



Preservation, discovery, mapping, online processing and mining of oceanographic data at OGS-NODC

Elena Partescano, Alberto Brosich and Stefano Salon



Some numbers on OGS data “archive”

GPS data for seismological studies

- present archive (15 sites, 11yrs, data@1Hz): 3TB/decade
- estimated development of archive (data@20Hz): 60TB/decade

Remote sensing data for risk assessment and hydrological modelling

- Archive of image files O(10) TB – private data, not archived

Amount of numerical data for seismological modelling (input/output)

- O(10) GB for 1 run, about 1TB for project (i.e. several runs with different inputs)
- COMPUTATIONS PRODUCED AT FERMI@CINECA

Amount of numerical data for physical-biogeochemical modelling (input/output)

- Climate n-Ensemble: (1yr, 2.7M nodes: 50GB) X 200yrs X n = 10TB x n (dn/dt>0)
- Operational: (17days, 3.3M nodes: 7GB zip) X 2 X 52weeks X 6yrs = 4.4TB
- Increase horiz. resolution (1/8° → 1/16°) → 4x amount
- COMPUTATIONS PRODUCED AT FERMI@CINECA

Amount of observation data from OGS-NODC

- Data+metadata from field observations (physical-chemical data) = 0.5 TB



OGS - NODC

- Members
 - A. Giorgetti, A. Altenburger, A. Brosich, M. Lipizer, E. Partescano, M. Vinci, (M. Eliezer)
- Principal activities:
 - Archiving and management of **hystorical data**
 - Archiving and management of **Real-Time data**

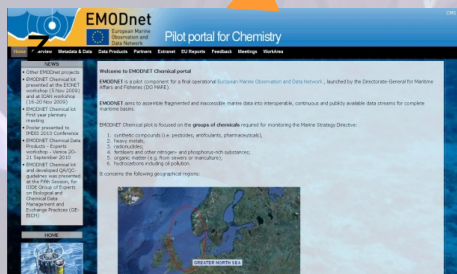
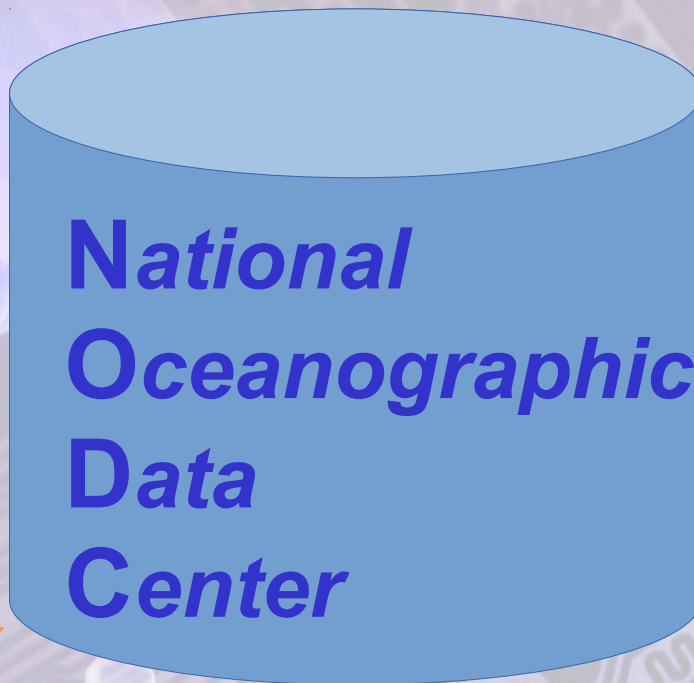


Current situation

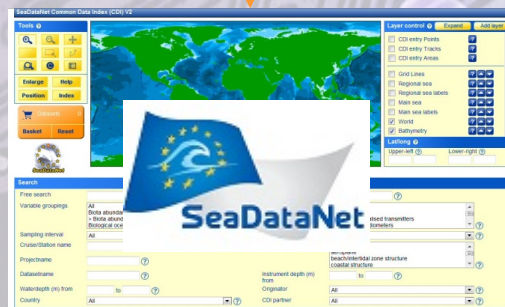
- Meta-data and data are stored in a **relational database**:
 - Historical Data (Oracle)
 - > 210 million of measurements (one table)
 - > 150 tables (especially meta-data)
 - Since 1889 until 2012
 - Mediterranean Sea and Black Sea
 - Mainly physical-chemical data
 - Real-Time Data (PostgreSQL)
 - Buoy MAMBO (Miramare)
 - FVG Regional Network (Protezione civile FVG)



Current situation



www.emodnet-chemistry.eu

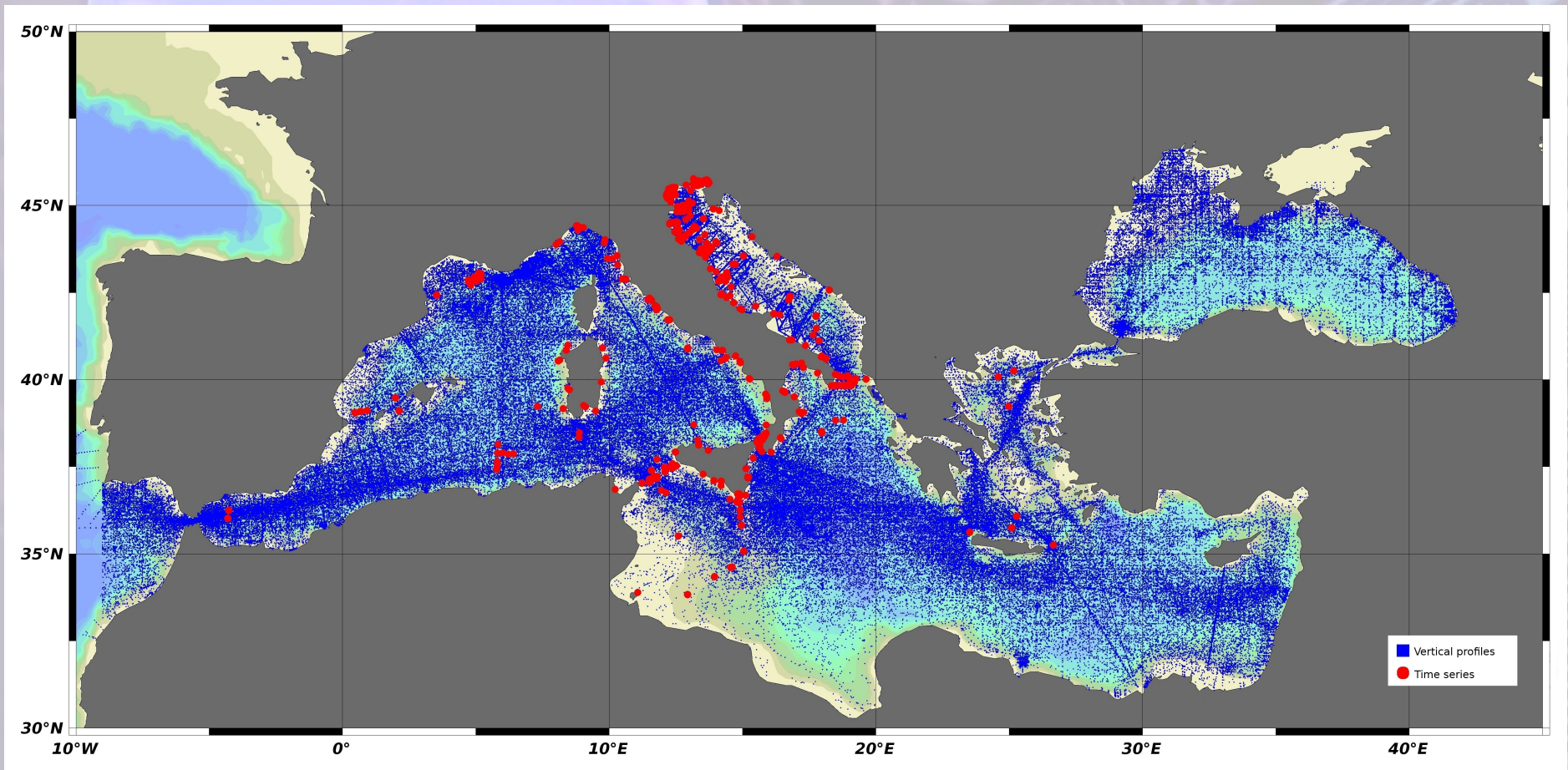


[SeaDataNet CDI Service](http://www.seadatanet.eu)

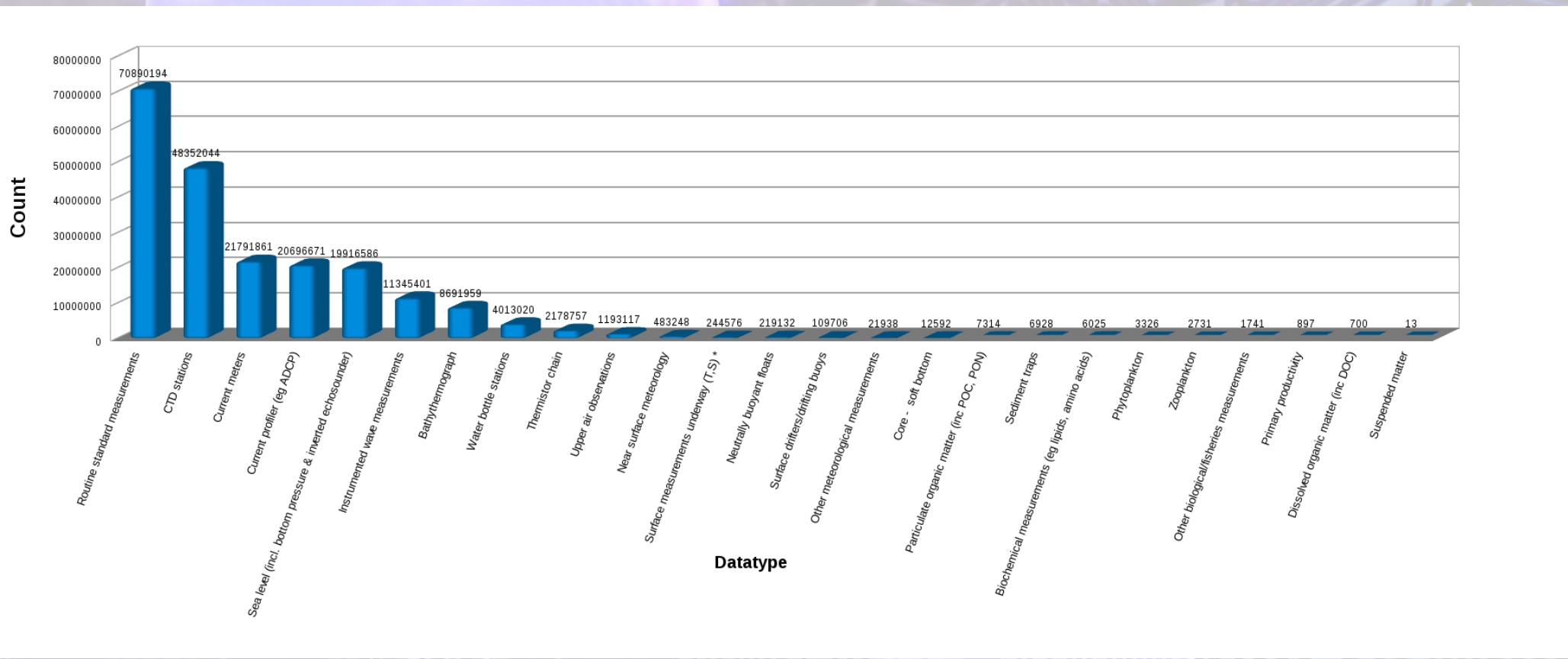


www.emodnet-hydrography.eu

Geographical distribution



Datatypes





Data features

- Database: “middle” size
- Unrepeatable data
- High cost of acquisition
- Abundance of meta-data
- High number of variables archived (>350)
- Data access especially read-only: analytical (OLAP), non transactional (OLTP)



User services

- Discovery
 - Searching data using metadata
- Mapping
 - Space/temporal mapping
- Online processing
 - On-Line Analytical Processing (OLAP)



Discovery

Or specify some search criteria...

Coordinates (spatial and temporal)

N *
W * E *
S *

From: *
To: *

Parameters

- AAMD - ALPHA-AMINOPEPTIDASE (μmol/(L-h))
- ABCP - ALPHA, BETA CAROTENES (milligram/m3)
- AGCO - SMALL AGGREGATE CONCENTRATION FLOCS,MACROFLOCS OR STRINGERS (type)
- AGT1 - AGGREGATE TYPE 1: FLOCS (presence)
- AGT2 - AGGREGATE TYPE 2: MACROFLOCS (presence)
- AGT3 - AGGREGATE TYPE 3: STRINGERS (presence)
- AGT4 - AGGREGATE TYPE 4: RIBBONS (presence)

Datatypes

- B01 - Primary productivity
- B06 - Dissolved organic matter (inc DOC)
- B08 - Phytoplankton
- B09 - Zooplankton
- B71 - Particulate organic matter (inc POC, PON)

Project

- AAS
- ADRIA
- ADRICOSM

Institutes

- ALBANIA - Akademia e Shkencave, Tirane
- ALBANIA - Hydrometeorological Institute, Tirana
- ALGERIA - ISMAL, Halieutique, Laboratoire de Plankton, Alger

Ships

- ALGERIA - MOHAMED SEDIK BENYAHIA
- BAHAMAS - MARISKA
- BAHAMAS - VENEZIA

Results

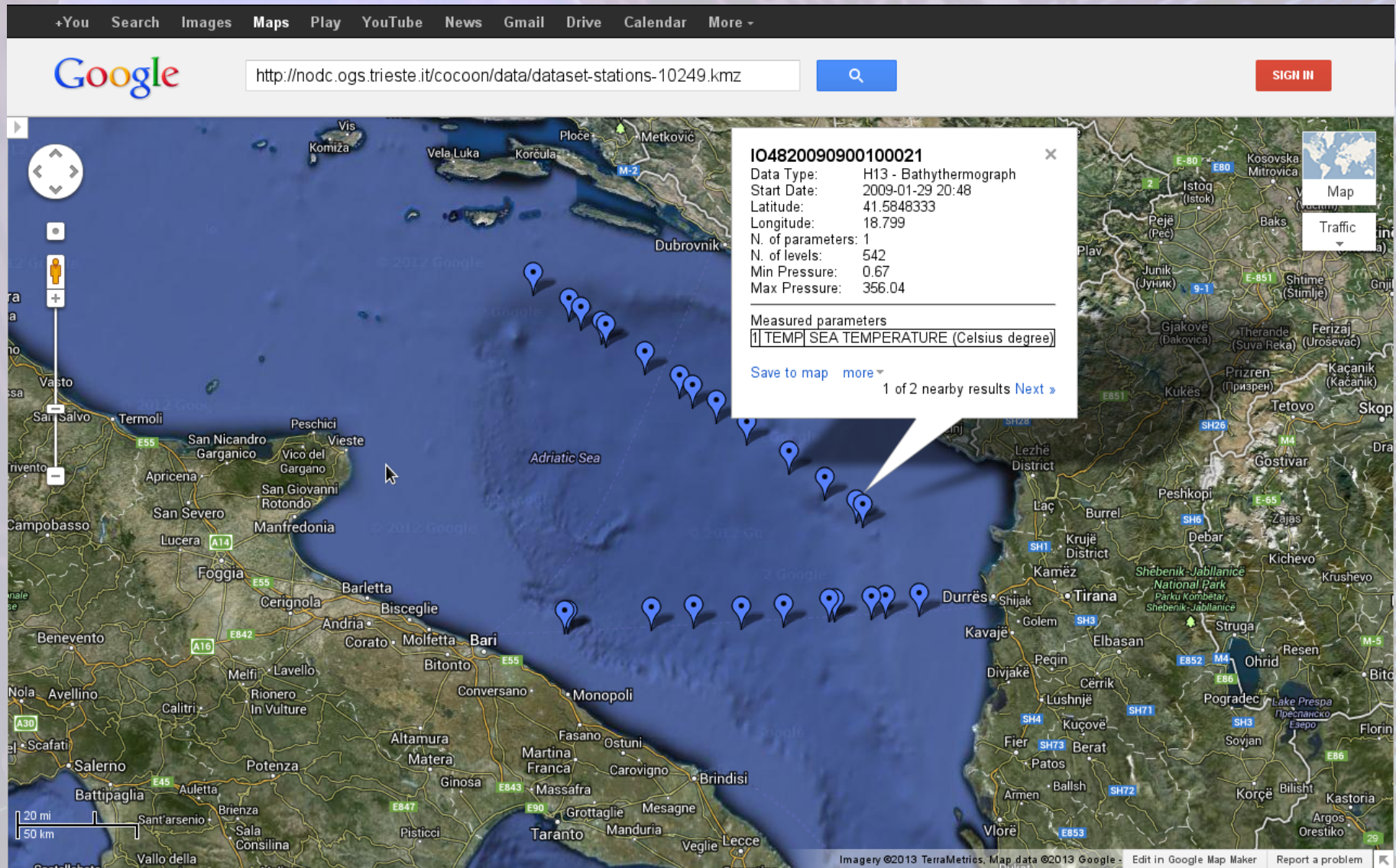
No data selected

You can download only public available data. Restricted data can be requested contacting us by email (nodc@ogs.trieste.it).

All information, data, ..., are freely available and free of charge under the owner (producer) policy conditions and their wide use for no profit science and education is encouraged. [Copyright @](#)



Mapping





OLAP



NATIONAL OCEANOGRAPHIC DATA CENTER

NODC Metadata Query uses Mondrian OLAP



Datatype	Sea Area	Project	Policy	Measures	
				NStations	NMeasures
-All Datatypes	+All Sea Areas	+All Projects	+All Policies	367,514	210,190,477
B01 - Primary productivity	+All Sea Areas	+All Projects	+All Policies	187	897
B06 - Dissolved organic matter (inc DOC)	+All Sea Areas	+All Projects	+All Policies	15	700
B08 - Phytoplankton	+All Sea Areas	+All Projects	+All Policies	334	3,326
B09 - Zooplankton	+All Sea Areas	+All Projects	+All Policies	357	2,731
B71 - Particulate organic matter (inc POC, PON)	+All Sea Areas	+All Projects	+All Policies	548	7,314
B72 - Biochemical measurements (eg lipids, amino acids)	+All Sea Areas	+All Projects	+All Policies	281	6,025
B73 - Sediment traps	+All Sea Areas	+All Projects	+All Policies	98	6,928
B90 - Other biological/fisheries measurements	+All Sea Areas	+All Projects	+All Policies	175	1,741
D01 - Current meters	+All Sea Areas	+All Projects	+All Policies	1,228	21,791,861
D05 - Surface drifters/drifting buoys	+All Sea Areas	+All Projects	+All Policies	38	109,706
D06 - Neutrally buoyant floats	+All Sea Areas	+All Projects	+All Policies	24	219,132
D09 - Sea level (incl. bottom pressure & inverted echosounder)	+All Sea Areas	+All Projects	+All Policies	189	19,916,586
D71 - Current profiler (eg ADCP)	+All Sea Areas	+All Projects	-All Policies	30,435	20,696,671
			Licensed	28,236	20,581,229
			Public	2,199	115,442
D72 - Instrumented wave measurements	+All Sea Areas	+All Projects	+All Policies	113	11,345,401
G04 - Core - soft bottom	+All Sea Areas	+All Projects	+All Policies	70	12,592
H09 - Water bottle stations	+All Sea Areas	+All Projects	+All Policies	93,579	4,013,020
H10 - CTD stations	+All Sea Areas	+All Projects	+All Policies	71,986	48,352,044
H13 - Bathythermograph	+All Sea Areas	+All Projects	+All Policies	167,264	8,691,959
H71 - Surface measurements underway (T,S) *	+All Sea Areas	+All Projects	+All Policies	5	244,576

Find: < Previous > Next Highlight all Match case



Performances

- Read-only (e.g., pre-input checking in real-time insert operations)
- Users services:
 - Mapping
 - OLAP



Future developments

- GIS services:
 - Spatial extensions into the database;
 - Implement WMS services (layers distribution)
 - Manage optional layers (bathymetry, coast lines, sea areas,..)
- Non-numerical data
- OLAP on data (not only on metadata)
- Database NoSQL (es. SciDB)



Thanks for your attention!

<http://nodc.ogs.trieste.it>