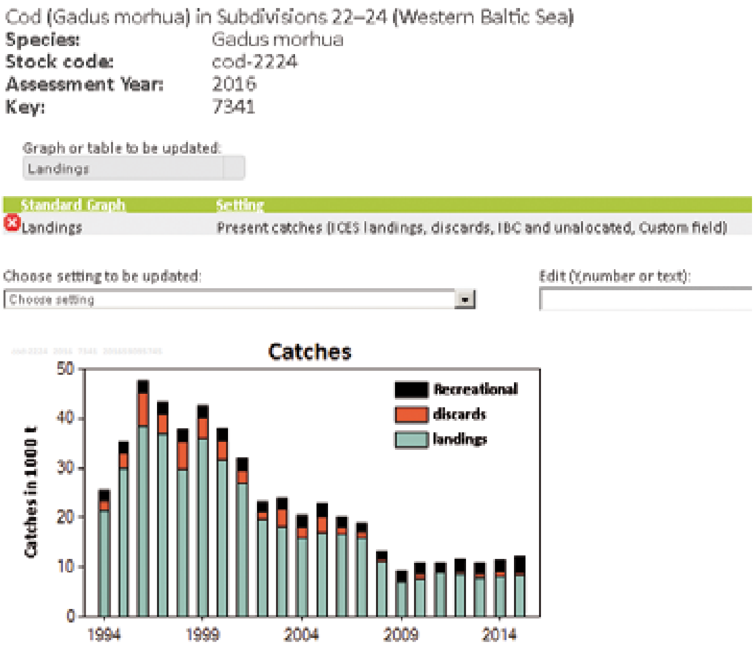


Fig. 1 - The user customizing the settings of the landings graph.

of metadata or presence of incorrect data. All these issues were solved while users continued to upload and disseminate data from the production database.

In 2015, with the system proving to be a success, the roll out to different types of assessments became the next logical development. This represented a big shift in the framework of the database. The system went from handling mainly controlled and well-structured data, to stocks where most of the data time series are heterogeneous (custom units and custom headers). This was accommodated by building a new XML template to allow the user to upload their own custom time series and to give them a little more freedom to build their own charts. The database was extended to encompass more types of information, for instance a more flexible set up including survey data considered representative for stocks where no standard assessment method is used.

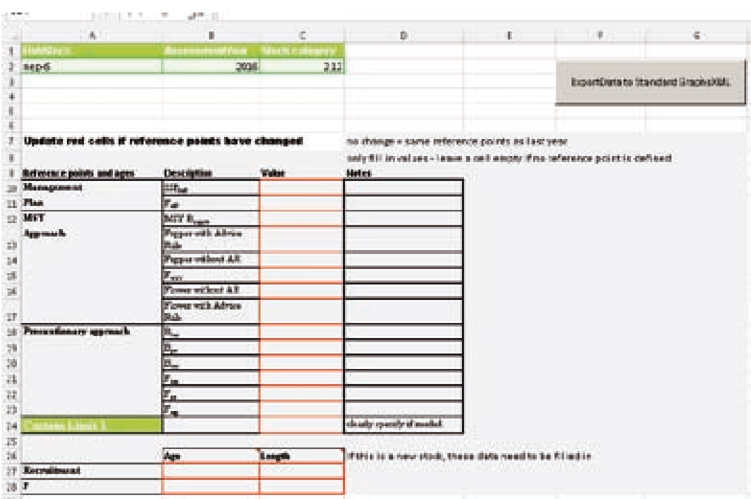


Fig. 2 - Stock Assessment Database Template.

Importing mechanism

The stock assessment database has its own standard reporting format, which is defined in XML and can be accessed by web interface or using a webservice. Users can follow the XML schema representing the stock assessment XML schema, or for users not comfortable with XML, there is an Excel template (Fig 2) with built in macros that exports the filled spreadsheet into the XML file following the XML Stock Assessment Database Schema (http://tinyurl.com/jczpvdd).

Outputs and visualization

The web interface for the visualization of the stock assessment charts data is based on three access levels:

1) Users without login can access the publicly available stocks.
2) Users with login can upload and adjust the settings of a stock.
3) ICES users can login, upload and manage the stocks (includes deleting stocks).

The system also offers another way of accessing the system, using webservices. The webservices (http://stock-assessment-graphs.ices.dk/webservices.aspx) can automatize most of the tasks in the stock assessment database. Users can upload and receive the data directly using the Simple Object Access Protocol (SOAP) standard. This has proven particularly useful because many users are using R scripting to manipulate the data, and can load the data directly via webservices.

ARCHIMEDE - A software for management of marine geological and biological samples and data

Catherine Borremans, Institut français de recherche pour l'exploitation de la mer (France), catherine.borremans@ifremer.fr

Vanessa Tosello, Institut français de recherche pour l'exploitation de la mer (France), vanessa.tosello@ifremer.fr

Ewan Pelleter, Institut français de recherche pour l'exploitation de la mer (France), ewan.pelleter@ifremer.fr

Jérôme Detoc, Institut français de recherche pour l'exploitation de la mer (France), jerome.detoc@ifremer.fr

Since 2003, IFREMER manages the marine geology database of Brest (BGMB) which includes information on geological samples (sediment cores and rocks) and geotechnical measurements collected by IFREMER teams and/or resources and related analyses (stratigraphy, lithology, chemistry, grain size). In the same way, the IFREMER Deep Sea Lab database (BIOCEAN) gathers data about biological, sediment and hydrological samples collected by IFREMER biologists and ecologists since 1967. These databases which include several thousands of operations and samples require having a better visibility for the scientific community through the improvement of existing software tools and services.

Indeed, according to the current scientific projects and future applications (e.g. management of mining licenses, implementation of the Nagoya Protocol, observatory of mineral raw materials, data dissemination), it becomes crucial to access to:

- the geographic position of sampling operations,
- the samples description (preservation, identifications,…) and related analyses,
- the samples physical location.

In this context, a project was initiated in order to:

- improve archiving and maintenance of marine samples collected during expensive marine cruises all over the world,
- make visible and accessible the available information (data and samples) to the whole scientific community,
- rationalise and optimise the logistic for cruise preparation for the French and international community.

As part of this, a study about management of marine geological and biological samples and related data in a unique system has been undertaken. The main objective is to save all information about the technical and scientific work during the operational phase in the field and in the lab. All metadata and data about handling and curating of sampled materials have to be collated as well. In that aim, new tools and services have been developed:

- A new database was designed by merging and adapting BGMB and BIOCEAN schemas. It allows management of multi-points operations, in-situ measures and observations. Different types and subsampling levels of samples (cores, fluids, organisms…) can be stored in relation with their physical place and transfers. The analyses on samples could also be managed following this new schema.

- The onboard software SEALOG allows recording and reporting events and operations and tracking samples from cruises, dives and moorings operated by biologists and geologists. Back on land, these data are loaded into the central database. However they must be completed after processing of samples and performing analyses.
- The graphical user interface connected to the new database, ARCHIMEDE, can be used in the lab, downstream the early acquisition at sea. This software was developed through four structuring functions: (1) Data input; (2) Visual display of data and query; (3) Data import; (4) Data export.
  So far these functionalities enable processing of information about surface operations, diving operations and samples (acquisition date-time, positions, equipment, originator, description…). The marine samples physical location and movements (transfers, loans…) can already be tracked with ARCHIMEDE whereas the features for the management of analyses on samples will be added by 2017. A module dedicated to geographical sites is planned to be integrated in a longer term.
- A website has been developped (soon available) in order to make visible and accessible to the whole scientific community all the information stored in the database. It is based on Sextant infrastructure (http://sextant.ifremer.fr/en/) which allows data display using the OGC Web Map Service protocol (WMS). Thanks to this website, scientists and general public will be able to locate geological and biological material from sea floor and archived in different French repositories. They will have access to related data (analyses) and images according to restriction rules.

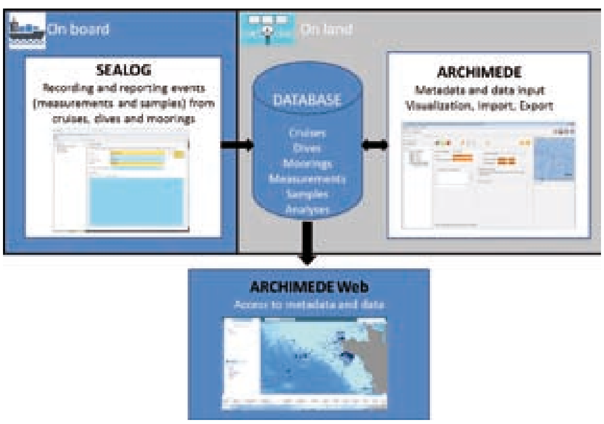


Fig. 1 - System for management of geological and biological samples and data.

This initiative was conducted through a collaborative relationship with field operators as well as laboratory staff in order to answer their needs and to offer user-oriented solutions.

Besides providing a way for scientists and data managers to enhance the currently existing data flows, this project aims to improve our capabilities in data and information security, management, use and delivery.

This is of vital importance for marine researchers, industry and policy-makers in the context of various studies, from the climate change prediction, marine resources assessment, to off shore engineering, and to make evidence-based decisions in France and globally.

The presentation will focus primarily on the software ARCHIMEDE, related developments and perspectives.

Biofloat: Python Software for Bio-Argo Float Data

Mike McCann, Monterey Bay Aquarium Research Institute (United States of America), mccann@mbari.org
Carole Sakamoto, Monterey Bay Aquarium Research Institute (United States of America), saca@mbari.org
Josh Plant, Monterey Bay Aquarium Research Institute (United States of America), jplant@mbari.org
Ken Johnson, Monterey Bay Aquarium Research Institute (United States of America), johnson@mbari.org

The Problem

Analyzing Bio-Argo profiling data from archive centers can be an onerous task. One must scan catalog lists for data assembly centers, then scan those sites for the NetCDF files and then read the data from many individual files before being able to perform any analysis or visualization. These difficulties hamper usefulness of the archive and limits access to only those familiar with NetCDF files and the ARGO data architecture.

A Solution

The Python programming language is gaining popularity as tool for analyzing large collections of data. The Scientific Python ecosystem (consisting of packages such as Numpy, Pandas, SciPy, Matplotlib, Xarray, and Jupyter) provides similar capabilites to commercial products such as

In [7]:
```
msdf = sdf.groupby(['wmo', 'year', 'month']).mean()

from biofloat.utils import o2sat, convert_to_mll
msdf['o2sat'] = 100 * (msdf.DOXY_ADJUSTED / o2sat(msdf.PSAL_ADJUSTED, msdf.TEMP_ADJUSTED))
msdf.head(10)
```

Out[7]:

			TEMP_ADJUSTED	PSAL_ADJUSTED	DOXY_ADJUSTED	lon	lat	o2sat
wmo	year	month						
1900650	2006	6	28.360667	35.855661	205.710002	-34.276000	3.892000	105.919919
		7	28.212667	35.563439	207.386667	-33.333750	4.579750	106.336035
		8	28.060556	35.031504	208.171110	-31.027000	4.595000	106.122748
		9	28.284222	34.886179	207.644444	-30.321667	5.486667	106.136009
		10	28.738000	34.431447	207.946665	-28.316667	6.555333	106.754944
		11	28.449334	35.208669	207.763334	-26.357667	6.368000	106.692811
		12	27.748222	35.490718	209.246665	-25.575333	6.333000	106.446434
	2007	1	27.297334	35.559545	209.997779	-26.867000	6.146000	106.101382
		2	26.830889	35.678819	211.642220	-28.320000	6.723333	106.206474
		3	26.726000	35.816099	212.120000	-28.167333	7.562667	106.357337

Fig. 1 - Example Jupyter Notebook showing processing of Bio-Argo data.

Matlab, yet is free and open source. To solve the problem of using data from the archive we developed the Python module “biofloat” that handles all the details of scanning catalogs, finding the files to load and then creating local caches of data. The biofloat module is free and open source and is available on GitHub at https://github.com/biofloat/biofloat. It can be installed from the Python Package Index (PyPI) with the command “pip install biofloat”. The biofloat module provides access to Bio-Argo data using simple Python method calls and delivers it in easy-to-use Pandas dataframes. The GitHub site has several Jupyter Notebooks which demonstrate using biofloat for various analyses and visulizations.

Oxygen Calibration Use Case

Takeshita *et al*. (2013) present a climatology based quality control procedure for dissolved oxygen calibration utilizing the World Ocean Atlas 2009 (WOA09). We implement a similar procedure in Python and fully document it in the Jupyter Notebook calibrate_all_oxygen_floats.ipynb, which is available on the GitHub site. Fig. 1 demonstrates listing the monthly binned

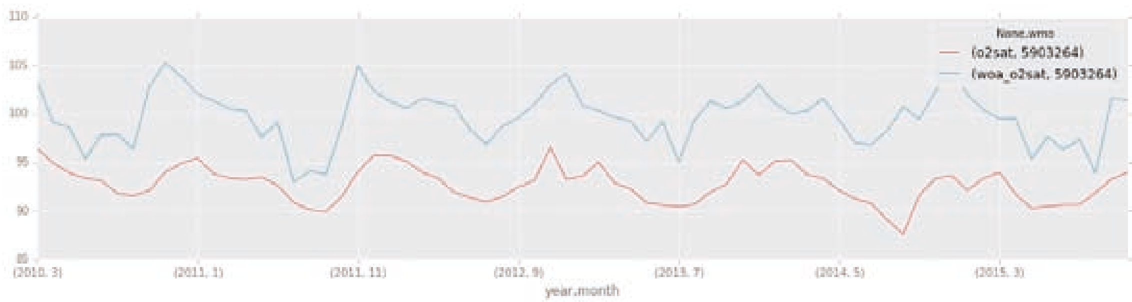


Fig. 2 - Over 5 years of surface oxygen saturation measured by Bio-Argo float 5903264 (blue) compared with oxygen saturation from the World Ocean Atas (red). The biofloat Python software enabled easy generation of this plot in a Jupyter Notebook.

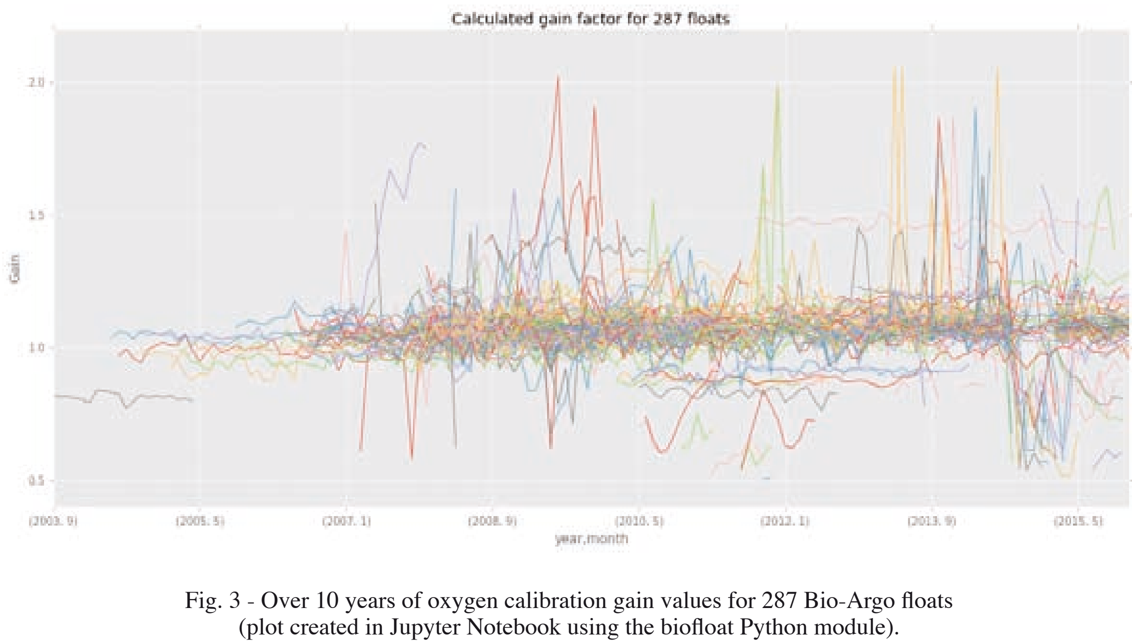


Fig. 3 - Over 10 years of oxygen calibration gain values for 287 Bio-Argo floats (plot created in Jupyter Notebook using the biofloat Python module).

surface float measurements in preparation for the WOA09 lookup. Fig. 2 shows surface oxygen saturation values for a single float compared with climatological values for the location and times of the float's trajectory. The ratio of these values is a gain factor that may be used to calibrate the float's oxygen profile. Fig. 3 shows the calculated oxygen calibration gain factors for 287 floats in the Bio-Argo archive. Biofloat is free to use for these kind of analyses, and source code contributions are welcomed to have it meet additional needs of the Bio-Argo user community.

Acknowledgements. Development of biofloat is supported by the David and Lucile Packard Foundation at the Monterey Bay Aquarium Research Institute.

References

Takeshita, Y., Martz, T.R., Johnson, K.S., Plant, J.N., Gilbert, D., Riser, S.C., Neill, C., and Tilbrook, B., 2013. *A Climatology-Based Quality Control Procedure for Profiling Float Oxygen Data*. Journal of Geophysical Research-Oceans, 118, 5640-5650. doi: 10.1002/jgrc.20399.

Database based Operational Eddy Detection in High-Frequency Radar Surface Velocities

Damir Ivankovič, Institute of Oceanography and Fisheries (Croatia), ivankovic@izor.hr
Vlado Dadič, Institute of Oceanography and Fisheries (Croatia), dadic@izor.hr
Stipe Muslim, Institute of Oceanography and Fisheries (Croatia), smuslim@izor.hr

Introduction

In the frame of project HAZADR (Strengthening common reaction capacity to fight sea pollution of oil, toxic and hazardous substances in Adriatic Sea) two high frequency radars are installed. Radars measure surface currents, wave height and direction every half hour. Measurements are shown on project web page (http://jadran.izor.hr/hazadr/).

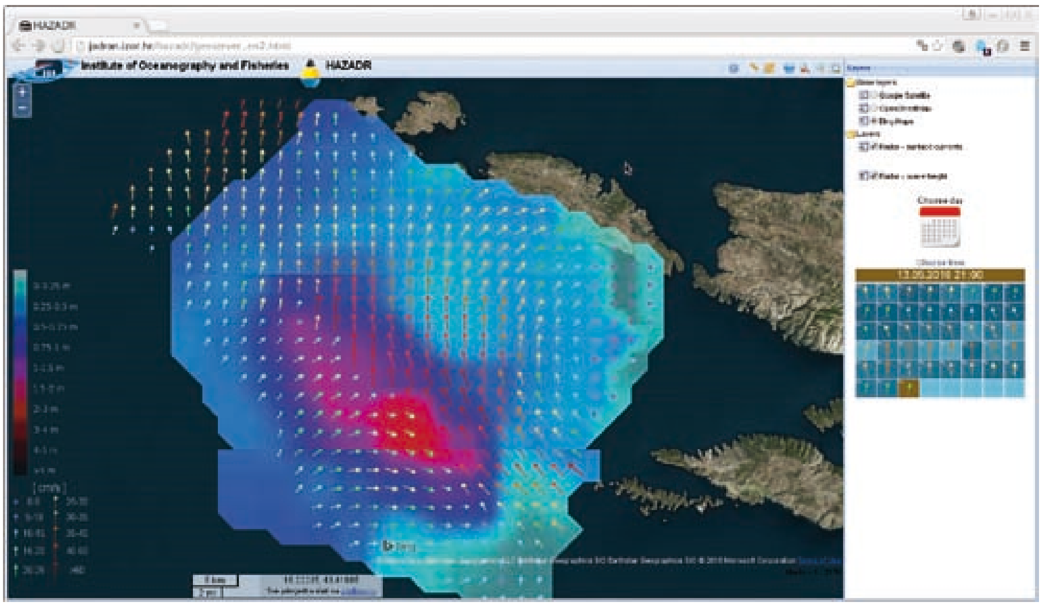


Fig. 1 - Visualisation Using Geoserver with database driven dynamic layers.

Eddy Detection

For data storage we use Oracle database 11.2 with Locator. Inside database are created procedures for eddy detection. Principle of detection is detection of changed directions in spatial fields of vectors. Procedure identifies eddy centre and provisional eddy borders onto only four main directions (North, East, South, West). For eddy detection field of daily averages is used. Also 6 hours average field can be used. Procedure is automatically executed after data insertion and averages calculation with crontab server mechanism. Detected eddies are also shown (number and type: cyclonic and anticyclonic) onto tables with averages.

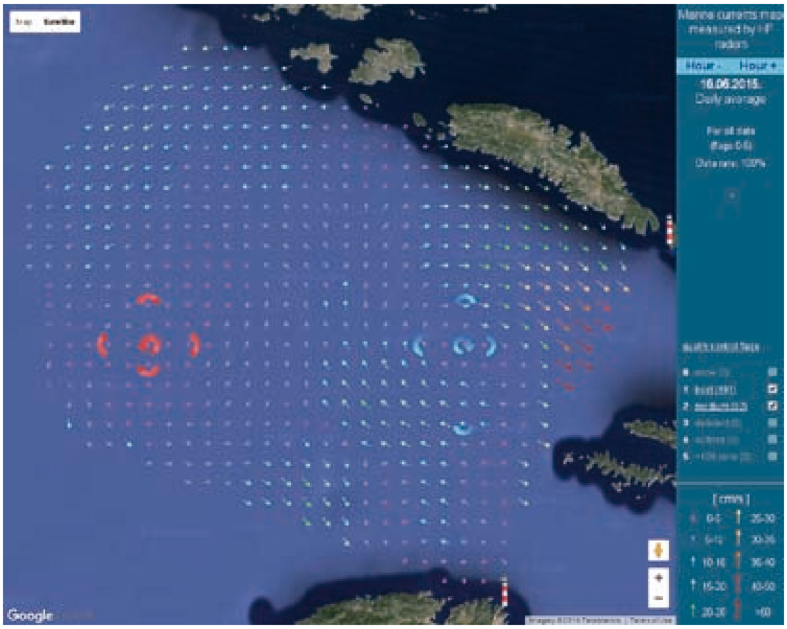
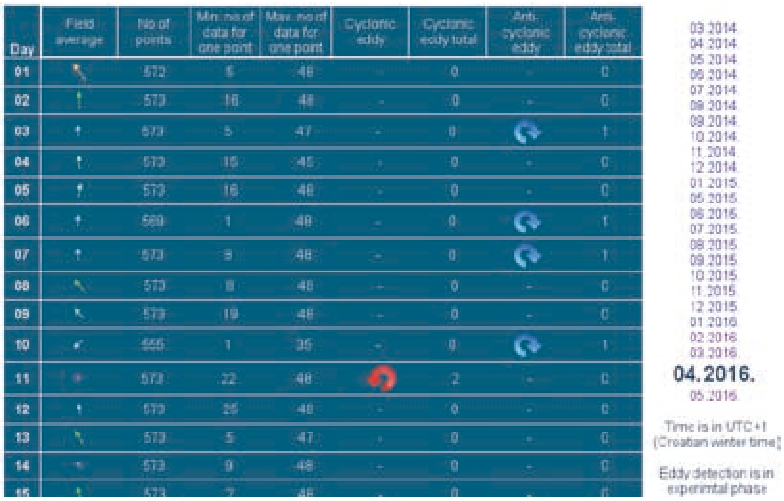


Fig. 2 - Detected eddies onto daily averaged field.



Day	Field average	No of points	Min. no of data for one point	Max. no of data for one point	Cyclonic eddy	Cyclonic eddy total	Anti-cyclonic eddy	Anti-cyclonic eddy total
01		572	5	48	-	0	-	0
02		573	16	48	-	0	-	0
03		573	5	47	-	0		1
04		573	15	45	-	0	-	0
05		573	16	48	-	0	-	0
06		569	1	48	-	0		1
07		573	3	48	-	0		1
08		573	8	48	-	0	-	0
09		573	19	48	-	0	-	0
10		555	1	35	-	0		1
11		573	22	48		2	-	0
12		573	25	48	-	0	-	0
13		573	5	47	-	0	-	0
14		573	9	48	-	0	-	0
15		573	7	48	-	0	-	0

Fig. 3 - Time axis of daily averages with eddies.

Future improvements and development

Procedure for eddy detection can be improved mainly to decrease false detected eddies. In smaller number of cases eddies are not detected. For detection improvement temporal sequence of detected eddies can be used (if for example eddy is detected in two 6h averages in the row, then one don't have eddy, and next have, probably in that one without eddy exists but it is not detected). Execution speed of procedure is very short, so procedure can be executed with every data input (every half hours). For now operational detection is onto daily averages, but until the end of the year will be also onto 6 hour averages.

Conclusion

Operational eddy detection can help us to improve our understanding of surface currents and improve surface currents predictions.

An analysis of biological data of the Georgian Black Sea coastal zone

Irine Baramidze, Environmental Pollution monitoring Branch of natural environmental agency, (Georgia), irine.baramidze@gmail.com

Tsisana Gvarishvili, Fisheries and Black Sea monitoring division of natural environmental agency, (Georgia), ciuri-gvarishvili@rambler.ru

Eteri Mikashavidze, Fisheries and Black Sea monitoring division of natural environmental agency, (Georgia), eteri-mikashavidze@rambler.ru

Guranda Makharadze, Batumi State University (Georgia) guranda.makharadze@yahoo.com

Kakhaber Bilashvili, Tbilisi State Universit (Georgia), wocen@telenet.ge

MadonaVarshanidze, WEFRI (Georgia),varshanidzem@yahoo.com

Basing on the biological information derived from SeaDataNet oceanographic data bases, for assessment of the current condition of biota of the Georgian Black Sea coastal waters, number of research activities has been carried out, as follow: identification of hydrochemical parameters,collection of phytoplankton,zooplankton,epifauna and bentic organisms, conservation, species identification, number and biomass determination, identification of dominant groups and species being essential for the state of ecosystems. In parallel with hydro-biological studies the hydrological and hydro-chemical indicators of sea water were determined. Hydrological parameters included pH, conductivity and salinity.

Studying of taxonomic composition, quantity, biomass and productivity of macroalgae gives a chance to make a conclusion on each waterpond condition. Phytoplankton plays a key role in the water quality formation.Zooplankton being as one of the most important thing in the feeding chain the studying of it assists to get an assessment of the water ecological condition.Epifauna is not considered only as a food for hydrobiota but it intensively activates in biofiltration process of the coastal zone. Despite of its Mediterranean origin the benthofauna of the Black Sea is 4-5 times insufficient.

With the result which we received by the study of phytoplankton we can conclude: There were 91 species and subspecies observed in samples of four seazon during whole years. 37 species of them were Diatoms, 30–Dinophyta, 11–Chromophyta, 6-Chlorophyta and 7–Cyanophyta. Dominant species mostly are fixe from Diatom group: Skeletonema coststum, Leptocylindrus minimus, Chaetoceros lorenzianus, Ch.affinis, Thalassionema nitzschioides.

Based on Epifauna data we can conclude that main groups composed of epifauna are Mollusca, Artropoda, Annelides, Platyhelminthes, Sarcodina, Tentaculata. The most abundant species were Mollusca (Bivalvias), Artropoda (Cructaceans) and Annelides (Popychaeta). Samples were collected in three stations: Tsikhisdziri, Green Cape and Kvariati (all stations along the coastal zone of the Black Sea, Georgia). In Kvariati samples were more diverse then others collected in Tsikhisdziri and Green Cape.

There were identified 46 species of zoobentos, in the bottom samples.Dominant speciese of this bottom settlement are: N. Cirrosa, from mollsca Ch.gallina and C. Cornea, arthropods C.pestai and B.improvisus.

CORA4.3: A global delayed time mode validated in-situ dataset

Tanguy Szekely, CORIOLIS R&D Centre National de la recherche scientifique (France), tanguy.szekely@ifremer.fr

Jérôme Gourrion, CORIOLIS R&D Centre National de la recherche scientifique (France), jerome.gourrion@ifremer.fr

CORA (Coriolis Dataset for ReAnalysis) is a delayed mode validated global temperature and salinity dataset provided by the Coriolis data center. This dataset provides delayed mode validated measurements from 1950 to 2015 and is used for the model reanalysis provided by the copernicus marine service (http://marine.copernicus.eu/).

Various new datasets have been added to CORA to enhnace the ocean data coverage in the 4.2 version. Among them, the French SHOM dataset (Service OceanorapHique de la Marine) have provided more than 3 million original profiles on the period 1950-2009.

The delayed mode data validation is based on two sets of tests. First, automatic tests set to detect and to flag obvious erroneous measurements. A second set of test is based on the comparison of the measurements with a minimum and a maximum temperature and salinity reference field. A warning is associated to each suspicious profile which is then manually checked by an oceanographer.

The next objective of CORA is to produce a copernicus CMEMS dataset by merging the best features of the CORA dataset and EN.4 datasets (www.metoffice.gov.uk/).

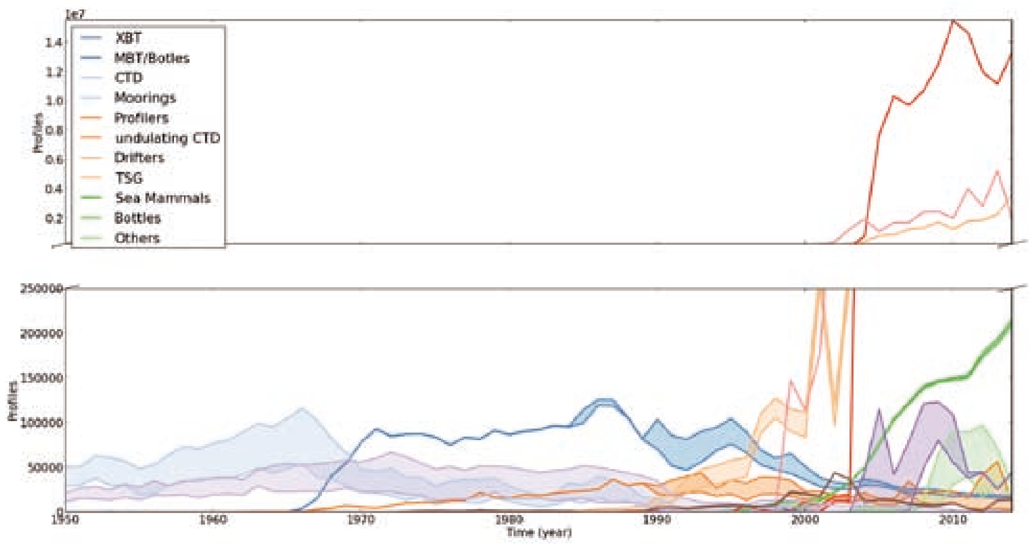


Fig. 1 - Yearly number of CORA profiles for various intrument types. For each intruments types, the lower line refers to CORA4.2 (released on april 2016) dataset and the upper line refers to the estimated number of profiles for the CORA4.3 dataset (to be released on april 2017).

EMODnet MedSea Checkpoint for Blue Economy efficiency and environment protection

Eric Moussat, Institut français de recherche pour l'exploitation de la mer (France), eric.moussat@ifremer.fr
Nadia Pinardi, Istituto Nazionale di Geofisica e Vulcanologia (Italy), n.pinardi@sincem.unibo.it
Giuseppe Manzella, Istituto Nazionale di Geofisica e Vulcanologia (Italy), giuseppe.manzella@ettsolutions.com,
Frédérique Blanc, Collecte Localisation Satellites Group (France), fblanc@cls.fr,
Simona Simoncelli, Istituto Nazionale di Geofisica e Vulcanologia (Italy), simona.simoncelli@ingv.it

The EMODNET checkpoints is a wide monitoring system assessment activity aiming to support the sustainable Blue Growth at the scale of the European Sea Basins by:

1) Clarifying the observation landscape of all compartments of the marine environment including Air, Water, Seabed, Biota and Human activities, pointing out to the existing programs, national, European and international
2) Evaluating fitness for use indicators that will show the appropriateness and availability of observation and modeling data sets and their roles and synergies based upon applications selected by the EU Directorate-General of Maritime Affairs and Fisheries (also known as DG MARE) to determine how the current infrastructure is compliant to the overarching aim of "collect once and use many times" philosophy of "Marine Knowledge H2020".
3) Prioritizing the needs to optimize the overall monitoring Infrastructure (in situ and satellite data collection and assembling, data management and networking, modeling and forecasting, geo-infrastructure) and release recommendations for evolutions to better meet the application requirements in view of sustainable Blue Growth.

Fig. 1 - EMODnet Central portal (http://www.emodnet.eu/): Sea Basin checkpoints.

The assessment is designed for:

- Institutional stakeholders for decision making on observation and monitoring systems
- Data providers and producers to know how their data collected once for a given purpose could fit other user needs
- End-users interested in a regional status and possible uses of existing monitoring data

The assessment is based on the development of selected end-user applications (“challenges”) which are of paramount importance for: (i) the Blue Economy sector (offshore industries, fisheries); (ii) marine environment variability and change (eutrophication, river inputs and ocean climate change impacts); (iii) emergency management (oil spills); and (iv) preservation of natural resources and biodiversity (Marine Protected Areas).

These applications have generated innovative products based on the existing observation landscape. The Medsea Checkpoint (http://www.emodnet-mediterranean.eu/) has attempted to base the fitness for use assessment of the input data used to build them on a more objective, quantitative and repeatable approach.

To do so, the Medsea checkpoint has based its approach on the ISO standards for Geographic Information (ISO 19157 Data Quality, ISO 19131 Data product specifications and ISO 19115 Metadata).

These standards provide not only the concept of quality but also a model - applicable both by data producers and data users - under which all differences of a dataset and a universe of discourse (the world of interest stated in specifications of production for a given purpose or of selection for a given use) can be:

- categorized
- evaluated
- recorded
- reported

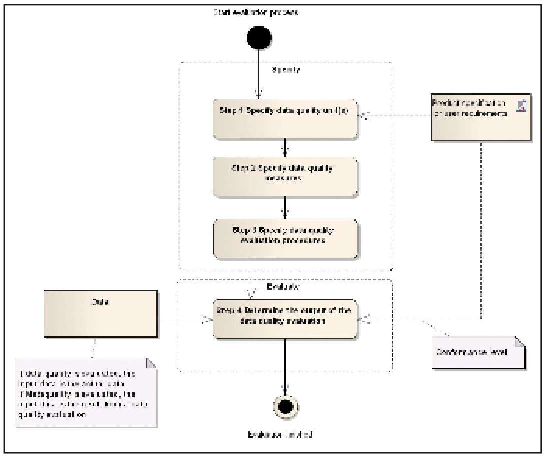


Fig. 2 - Assessment work flow according to ISO 19157.

The fitness for use of the input datasets are assessed using 2 categories of criteria to determine how these datasets fits the user requirements which drive them to select a data source rather than another one and to show performance and gaps of the present monitoring systems :

- data appropriateness i.e. the extent to which data (“what”) fits the user needs
- data availability i.e. the extent to which data can be discovered and accessed by users (“how”).

The most relevant quality criteria to evaluate the fitness for use of input data in term of appropriateness have been selected among the ISO 19157 quality elements according to the product specifications provided by the partners of each “challenge” (application). A series of quality measures corresponding to these criteria have been specified according to the standard. They are common to the applications of the checkpoint.

The output of the evaluation is a series of indicators which are derived automatically from the differences between expected values provided in the specifications of the product and the values determined from the actual product and from the input data sets used to build it.

9 quantitative indicators of the fitness for use of input data in term of "completeness", "consistency" and "accuracy" have been defined for the Medsea checkpoint input data. These indicators allow to give an overview of the extent to which input data fit the user needs in term of:

- observations (parameters),
- spatial and time coverage,
- spatial and time resolution,
- thematic accuracy
- temporal validity.

A similar process is applied to determine the fitness for use in term of data availability.

All information including product and input data selection specifications, quality measure and indicator definitions and values, are stored in a GIS platform in line with the ISO 19115/ISO19157 standards (Fig. 2) and made available by the means of front-end Web interfaces to discover:

- the input data used by the challenges (Input Datasets Catalogue),
- the innovative products generated by the challenges (Target Product Catalogue)
- the assessment indicators and adequacy information of each product and input data sets
- the Data Adequacy Reports which integrate both quantitative and non-quantitative – expert knowledge based - information - on the fitness for use of the input data.

The adoption of the SeaDataNet Common Vocabularies - governed by a combined SeaDataNet and MarineXML Vocabulary Content Governance Group - for the input data classification allows to display the results by discovery parameter relevant to the oceanographic and wider community.

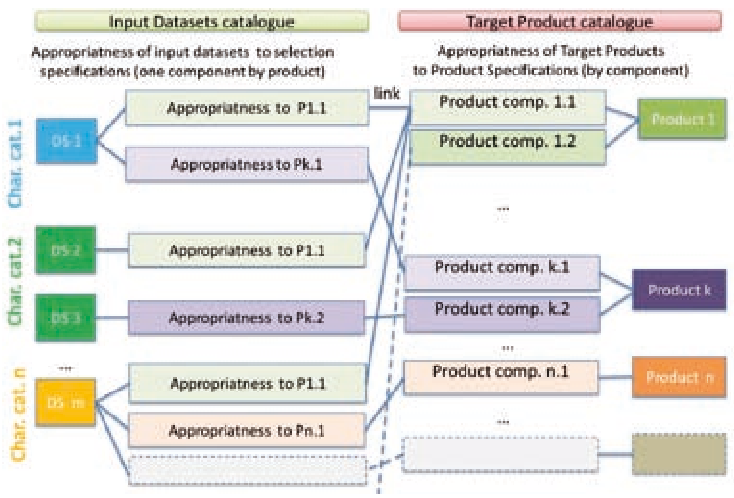


Fig. 3 - Data quality information in database. Fitness for use of input data are recorded by product and by application (challenge). Collection of input data are described by "characteristic" (parameter).

In conclusion, the application of the ISO based methodology for data quality assessment and of the SeaDataNet Common Vocabularies provides a reference framework to measure and to compare in a more objective way the 'data adequacy' from one application to another one and from one Checkpoint to another.

This initial Checkpoint Service paves the way for the creation of a virtuous loop to optimize the monitoring systems in function of user needs, to recycle the existing input data sets, in all basins in Europe and to further a global assessment: Atlantic and Black Sea Checkpoints have adopted the same methodology.

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

Bringing ocean observations to the classroom - student access to open data

Xavier Hoenner, Integrated Marine Observing System (Australia), xavier.hoenner@utas.edu.au
Sebastien Mancini, Integrated Marine Observing System (Australia), Sebastien.mancini@utas.edu.au
Roger Proctor, Integrated Marine Observing System (Australia), roger.proctor@utas.edu.au
Jason Everett, Sydney Institute of Marine Sciences (Australia), jason.everett@unsw.edu.au
Kerrie Swadling, Institute of Marine and Antarctic Studies (Australia), kerrie.swadling@utas.edu.au

For the past 4 years the Sydney Institute of Marine Science, a partnership of four Australian Universities (Macquarie University, the University of NSW, the University of Sydney and the University of Technology Sydney) has been running a Master's degree course called Topics in Australian Marine Science (TAMS). This course is unique in that the core of the course is built around research infrastructure – the Integrated Marine Observing System (IMOS). IMOS, established in 2007, is collecting unprecedented volumes of multi-disciplinary oceanographic data in the ocean and on the continental shelf which is made freely available across the web; IMOS frequently runs 'data user workshops' throughout Australia to introduce scientists and managers to the wealth of observations available at their fingertips.

The SIMS Masters course gives students an understanding of how different measurement platforms work and they explore the data that these platforms collect. Students combine attending seminars and lectures with hands on practical and personal assignments, all built around access to IMOS data and the many tools available to visualise and analyse. The course attracts a diverse class with many mature students (i.e. > 25 years old) from a range of backgrounds who find that the ease of discovering and accessing data, coupled with the available tools, enables them to easily study the marine environment without the need for high level computational skills. Since its inception the popularity of the course has increased with > 30 students undertaking the subject in 2016. The consensus from students and lecturers is that integrating 'real' observations into the classroom is beneficial to all, and IMOS is seeking to extend this approach to other university campuses. Similarly, the Institute of Marine and Antarctic Studies at the University of Tasmania is building IMOS data access into its Marine Ecology Masters course.

The talk will describe the experiences from the TAMS & IMAS courses and highlight the IMOS approach to data discovery, availability and access through course examples.

Ireland's Integrated Digital Ocean

Adam Leadbetter, Marine Institute (Ireland), adam.leadbetter@marine.ie
Eoin O'Grady, Marine Institute (Ireland), eoin.ogrady@marine.ie
Nigel Burke, Marine Institute (Ireland), nigel.burke@marine.ie

Introduction

Ireland has a marine resource extending to approximately 10 times its land area. With a renewed national focus on marine development a large number of agencies are operating in the marine sector. In parallel new digital technologies such as cloud computing and the internet of things are creating new opportunities. According to the European Commission "the digital economy…is the single most important driver of innovation, competitiveness and growth" and "how European businesses adopt digital technologies will be a key determinant of their future growth".

If applied effectively digital technologies can enable new marine research and innovation to support sustainable growth in Ireland's marine sector. Porter and Heppelmann (2015) say that "smart, connected products require a whole new supporting technology infrastructure" with the ability to integrate data from multiple sources being particularly important. Digital connectivity is also enabling new approaches to science and innovation with Open Innovation 2.O being "a new paradigm based on principles of integrated collaboration, co-created shared value, cultivated innovation ecosystems, unleashed exponential technologies, and extraordinarily rapid adoption".

Fig. 1 - The evolving innovation landscape (Curley and Salmellin).

In the marine sector digital services can leverage existing and future data assets to provide more timely and relevant information to support offshore operations, research and innovation, and to support the development of new marine-related public services. However the task of creating a coherent picture of the available data and information from these organisations is challenging, as the data originates from many different organisations and in many different formats.

In response to this need, and recognising that users and services often require direct access to data and information rather than to a catalogue, the Integrated Digital Ocean platform and associated portal is being developed with data contributions from fifteen organisations across Ireland.

Fig. 2 - The Digital Ocean logo.

The Digital Ocean Platform

The Digital Ocean Platform is designed to make it as simple as possible for data providers to contribute their results to the portal. As such, there are no minimum data requirements for a provider to join the network beyond providing a web-based interface for access to the observations or information.

The portal connects to various web services (including: ERDDAP from the Marine Institute; custom Application Programming Interfaces from the Commissioners of Irish Lights and Dublin City University; Open Geospatial Consortium Web Map and Web Feature Services from University College Cork and the Environmental Protection Agency) via a series of ingestion scripts.

These scripts allow the display of the data from the multiple originators in a user interface built on Open Source components (Fig. 3): jQuery for the general UI components; Leaflet for the mapping interface; HighCharts for graphing; and three js for WebGL display of 3D modelling of bathymetric features such as bays and shipwrecks and dynamic representation of oceanographic model results.

Fig. 3 - Open Source components of the Irish Digital Ocean porta UI.

The data is presented to the end user through a combination of maps, charts and card widgets (Fig. 4). These cards combine in-situ, model and archived data in a range of formats. They are reusable and can be quickly integrated into other applications.

Map layers come from a range of sources such as the INFOMAR seabed survey, Failte Ireland (the Irish tourist Board) and Dublin City Council's Dublin Bay Biosphere project.

All inputs to the Digital Ocean platform are monitored for data outages in a traffic-light dashboard system.

Fig. 4 - Irish Digital Ocean cards displaying various types of information (see www.digitalocean.ie)

The Digital Ocean network

Whilst the Marine Institute has coordinated the Digital Ocean network to date, there are a large number of partner organisations involved in Ireland's Integrated Digital Ocean, including:

- Commissioners of Irish Lights
- Dublin Bay Biosphere
- Dublin City University
- Electricity Supply Board
- Environmental Protection Agency
- INFOMAR seabed mapping project
- Irish Underwater Council
- National Parks and Wildlife Service
- SmartBay Ireland

- Sustainable Energy Authority of Ireland
- University College Cork

This network of data providers has been a key focus of the Irish Digital Ocean project, and while the approach needs to be further elaborated a long-term partnership approach is envisaged.

Conclusions

Ireland's Digital Ocean provides a flexible integrated platform to explore in data and information about Ireland's marine environment originating from a range of organisations. A significant focus has been on building the initial network of data contributors and developing the pilot portal to illustrate the concept to potential data providers and end users.

The current pilot phase of the portal extends beyond the cataloguing of data to the integrated display of data. The next development phase includes a data access broker which will negotiate various data services to allow end users or services to connect directly access data they have discovered. A series of "widgets" modelled on those provided by Facebook and Twitter will allow the re-use of information components in partner's websites, thus providing ready access to targetted information.

References

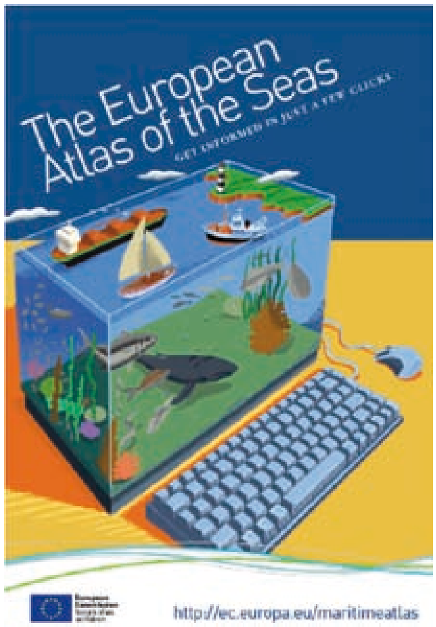
European Commission, 2016. *"The importance of the digital economy"*, [Online] Available at: http://ec.europa.eu/growth/sectors/digital-economy/importance/.

Porter, Michael E., Heppelmann, James E., 2015. *"How Smart Connected Products Are Transforming Companies"*, Harvard Business Review, PP.96-114.

Curley, M., Salmellin, B., 2013. *"Open Innovation 2.0: A New Paradigm - White Paper"*, [Online] Available at: https://ec.europa.eu/digital-single-market/node/66731.

A picture is worth a thousand words: the European Atlas of the Seas

Vittorio Barale, European Commission, Joint Research Centre (EU), vittorio.barale@jrc.ec.europa.eu



The European Atlas of the Seas – originally foreseen by the Integrated Maritime Policy (IMP) of the European Union (EU) as a simple public-oriented education and communication tool – today represents an experiment of knowledge brokerage in the marine world, mediating between the research domain and a varied group of practitioners, in need of using specialized information. The overall idea is to convert complex scientific data about coasts, seas and oceans into graphical form, so that non-specialists may access (and fuse) available products, without having to transfer or process large amounts of data, tasks that may require unavailable ad hoc technical skills. This may have a particular positive effect on EU rules and strategies related to marine/maritime matters, when the Atlas beneficiaries are policy-makers.

Examples of research innovations set for policy support are those linked to fisheries and Maritime Spatial Planning (MSP). In the first case, the Atlas' maps and charts are offering a new integrated approach to fisheries research results, previously available only in lengthy reports, aiming to improve data display and analysis and to monitor progress of policy implementation. In the second, the Atlas' EU-wide thematic map collection delivers a common baseline that can help EU Member States getting started on the MSP Directive requirements. As this is widely seen as a pre-requisite for Blue Growth, future research extending current projects results and facilitating MSP applications will help combine protection and sustainable use of the European Seas' ecosystem resources.

The Atlas is poised to become an advanced visualization and mapping tool, in the framework of the current effort to simplify and streamline the plethora of Marine Information Systems available online. The many initiatives in this field, either dealing with a holistic view of the sea or with specific maritime sectors, are a symptom of an unanswered requirement by a wide user community. Most initiatives would benefit from a unifying element, such as the Atlas is becoming, to ensure substantial savings in organizing and delivering data, and giving a second life to past scientific projects. Ultimately, the Atlas aims to help policy-makers cope with the incumbent data, information and knowledge overload, providing expertise in the use of modern, analytical tools.

The EU Atlas of the Seas is available at http://ec.europa.eu/maritimeaffairs/atlas/index_en.htm.

The EyeOnWater concept: Marine observations via participative science

Peter Thijsse, Marine Information Service (The Netherlands), peter@maris.nl
Marcel Wernand, Royal Netherlands Institute for Sea Research (The Netherlands), marcel.wernand@nioz.nl
Hans van der Woerd, Vrije Universiteit Amsterdam (The Netherlands), h.j.vander.woerd@vu.nl

EyeOnWater

The EyeOnWater (EOW) website and adjacent colour App helps to assess the colour of natural waters. The App was developed within the FP7 EU-Citclops consortium to estimate the colour of water bodies by citizens via an old oceanographic colour standard, the Forel-Ule (FU) scale. The colour of the waterbody is compared with 21 on-screen colour bars.

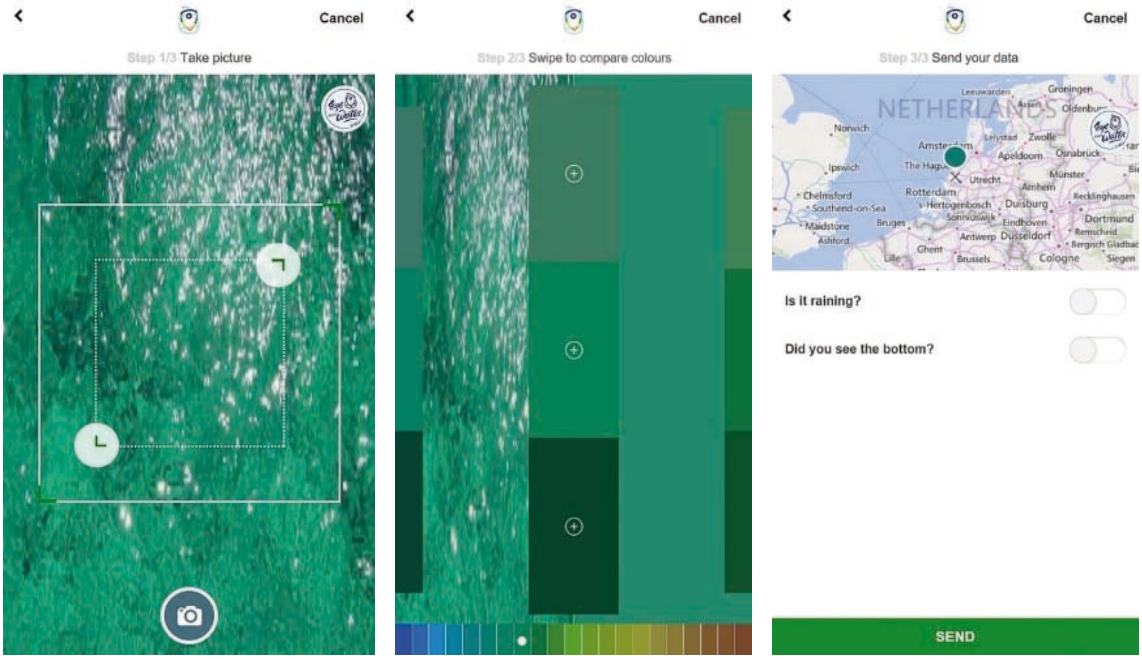
Smartphones are nowadays everywhere. Expensive observation satellites and in-situ measuring stations can now be complemented by low-cost EOW colour observations carried out by citizens. Data will be used by scientists (oceanographers, limnologists) and water authorities for statistical and long-term analysis in conjunction with climate research.

The Eyeonwater website www.eyeonwater.org and app (IOS and Android) have been launched end 2015. This presentation focusses on the implementation of the app, validation and the possible use of the data.

How does it work?

Water quality of natural waters can be estimated via their colour and transparency, since these optical properties are affected by the substances that are either suspended or dissolved in it. The EOW App facilitates the monitoring of colour changes of natural waters around you. The App was developed within the EU-Citclops consortium to estimate the colour of water bodies by citizens. The App is based on an old oceanographic colour standard, the Forel-Ule (FU) scale. The colour of the waterbody is compared with 21 on-screen colour bars.

The user takes a picture of the observed water surface, sun in the back, shining over one of the shoulders, then zooms-in at the part of the image that resembles the observed water colour best. The FU scale is then displayed on-screen and the user has colour-match and pick one of the coloured bars resembling the colour of the water (either by comparing the picture just taken or by looking at the water surface directly). When a colour is matched the SEND screen appears, showing your position and data can be uploaded to the EOW portal pressing SEND (data transfer only in reach of Wi-Fi).



Once the data is uploaded to the EOW portal one can create a user-account and sign-in at www.eyeonwater.org to look at the data. Once clicked on a data point info will show-up containing meta data, the FU colour-match and the FU calculated from the zoomed image (in expert mode also MFU). You can compare with:

i) Previously uploaded data, near-by data by others
ii) Historic data (1890 – 2000), over 200,000 observations

Scientific value of Eyeonwater data

Citizen data is various in quality and should not be used and distributed to wider community without validation. We will present the innovative methods how EOW data is validated and standardised.

The observations from the app are sent to the server and stored. The data concerns the image taken, the FU-index as marked by the user, and a set of metadata important for quality assessment: location, date/time, device type, angle, azimuth angle, installation ID etc.

EOW validates incoming data in two ways:

- Innovative algorithms were developed and published in the scientific literature to retrieve the colour index from the images taken with the App. We will present the results of these algorithms and how they can be used for quality control of the colour information provided by the App user.
- All app observations are visible on the Eyeonwater website. Other citizens are involved to correct each other. Users can flag a certain observation as “possibly not ok”, and a message is sent to an expert who has the final verdict. Simple but strong.

We will present the steps taken to fit this type of data seamlessly in the stream of data collected by operational coastal monitoring stations (a.o making use of OGC INSPIRE compliant services and INSPIRE compliant marine standards provided by SeaDataNet/EMODNet).

Currently actions are undertaken to expand the EyeOnWater concept to other parameters, and make a connection to water management organisations for uptake of the dataand validation in actual use cases.

KAPTAN - A smartphone application for mariners

Aldo Drago, University of Malta, Physical Oceanography Research Group (Malta), aldo.drago@um.edu.mt

Audrey Zammit, University of Malta, Physical Oceanography Research Group (Malta)

Raisa Tarasova, University of Malta, Physical Oceanography Research Group (Malta)

Adam Gauci, University of Malta, Physical Oceanography Research Group (Malta)

Anthony Galea, University of Malta, Physical Oceanography Research Group (Malta)

Joel Azzopardi, University of Malta, Department of Intelligent Computer Systems (Malta)

Giuseppe Ciraolo, University of Palermo, Department of Civil, Environmental and AeroSpace Engineering (Italy)

Fulvio Capodici, University of Palermo, Department of Civil, Environmental and AeroSpace Engineering (Italy)

As the CALYPSO Follow On project (www.capemalta.net/calypso) came to completion in December 2015 with the addition of a fourth HF radar station in Ragusa Harbour, another major milestone was being accomplished with the launching of KAPTAN, an integrated service of met-ocean information delivered online and on smartphone to aid sea farers navigating in the proximity of the Maltese Islands and south of Sicily for planning their journeys and ensuring safer trips. The initiative follows the trail of efforts by the Physical Oceanography Research Group (ex PO-Unit) of the Dept. of Geosciences within the University of Malta to deliver services deriving from operational oceanography and meteorology to users, not only at the level of national stakeholders that require data and information for their routine operations, but also to the general public by making use of popular media and affordable smart technologies.

KAPTAN (the equivalent Maltese word for 'Captain' in English), provides an aid to sea farers navigating in the proximity of the Maltese Islands and southern Sicily to plan their journeys and to monitor meteo-marine conditions in real time for safer trips. Just a few clicks on a smartphone application leads users to a suite of sea and weather data in the form of interactive maps providing instantaneous user friendly and user defined access to prevailing conditions at sea as well as short term past and forecast information.

The data for this integrated service to mariners is mainly derived from the CALYPSO HF radar observing system, consisting of a network of CODAR SeaSonde installations on the northern Maltese and southern Sicilian shores at four selected sites. High resolution weather and marine numerical models run at the University of Malta specifically for the Malta-Sicily Channel together with satellite observations of the area, provide a full suite of very local reporting, and complement other weather forecasts derived from GFS/WRF models and local weather stations.

The services are delivered on KAPTAN using Google Maps API and are composed of six components: sea surface currents on 2D Eulerian maps; sea surface currents along transects; sea current drift; Atmospheric forecasts – wind vector, air temperature, precipitation, mean sea level pressure; sea surface temperature - observed and forecast; sea wave conditions – significant wave height, wave direction, peak period and mean period.

When users open KAPTAN, the smartphone device sends requests for data, over an Internet connection, to a Simple Object Access Protocol (SOAP) web-based service. The extraction from the available data sources (observations, models, local and third party) is performed through text files and maps that are generated by dedicated Matlab functions; they contain data belonging to one time frame or a group of related time frames. The file names have a prefix showing the service that they relate to, and a date/time stamp to show the time relating to the data they contain.

All services are implemented as a Java web application. Requests to consume the service are transmitted in EXtensible Markup Language (XML) format using HTTP and include strings carrying the required input parameters.

The KAPTAN smartphone application can be downloaded for free for both Android and iOS

Fig. 1 - Snapshots of the KAPTAN online and smartphone service delivery.

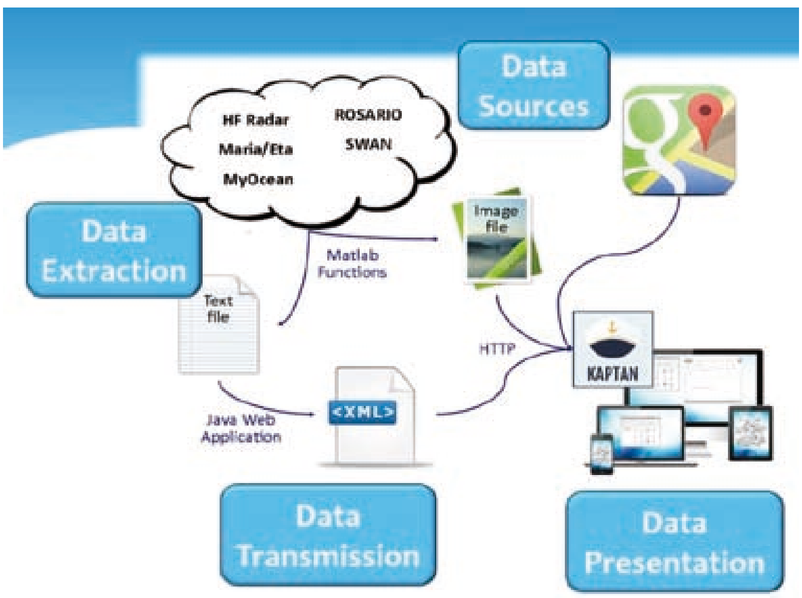


Fig. 2 - Schematic of data flow: sourcing and extraction from observing and modelling platforms to data transmission and presentation on the web and smartphone interfaces.

devices (Google Play and App Store on iTunes respectively). The same services are also available online on www.capemalta.net/calypso/kaptan.

Acknowledgements. The CALYPSO and CALYPSO Follow On projects were partially funded by the ERDF Italia – Malta Programme, Cohesion Policy 2007-2013. Besides the Physical Oceanography Research Group (project lead) from the University of Malta, the partnership comprised four Sicilian partners: University of Palermo (Polo Territoriale Universitario della Provincia di Trapani) as Sicilian Focal Point, L'Istituto per l'Ambiente Marino Costiero del Consiglio Nazionale delle Ricerche (IAMC-CNR), University of Catania (CUTGANA) and ARPA SICILIA (Agenzia Regionale per la Protezione dell'Ambiente), and three other Maltese partners: Transport Malta, Armed Forces of Malta and Civil Protection Department.

References

Drago A., J. Azzopardi, A. Gauci , R. Tarasova, G. Ciraolo, F. Capodici, S. Cosoli and M. Gacic, 2013. *'Sea Surface Currents by HF Radar in the Malta Channel'*, Rapp. Comm. Int. Mer Medit., 40, p144.

Drago A., G. Ciraolo, F. Capodici, S. Cosoli, M. Gacic, P.-M. Poulain, R. Tarasova, J. Azzopardi, A. Gauci, A. Maltese, C. Nasello and G. La Loggia, 2015. *'CALYPSO – An operational network of HF radars for the Malta-Sicily Channel'*, Proceedings of the Seventh International Conference on EuroGOOS, 28-30th October 2014, Lisbon, Portugal. Edited by H. Dahlin, N.C. Fleming and S. E. Petersson. First published 2015. EuroGOOS Publication no. 30. ISBN 978-91-974828-9-9.

Insights from the development of a downstream web-service to visualize ocean and meteorological forecast data at Iberian Atlantic beaches

Pablo Otero, Instituto Español de Oceanografía Madrid (Spain), pablo.otero@md.ieo.es
Gonzalo Gonzalez-Nuevo, Instituto Español de Oceanografía Vigo (Spain), gonzalo.gonzalez@vi.ieo.es
Elena Tel, Instituto Español de Oceanografía Madrid (Spain), elena.tel@md.ieo.es
Manuel Ruiz-Villarreal, Instituto Español de Oceanografía A Coruña (Spain), manuel.ruiz@co.ieo.es
José Manuel Cabanas, Instituto Español de Oceanografía Vigo (Spain), jmanuel.cabanas@vi.ieo.es

Introduction

Coastal modeling and observation, data management as well as the development of marine core and downstream services for intermediate and end-users, are critical components aligned with the EU's Blue Growth strategy. The marine knowledge value chain may focus on areas with special potential to foster a smart, sustainable and inclusive Europe, as for example, the coastal and maritime tourism, the biggest maritime sector in terms of gross value added and employment (ECORYS, 2013).

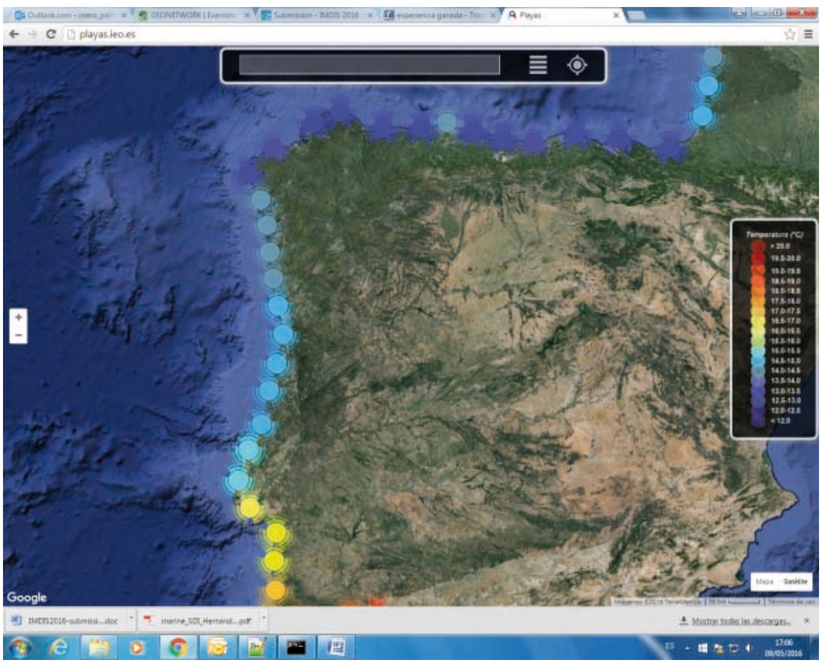


Fig. 1 - Home page of the application playas.ieo.es.

This is particularly true in countries with favourable climate conditions and long coastlines like Spain and Portugal.

In 2013, the Instituto Español de Oceanografía (IEO) developed a maritime data downstream service (http://playas.ieo.es) particularly focused on providing sea surface temperature (SST) at Iberian Atlantic beaches, a parameter that, although basic for the scientific community, it was certainly rather unusual in those weather forecast reports to the wider public. Thus, although the web-service enhances the visualization of SST at more than one thousand beaches, it also provides 3-day forecast of a set of other meteo-ocean variables (air temperature, wind speed and direction, wave height, period and direction, and tides). The service is accompanied of a dedicated page in Facebook to spread marine knowledge directly related to SST. The main elements that compose this service and the experience gained during these years are described.

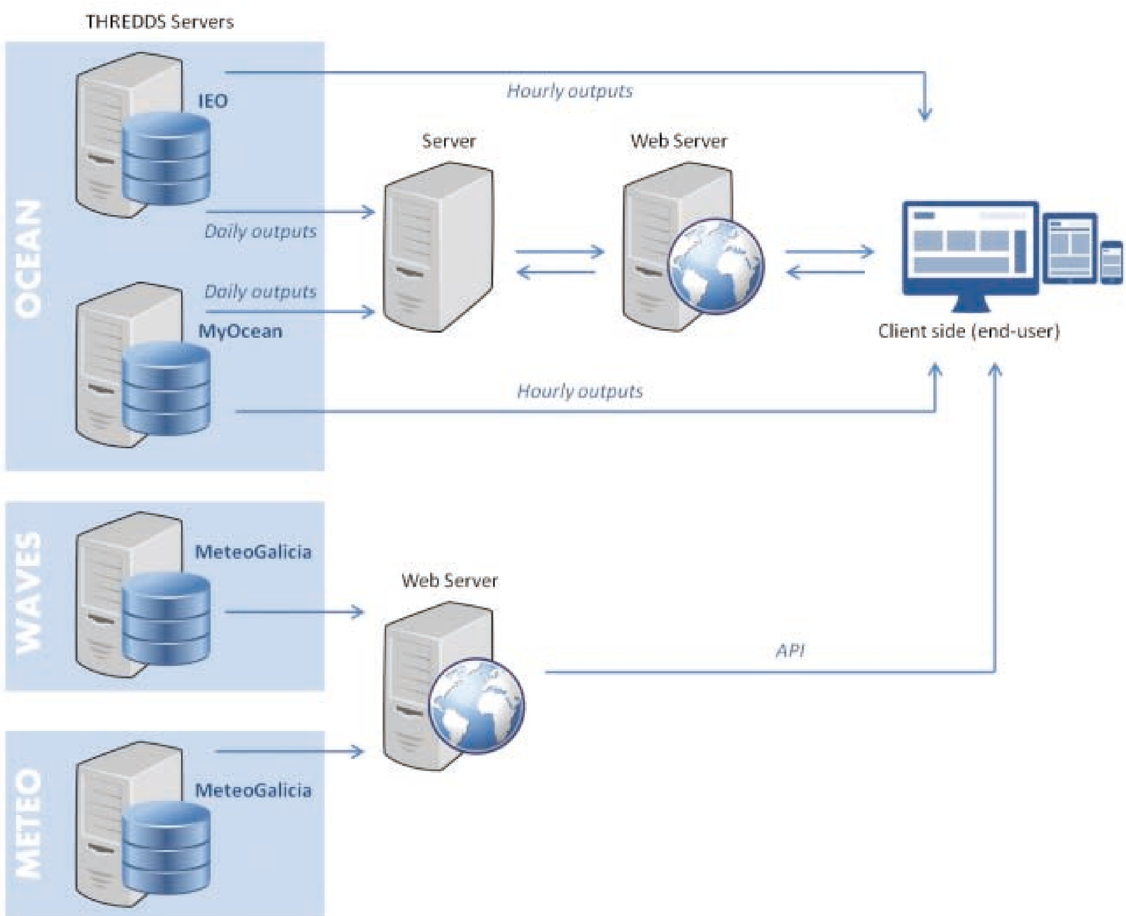


Fig. 2 - Operation scheme.

Methods

The main parameter of the web service is the SST. Data are obtained from a realistic high resolution (~1.3 km) configuration of a numerical ocean model, operated and mantained by IEO as part of the RAIA Ocean Observatory (http://www.marnaraia.org/). Supplementary details of the model configuration and validation can be found in Otero et al. (2009, 2013). In areas where the model configuration does not provide results or during out-of-service periods, data are obtained from the ~3 km ocean model configuration of the NorthEast Atlantic provided by Copernicus service (http://marine.copernicus.eu/) and operated by Puertos del Estado (http://www.puertos.es/).

Wave and meteorological data come, respectively, from regional configurations of the WW3 (Wave Watch III) and the WRF (Weather Research Forecast) models operated and maintained by MeteoGalicia (http://www.meteogalicia.es/). Data are easily accessible through an API service.

The downstream web-service has been programmed by using HTM5/JS/CSS technology and adapted for optimum visualization in mobile devices. When the end-user access to the application, SST data interpolated at beach locations (stored in a JSON file) are visualized using an optimum colour scale. To facilitate the visualization, data are clustered in dependence with the zoom level. The data shown at this stage comes from daily-averaged outputs of the ocean models that were previously downloaded to IEO servers. However, if the user clicks over a specific beach, a table is opened to show hourly data and the forecast. Queries are performed directly by the client to THREDDS Data Servers and the API provided by MeteoGalicia. The aim is to provide data at 1 h of temporal resolution. In case of delay or failed connection, the data shown in the web service correspond with the daily-averaged outputs.

The web-service makes easy to search for the warmest (or the coldest) beach in the visualized area, which turns attractive to localize optimum bathing waters in areas with strong SST gradients like the *rías* (inlets in NW Spain).

Results and conclusions

The development of the application playas.ieo.es has been a good opportunity to test and improve different data services aligned with the current state-of-the-art in operational oceanography. Although this web app has been initially deployed as a Beta version and no particular advertisement and promotion have been done, from 2013 on, more than 30K users have accessed to the application, particularly during summer, when the service experiences an average of 140 daily connections. In addition, the App has a link to a Facebook page with curious or relevant information for the general public. Explanations about massive jellyfish occurrence, upwelling or local strong coastal currents are done by experts in a scientific dissemination effort. Currently, this downstream service is being used for Sondara Solucions, a small Spanish enterprise, to provide additional information in their smart beach products, becoming an example of those priorities drawn in the EU's Blue Growth strategy.

References

ECORYS, 2013. *Study in support of policy measures for maritime and coastal tourism at EU level*. Specific contract under FWC MARE/2012/06 -SC D1/2013/01-SI2.648530. Final Report. Client: DG Maritime Affairs & Fisheries. Rotterdam/Brussels, 15 September 2013. Available at: http://ec.europa.eu/maritimeaffairs/documentation/studies/documents/study-maritime-and-coastal-tourism_en.pdf

Otero, P., M. Ruiz-Villarreal and A. Peliz, 2009. *River plume fronts off NW Iberia from satellite observations and model data*. ICES Journal of Marine Science, 66, doi:10.1093/icesjms/fsp156

Otero, P., M. Ruiz-Villarreal, L. García-García, G. González-Nuevo and J.M. Cabanas, 2013. *Coastal dynamics off Northwest Iberia during a stormy winter period*. Oceans Dynamics 63(1), 115-129.

POSTERS

Development and use of HELCOM COMBINE phytoplankton dataset in SeaDataNet compliant format for the Baltic Sea

Seppo Kaitala, Finnish Environment Institute (Finland), Seppo.Kaitala@ymparisto.fi
Sirpa Lehtinen, Finnish Environment Institute (Finland), sirpa.lehtinen@ymparisto.fi
Sanna Suikkanen, Finnish Environment Institute (Finland), sanna.suikkanen@ymparisto.fi
Jani Ruohola, Finnish Environment Institute (Finland), jani.ruohola@ymparisto.fi
Hermanni Kaartokallio, Finnish Environment Institute (Finland), hermanni.kaartokallio@ymparisto.fi

Marine environmental management as well as scientific demands of the development in marine ecology needs a biological data system to store, archive, integrate and update marine biological, especially taxonomical information. A biological database should provide detailed information on procedures on sampling, preservation, analysing, taxonomical identification, and biovolume determination, in order to be able to select comparable results. For taxonomical identification, an authoritative and comprehensive general list of valid and synonymized names of marine organisms is required. For that purpose, the World Register of Marine Species (WoRMS) was developed as an offspring of European Register of Marine Species (ERMS) maintained at the Flanders Marine Institute (VLIZ). AphiaID, specific identification number for each marine taxon is the key parameter to identify taxa in the WoRMS.

In addition to a commonly acknowledged list of valid taxa names, common rules for determination of biovolumes are required. HELCOM Phytoplankton Expert Group (PEG) has compiled the list of taxa and biovolume (*Biovolume file*) for the Baltic Sea phytoplankton. In the HELCOM PEG *Biovolume file*, biovolumes are calculated using agreed geometrical shapes and various size classes for each species. The ongoing tasks of PEG group include updating the HELCOM COMBINE manual for Baltic Sea phytoplankton monitoring procedures, and updating the *Biovolume file* with the addition of AphiaID and AphiaID link according to WoRMS, addition of species synonyms, and addition of new species and size classes. Currently the PEG list includes 916 phytoplankton taxa with an individual AphiaID and AphiaID link, as well as information on trophy, geometric shape, size class, cell dimensions, biovolume calculation formula and carbon content of each taxon.The HELCOM COMBINE phytoplankton data forms an important biodiversity data set with continuous data flow from the Baltic Sea area and it is managed by HELCOM data host ICES.

Parallel tasks are included in other European projects as SeaDataNet, EMODnet Biology and JERICO-Next. The collaboration between projects is desired and the data standards are under development. ODV format is a candidate for a common data standard in marine biological taxonomy. Also common software would help the standardization on the reporting. Use of NEMO software provided by SeaDataNet project is demonstrated to manage marine phytoplankton observations to produce data output in ODV format coupling the need for HELCOM reporting.

The Black Sea GIS developed on basis of free software

Elena Zhuk, Marine Hydrophysical Institute RAS (Russian Federation), alenixx@gmail.com
Alexey Khaliulin, Marine Hydrophysical Institute RAS (Russian Federation), khaliulin@yahoo.com
George Zodiatis, Cyprus Oceanography Centre, UCY (Cyprus), gzodiac@ucy.ac.cy

Using and developing up-to-date methods, tools, and technologies as well as assuring automatic processing and representing oceanographic data and knowledge, with geoinformation systems (GIS) being in the first place, are the most efficient instrument of providing users with various oceanographic information [1].

System structure

While creating the Black Sea GIS, the primary attention was paid to selecting free and platform independent software capable to provide necessary functionality [2].

The software for online data access and visualization was developed on the basis of client server architecture. Its server part includes:

- oceanographic database (ODB) developed at MySQL
- map service, which was made at Mapserver
- python-modules performing data processing
- php-modules providing interaction between server and client applications.

Functionality

Client part is a user-friendly interface providing the following functions:

- Export of oceanographic data in ODV format.
- Selection of oceanographic data on rectangular region, temporal interval, season, cruises, type of measurements, hydrological and/or hydrochemical parameters; and visualization of required oceanographic data (presenting stations on the map, making profiles of oceanographic parameters, metadata and data output (Fig. 1).

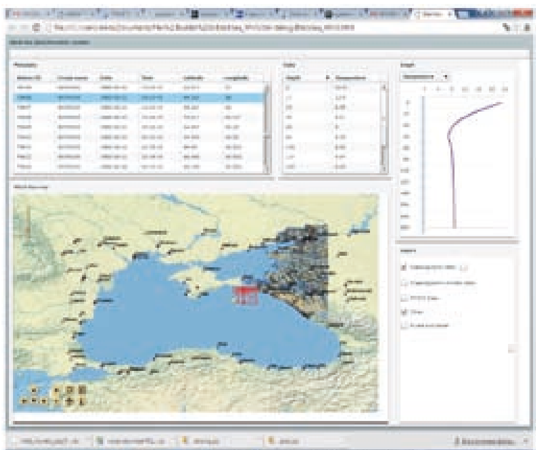


Fig. 1 - An example of displaying the result of selecting oceanographic data.

- Selection and visualisation maps of the climatic atlas in shape formats. The maps can be selected on certain parameters (seawater temperature, salinity, density, heat storage, oxygen concentration in seawater, aerobic waters low boundary location, and hydrogen sulfide zone upper boundary location), levels, seasons, and years; and visualized as both bitmaps and isolines.

- Selection and display MODIS AQUA satellite images in GeoTiff format for any accessible date on the following parameters: sea surface temperature, chlorophyll a concentration, and water leaving radiation.
- Selection and displaying social and economic data for coastal regions.
- Overlay different data types (Fig. 2).

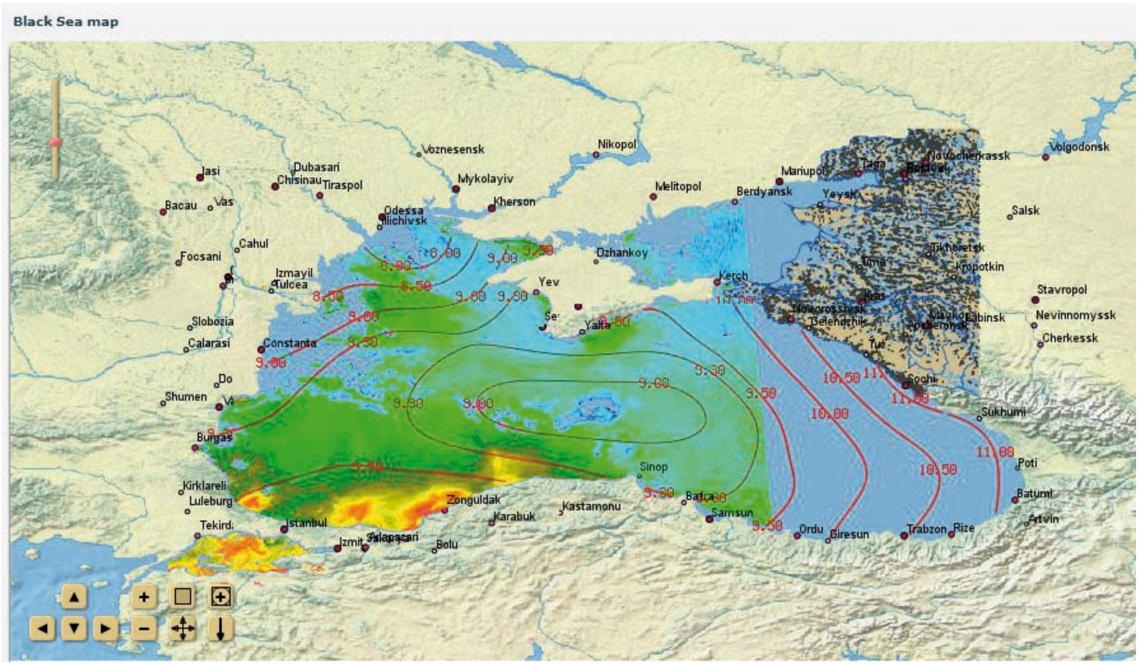


Fig. 2 - Overlaying a climatic atlas map and a satellite image.

The openness of the Black Sea GIS architecture gives excellent capabilities for extending its functionality, adding parameters, updating the existing and connecting new databases. At present, the work on improving the Black Sea online GIS is continued.

References

[1] Eremeev V., Khaliulin A., Ingerov A., Zhuk E., Godin E., PlastunT. *"Current state of the oceanographic data bank at MHI NAS of Ukraine: software",* Marine Hydrophysical Journal, Vol.2, 2014, p.54-66.

[2] Zhuk E., Khaliulin A., Ingerov A., Godin E. *"The experience of Black Sea GIS developing on basis of free software"*, GEOPROFI, Vol. 2, 2016, p. 36-39.

Medclic: the Mediterranean in one click

Charles Troupin, Balearic Islands Coastal Observing and Forecasting System (Spain), ctroupin@socib.es
Biel Frontera, Balearic Islands Coastal Observing and Forecasting System (Spain), bfrontera@socib.es
Joan Pau Beltran, Balearic Islands Coastal Observing and Forecasting System (Spain), jbeltran@socib.es
Andreas Krietemeyer, Balearic Islands Coastal Observing and Forecasting System (Spain), akrietemeyer@socib.es
Kristian Sebastian, Balearic Islands Coastal Observing and Forecasting System (Spain), ksebastian@socib.es
Sonia Gómara, Balearic Islands Coastal Observing and Forecasting System (Spain), sgomara@socib.es
Miquel Gomila, Balearic Islands Coastal Observing and Forecasting System (Spain), mgomila@socib.es
Romain Escudier, Balearic Islands Coastal Observing and Forecasting System (Spain), rescudier@socib.es
Mélanie Juza, Balearic Islands Coastal Observing and Forecasting System (Spain), mjuza@socib.es
Baptiste Mourre, Balearic Islands Coastal Observing and Forecasting System (Spain), bmourre@socib.es
Angels Garau, Balearic Islands Coastal Observing and Forecasting System (Spain), agarau@socib.es
Tomeu Cañellas, Balearic Islands Coastal Observing and Forecasting System (Spain), tcanellas@socib.es
Joaquín Tintoré, Balearic Islands Coastal Observing and Forecasting System (Spain), jtintore@socib.es

The Medclic project

"Medclic: the Mediterranean in one click" is a research and dissemination project focused on the scientific, technological and societal approaches of the Balearic Islands Coastal Observing and Forecasting System (SOCIB, www.socib.es). It is a collaboration with *"la Caixa"* Foundation.

SOCIB aims at research excellence and the development of technology which enable progress toward the sustainable management of coastal and marine environments, providing solutions to meet the needs of society. Medclic goes one step forward and has two main goals:

1. at the scientific level: to advance in establishing and understanding the mesoscale variability at the regional scale and its interaction, and thus improving the characterization of the *"oceanic weather"* in the Mediterranean;
2. at the outreach level: to bring SOCIB and the new paradigm of multi-platform observation in real time closer to society, through scientific outreach.

Fig. 1 - Home page of the Medclic project.

The role of the Data Centre

SOCIB Data Centre is the core of the new multi-platform and real time oceanography and is responsible for directing the different stages of data management, ranging from data acquisition to its distribution and visualization through web applications.

The implemented system relies on open source solutions and provides data in line with international standards and conventions (INSPIRE, netCDF Climate and Forecast...). In addition, the Data Centre has implemented a REST web service, called *Data Discovery*. This service allows data generated by SOCIB to be integrated into applications developed by the Data Centre itself or by third parties, as it is the case with Medclic. It also facilitates the transfer of data from SOCIB to other international portals, such as EMODnet-Physics, CMEMS INSTAC or MONGOOS.

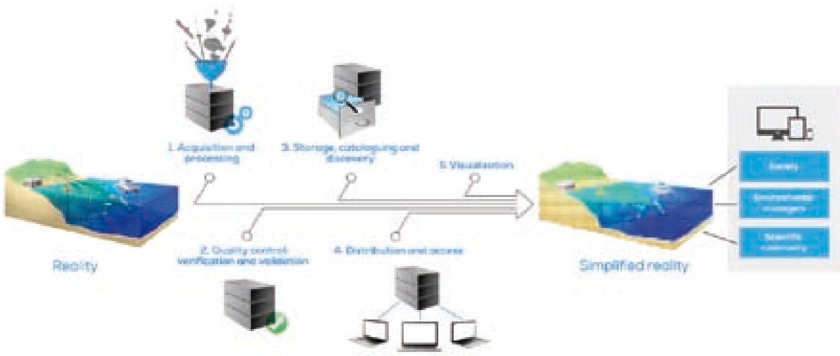


Fig. 2 - Connexions and roles of the Data Center.

Results

Relying on this data distribution, the new web Medclic, www.medclic.es, constitutes an interactive scientific and educational area of communication that contributes to the rapprochement of the general public with the new marine and coastal observing technologies.

Thanks to the Medclic web, data coming from new observing technologies in oceanography are available in real time and in one click for all the society. Exploring different observing systems, knowing the temperature and swell forecasts, and discovering the importance of oceanographic research will be possible in a playful and interactive way.

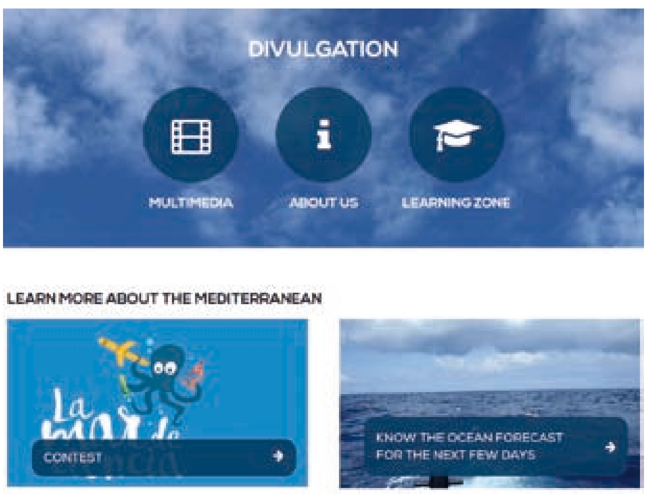


Fig. 3 - Outreach section of the Medclic project.

Evolution of physical oceanography observations in Finland

Pekka Alenius, Finnish Meteorological Institute (Finland), riikka.hietala@fmi.fi
Riikka Hietala, Finnish Meteorological Institute (Finland), riikka.hietala@fmi.fi
Kimmo Tikka, Finnish Meteorological Institute (Finland), riikka.hietala@fmi.fi

Systematic physical oceanography measurements in Finland date back to 1887 when a permanent sea level station was founded in Hanko. Regular research cruises were started in 1898. Also the sea ice observation network was established at the end of 1800's.

The Finnish sea level station network got its second station, Helsinki in 1904. Ten more stations were built in 1920's, then one in 1933 and the newest in 2014. The old stations recorded sea level continuously on paper rolls. That data is digitized in four hours intervals till 1970 and with one-hour interval since then. Nowadays the stations record data in digital form with one-minute intervals. The remarkable thing is that it is possible to go back to the original recordings and digitize them again. That has been done in recent studies of meteotsunamis. Those have been identified between the digitized time steps of data.

The coasts of Finland are in many places full of islands and bank and therefore in addition to lighthouses also light ships were in use the first half of 1900's during ice-free seasons. The ships were anchored in fixed positions. Regular daily observations of weather parameters and sea temperature and currents were part of the life of those vessels. Salinity was also observed by taking water samples. In addition to the light ships, oceanography observations were made on fixed oceanographic stations, too. Those stations were located to as deep as possible places near the coast near to pilot stations or other places in archipelagos where people lived. The number of such fixed stations varied from year to year considerably. After 1950's the light ships were no more needed and the number of people living at the outer edges of the archipelago diminished, which led eventually to drastic decrease of the number of fixes stations. At present there are only five such left and instead of water samples the CTD is used to get the salinity and temperature profile.

The Finnish research cruises were done first with steam ship Suomi, then with Nautilus from 1903 end of 1930's. In 1938 a new ship was build for marine research and it was named Aranda. Because of the Second World War it did only 1 research cruises. After the war it was delivered to Soviet Union as part of the war reparations. In 1953 a combined research and passenger ship Aranda was taking into used. It served first research in the summers and passengers in the archipelago in late autumns and in spring when there was still ice on the sea. In 1970's the ship was entirely used for research and the cruises extended to cover wintertime, too. The old Aranda was replaced in 1989 with a new Aranda, which is still in use. It was built to be usable in all oceans and it cruised to Weddell Sea in 1989 when Finland joined the Antarctic Treaty. The present Aranda was planned with scientists and built entirely for research purposes.

The building of Aranda 1989 was a milestone in Finnish oceanography because then the oceanographic data systems were renewed, too. The ship included multi scientific database system, laboratory information management system, cruise planning system and internal and external information's systems. Those included the delivery of data to land during the cruises and once a day updates of the position map and weather conditions on Aranda in the Aranda Internet pages. The database system included automatic compression of data to files that the crew sent to the institute via e-mail. On land the files were decompressed and data put to the database. Thus the "rolling deck to repository" was reality in early 2000's.

Automatic real-time observation began in 1980's when the use of wave buoys at fixed position(s) was begun after several years of measuring the wave conditions in different sea areas for research and statistical purposes. At present semi permanent wave buoy stations are in the northern Baltic Sea proper, Gulf of Finland, Bothnian Sea and Bothnian Bay. These buoys are larger Wave Rider buoys. Smaller wave buoys are in use on research projects.

Fig. 1 - Taking samples from Nautilus, FMI's Argo-buoy, Aranda and glider. (Photos from FMI's archives).

The drifting buoys have been used in 2010's for surface drift research, for ice research and nowadays they are used as semi permanent temperature buoys in anchored mode, too.

FMI took as the first institute, the Argo buoys into use in the Baltic Sea. As a member of Euro-Argo, Finland, has been using Argo buoys in the Bothnian Sea since 2012 and in the Baltic Sea Proper since 2013. These buoys have two-way satellite communication system and they transmit data once a week. The buoys are replaced once a year. Thus the Baltic Sea Argo buoys can be used again and again.

The newest step in physical oceanography observations in the taking into use of underwater glider. FMI was a participant in European glider project GROOM and tested with Spanish colleagues a glider in the Bothnian Sea and Archipelago Sea. After these successful tests, FMI has got into its repertoire a Slocum shallow sea glider as part of the Finnish marine research infrastructure FINMARI.

A permanent ICOS station has been established at the Utö Island in the outer edge of the Archipelago Sea towards the Baltic Sea proper. That station has many oceanographic measuring systems included, too. The next step will be a bottom-mounted cabled underwater profiling system. It can operate all year round, a great advantage in areas with ice-cover in winter.

As a conclusion we can state that there has been success stories and obstacles to over come in Finnish physical oceanography observations. Active participation in international observation campaigns and systems, and in organizations like ICES, IOC and HELCOM, and in international oceanographic data system development projects like SeaDataNet and EMODnets, have been vital for Finnish oceanography, too.

Authors Index

A

Adam Marielle 127, 201
Akylas Evangelos....................................160
Alenius Pekka 244
Alvera Aida 39
Amaike Ken-ichi 197
Andreasson Arnold..................................110
Appeltans Ward.......................................155
Arnold Matt...137
Atkins Natalia 32
Azelmat Hamza.......................................116
Azzopardi Joel 231

B

Babau Jean-Philippe..................................26
Bailly Nicolas..155
Balem Kevin 149
Balopoulou Stavroula................................98
Barale Vittorio...227
Baramidze Irine.......................................216
Barth Alexander 39, 203
Bastianini Mauro.....................................104
Beckers Jean-Marie....................................39
Beg Paklar Gordana 45
Beja Joana 145
Belhassen Malika 158
Belov Sergey 94, 126
Beltran Joan Pau 43, 242
Bester Marthan..152
Bilal Abdallahi 26
Bilashvili Kakhaber 216
Bisquay Hervé...54
Błachowiak-Samołyk Katarzyna 106
Blain Peter...118
Blanc Frédérique 178, 218
Bock Steffen..128
Boehme Lars 152
Bornemann Horst 152
Borremans Catherine 73, 116, 209
Bradshaw Elizabeth 66, 162
Brasseur Sophie 152
Breitbach Gisbert 151
Bricher Pip 145
Brink Richard..143
Brosich Alberto 34, 171
Buch Erik 121
Buck Justin..16, 30
Burger Eugene................................108, 139
Burke Nigel 224
Busato Alessandro.....................................21

C

Cabanas Jose Manuel.........................68, 234
Cabrero Agueda 68
Cañellas Tomeu.......................................242
Cano Daniel 64
Capodici Fulvio.......................................231
Cardin Vanessa..34
Carval Thierry 30, 149
Catarino Nuno 77
Caumont Hervé 77
Charrassin Jean-Benoit 152
Ciraolo Giuseppe 231
Claus Simon 176
Coatanoan Christine................................164
Corgnati Lorenzo 133
Coro Gianpaolo...................................23, 51
Costa Dan..152
Crise Alessandro 133
Cusick James...145

D

Dadić Vlado 45, 214
Darroch Louise..16
de Alfonso Alonso-Muñoyerro Marta......149
de Bruin Taco 143

Degraer Steven..199
de Koster Ronald.....................................143
De Leo Francesco 70
Deneudt Klaas...155
De Pooter Daphnis 155
Détoc Jérôme 116, 209
Diggs Steve 145
Dillo Ingrid..13
Diviacco Paolo 21, 123
Drago Aldo..231
Drudi Massimiliano 133
Dugenne Mathilde...................................158
Dwyer Ned 96

E

Edmunds Rorie..13
Eliezer Menashè......................................155
Escudier Romain.....................................242
Evans Gaynor..132
Everett Jason 223

F

Faber Claas..192
Fedak Mike 152
Feistel Susanne..128
Fernández Vicente...................................146
Fichaut Michèle 91, 164, 181
Finney Kim 32
Foglini Federica 70, 135
Forneris Vega 133
Fox Peter 21
Frontera Biel 43, 242
Fruehauf Dan 118
Fujioka Ei..155
Fuller Rob 113

G

Galea Anthony 231
Garau Angels...242
Garcia Sanchez Oscar 123
Gauci Adam 231
Genova Françoise......................................13
Gertman Isaac 56
Giorgetti Alessandra.....34, 39, 133, 155, 171
Glaves Helen 89, 141
Godin Eugeny 56
Godøy Øystein 100
Goffin Annelies 127
Goldstein Philip 155
Gómara Sonia...................................43, 242
Gómez Jose Manuel 70
Gomila Miquel 43, 242
González-Nuevo Gonzalo..........68, 130, 234
González-Pola César 64
Gordon Kathy..162
Gorringe Patrick..............121, 146, 151, 170
Gourrion Jérôme 29, 217
Grande Valentina...............................70, 135
Grandi Alessandro...................................166
Grégori Gérald 158
Griffiths Charly 137
Grosso Nuno 77
Guinet Christophe 152
Guyot Corentin..149
Gvarishvili Tsisana..................................216

H

Haelters Jan...199
Hall Ailsa 152
Hall Mike 137
Hammarklint Thomas 149
Hannant Terry 102
Harcourt Robert 152
Harrison Sandra 80
Harscoat Valérie 186
Hendriksen Gerrit..............................39, 203
Hernandez Francisco........................155, 176
Hibbert Angela 162
Hietala Riikka 244
Hindell Mark...152
Hoenner Xavier 223
Holdsworth Neil..............................168, 206
Huang Thomas 111
Hughes Esther 137

I

Ingerov Andrey 56
Iona Sissy 98, 164, 166, 181

Isaeva Elena 56
Ishii Masayoshi 197
Ivanković Damir 45, 214

J
Jakobsson Martin 195
Jauniaux Thierry 199
Jelavić Dalibor 45
Jevrejeva Svetlana.............................66, 162
Johannessen Truls 189
Jones Steve..189
Johnson Ken...211
Jürgensmann Susanne 128
Juza Mélanie 242

K
Kaartokallio Hermanni............................239
Kaitala Seppo 239
Kande Oumar 26
Karaolia Andria 58
Karigome Yasuhiko.................................197
Kerckhof Francis.....................................199
Khaliulin Alexey 56, 240
Kim Sung-Dae 55
Kizu Shoichi..197
Kokkinaki Alexandra 16, 62
Kobelev Alexander..................................124
Konshin Georgij 183
Kovacs Kit M. 152
Kozyr Alex 139, 145,188
Krassovski Misha....................................188
Krietemeyer Andreas 43, 242
Kruglov Artem 135
Kuragano Tsurane 197

L
Lagring Ruth 127, 201
Lahbib Soumaya 158
Lardner Robin 48
Lavín Alicia...64
Leadbetter Adam.....................................113
Lear Dan...137
Lehtinen Sirpa 239
Leone Rosemarie 70
Leposhkin Oleksandr 135
Lewis Mirtha 155
L'Hours Hervé...13
Libes Maurice 158
Lipizer Marina 34, 155, 171
Lisovskyi Richard 135
Loubrieu Thomas 30, 116
Lowry Roy 62
Lurton Xavier..54
Lydersen Christian 152
Lykiardopoulos Angelos 98

M
Mackay Kevin...155
Makharadze Guranda 216
Mancini Sebastien.............................32, 223
Mangin Antoine 77
Mantovani Carlo 133
Manzano Muñoz Fernando 149
Manzella Giuseppe..........................170, 218
Marelli Fulvio 70
Marinova Veselka....................................149
Marks Karen M. 195
Martinez Inigo...206
Mathieu Grégory 54
Matthews Andy 66, 162
McCann Mike 211
McMahon Clive 152
Meillon Julien 75
Mehrtens Hela 192
Menegon Stefano 104
Merceur Frédéric................................30, 73
Michelsen Christian 102
Michida Yutaka 197
Mieruch Sebastian.....................................60
Mikashavidze Eteri 216
Mikhailov Nikolai.............................94, 126
Minster Jean-Bernard................................13
Minuzzo Tiziano 104
Mokrane Mustapha 13, 80
Moncoiffe Gwenaelle..............................155
Morgado Cristina 206
Morvik Arnfinn 102
Mourre Baptiste 242

Moussat Eric 178, 218
Muñoz Cristian..43
Munoz Raquel...123
Muslim Stipe.....................................45, 214

N

Nakano Toshiya.......................................197
Napolitano Umberto..................................23
Nauw Janine..143
Neprokin Oleksandr................................135
Nexelius Nils...110
Nieuwenhuis Jan 143
Nikolaidis Andreas.............................58, 160
Nikolaidis George 160
Nikolopoulou Stamatina 155
Nolan Glenn 146
Novellino Antonio...................151, 170, 181

O

O’Brien Kevin.................................108, 139
O’Grady Eoin..................................113, 224
Olsen Are 139
Osborne Mike..83
Otero Pablo 68, 130, 234

P

Pagano Pasquale...................................23, 51
Paglialonga Lisa......................................192
Palasov Atanas 178
Panichi Giancarlo......................................51
Park Hyuk-Min 55
Parner Hjalte 37
Partescano Elena 34, 133
Pelleter Ewan 209
Penna Pierluigi 104
Perivoliotis Leonidas 149
Petersen Willi 151
Petit de la Villéon Loïc............................149
Pfeil Benjamin 108, 139, 189
Picard Baptiste 152
Pinardi Nadia 178, 218
Pinto Carlos....................................168, 206
Plant Josh 211
Pobitzer Armin 102
Politi Eirini..77
Ponsoni Leandro 143
Populus Jacques 178
Postlethwaite Clare F. 132
Pouliquen Sylvie 149, 170, 186
Proctor Roger....................32, 118, 145, 223
Provoost Pieter 155

Q

Quimbert Erwann......................................75

R

Rauch Shannon 155
Reinke Susanne.......................................151
Reverdin Gilles 152
Richter Thomas.......................................143
Rickards Lesley 13, 66, 80, 132, 162, 170, 195
Riminucci Francesco...............................104
Ringheim Lid Sjur...................................149
Rodríguez Carmen 64
Roquet Fabien 152
Roubicek Andres.....................................155
Ruiz Villareal Manuel 64, 68, 234
Rújula Miquel Àngel.................................43
Rumes Bob..199
Ruohola Jani..239

S

Sá Sandra 96
Saadatnejad Bard 100
Sagen Helge 102
Sakamoto Carole.....................................211
Sammari Cherif.......................................158
Santinelli Giorgio..............................39, 203
Sarretta Alessandro 104
Sato Kanako 197
Satra Le Bris Catherine.............................75
Scarrott Rory 77
Schaap Dick 89, 91, 164, 170, 174, 181
Schirnick Carsten 192
Schlitzer Reiner.................................60, 164
Schoute Barbara 206
Scory Serge 123, 181, 199
Sebastian Kristian 242

Seglinsh Valdis..183
Serra Imma..43
Shimizu Yugo..197
Simoncelli Simona...................162, 166, 218
Simons Bob...108
Sinquin Jean-Marc 54, 123
Sistermans Wil 143
Smyth Damian 113
Som de Cerff Wim 41
Somavilla Raquel.......................................64
Sorribas Jordi 123
Spears Tobias 94
Springer Pina...192
Stanghellini Giuseppe 104
Stegen Landa Camilla..............................189
Stojanov Yvan 123, 127, 201
Strobbe Francis 201
Strömbergr Patrik....................................110
Stylianou Stavros 48, 58, 160
Suga Toshio...197
Suikkanen Sanna.....................................239
Suzuki Toru...197
Swadling Kerrie 223
Szekely Tanguy 29, 217

T

Tacoma Marten 143
Tamm Susanne..149
Tanhua Toste 139
Tarasova Raisa 231
Tarot Stephane 149
Tel Elena 64, 68, 130, 234
Telszewski Maciej...................................139
Terra Homem Miguel................................77
Thijsse Peter......................................41, 228
Thomas Rob 62
Thorel Arnaud 116
Thyssen Melilotus...................................158
Tikka Kimmo 244
Tintoré Joaquín 43, 149, 242
Tokotoko Jannaï...26
Torres Carlos 155
Tosello Vanessa.......................................209
Toussaint Elise 199
Treguer Mickael...75
Tronconi Cristina 133
Troupin Charles...................39, 43, 149, 242
Tyberghein Lennert.................................127

V

Van de Putte Anton...........................145, 155
Van der Woerd Hans 228
Vandenberghe Thomas.....123, 127, 199, 201
Vandepitte Leen 155, 176
Vanhoorne Bart 155
Vardigan Mary 13
Varshanidze Madona...............................216
Viazilov Evgenii........................85, 124, 126
Viloria Amaia 64
Vincent Cécile 152
Vinci Matteo....................................155, 171
Vitorino Pino Bruno................................176
Voronsov Alexander................................126
Vrignaud Christophe 54

W

Wambiji Nina 155
Watelet Sylvain 39
Waumans Filip 176
Weatherall Pauline 195
Wehde Henning.......................................149
Welsh Jennifer...143
Wernand Marcel......................................228
Wichorowski Marcin..........................96, 106
Williams Hannah.....................................132
Williams Simon.......................................162
Wood Chris 18
Woodworth Philip 162
Worely Steven 108

Y

Yasuda Tamaki 197
Yoritaka Hiroyuki 197

Z

Zammit Audrey 231
Zhuk Elena.....................48, 56, 58, 160, 240
Zodiatis George....................48, 58, 160, 240