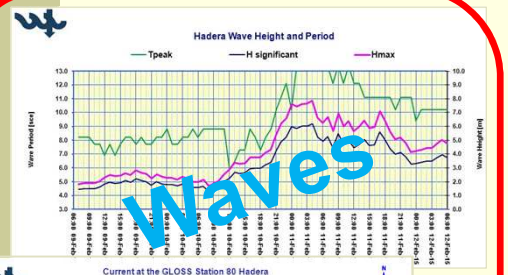
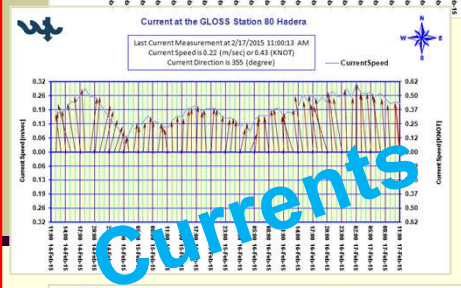
Implementation of SeaDataNet strategy in development of Israel Marine Data Center (ISRAMAR) and PERSEUS marine cast database

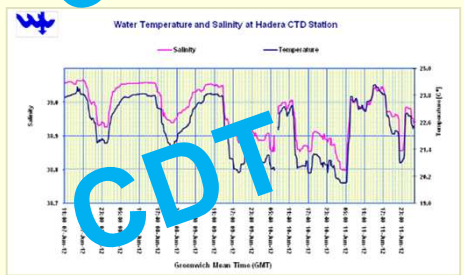
how SeaDataNet has influenced on data management methods: improvements-benefits, changes, what more could be done

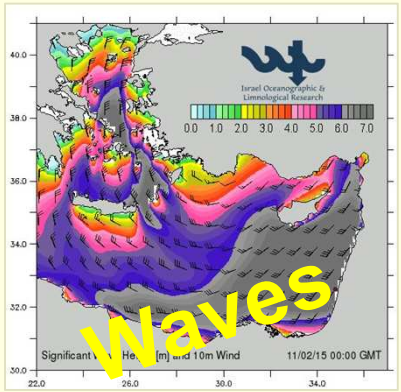
Isaac Gertman, Eyal Greengrass

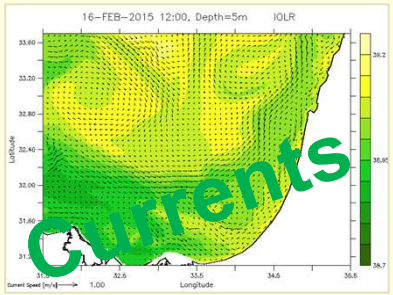
Israel Marine Data Center (ISRAMAR)

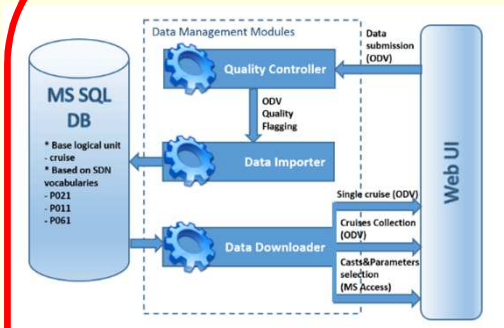


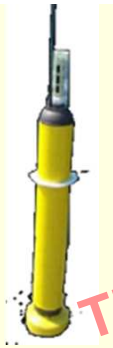


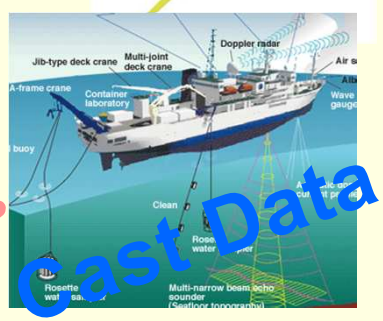












Major databases with historical observations

- **Cast database**

  contains vertical observed profiles of physical, chemical and biological data
- **TimeSeries databases**

  contains time series of physical, chemical and biological data observed at fixed stations
- **Biological data base**

  contains information regarding marine species in the South Eastern Mediterranean
- **Heavy metal database (in process of reorganization)**

  contains information regarding concentrations of heavy metals in organisms and sediments from Israel's Mediterranean coastal waters.
- **Dead Sea Hydrometeorological data**

  contains data from open sea meteorological and hydrological stations

Result of SeaDataNet influence in brief: **Redevelopment of all databases**.

- From a scientist oriented database to a regional multi parametric DBs connected with European Distributed Network of Oceanographic data.

Mediterranean Cast Data base

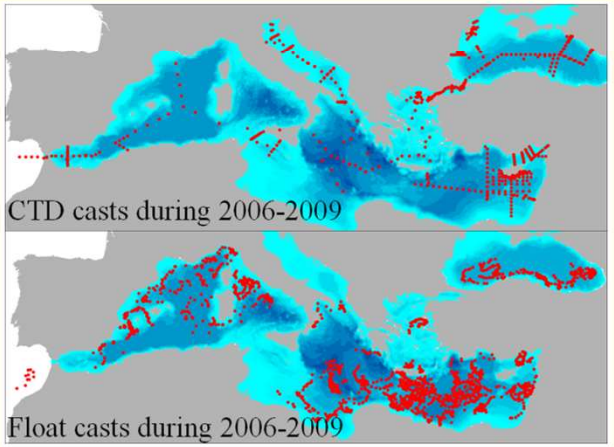
- An oceanographic cast or oceanographic profile is a logical unit of physical and chemical parameters of sea water obtained from the water body beneath the ship by in situ measurements or by analysis of water samples.
- To provide in situ measurements and water samples oceanographers used oceanographic bottles equipped with reversing thermometers. In the 1970s CTD probes with carousel water samplers replaced bottles.

From MEDAR/MEDATLAS collection to PERSEUS Cast DB

Extractions from Public available Data bases:

Soviet POEM WOD09 NATO TU BS DYFAMED SESAME

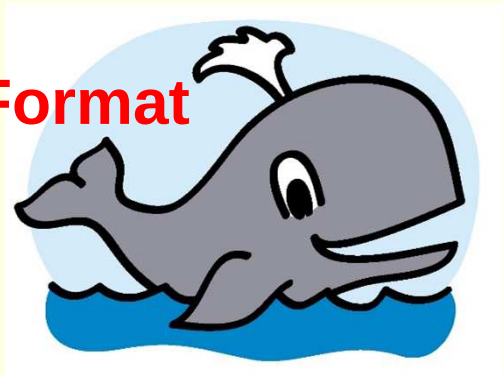
- Import to MS ACCESS all MEDATLAS cast data
- Periodical import of rescued historical data
- On–line submission interface for remote data providers
- GIS like query interface
- During the period 2006-2011 MEDACC was significantly extended by rescued historical data and by data observed within the SESAME framework.
- Continuation of the DB development and data collection in framework PERSEUS project (2012-2015)

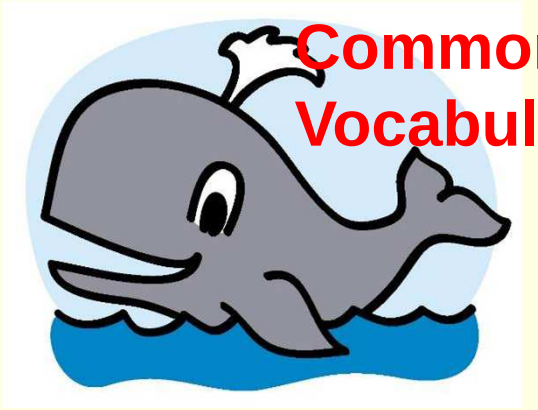


Why it was necessary to develop PERSEUS CAST DB

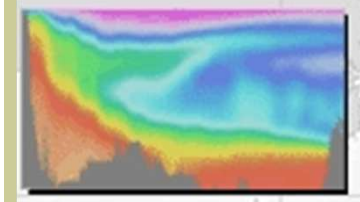
- World data centers (NOAA, SEADATANET) support long-term archiving and distribution of global data.
- Scientists and stakeholders who are not experienced in data management may find it hard to handle the online interfaces;
- Data access may be complicated and time consuming.
- Resulting data will often require further processing for aggregation and harmonization.
- A project oriented, regional DB can be easier in data import, quality control, selection and in organization harmonized export.
- BUT all improvement are possible only when two WHALES are developed:

Data Format





Common Vocabularies



Ocean Data View

Sitemap Imprint

AWI

Reiner **Schlitzer,** http://odv.awi.de/

Ocean Data View (ODV) is a software package for the interactive exploration, analysis and visualization of oceanographic and other geo-referenced profile, time-series, trajectory or sequence data. ODV runs on Windows (7, Vista, XP, 9x, Me, NT, 2000), Mac OS X, Linux, and UNIX (Solaris, Irix, AIX) systems. ODV data and configuration files are platform-independent and can be exchanged between different systems.

Use ODV to produce:

- property/property plots of selected stations,
- scatter plots for sets of stations,
- color sections along arbitrary cruise tracks,
- color distributions on general isosurfaces,
- temporal evolution plots of tracer fields,
- differences of tracer fields between repeats,
- geostrophic velocity sections,
- animations (3MB).

COMMON VOCABULARIES

Use of common vocabularies in all metadatabases and data formats is an important prerequisite towards consistency and interoperability. Common vocabularies consist of lists of standardised terms that cover a broad spectrum of disciplines of relevance to the oceanographic and wider community. Using standardised sets of terms solves the problem of ambiguities associated with data markup and also enables records to be interpreted by computers. This opens up data sets to a whole world of possibilities for computer aided manipulation, distribution and long term reuse.

http://seadatanet.maris2.nl/v_bodc_vocab_v2/welcome.asp

				Version	Members
P01		BODC Parameter Usage Vocabulary	BODC PUV	462	30578
P02	View	SeaDataNet Parameter Discovery Vocabulary	SeaDataNet PDV	84	447
P03	View	SeaDataNet Agreed Parameter Groups	SeaDataNet APG	21	64
P04		Global Change Master Directory Science Keywords V5	GCMD Science Keywords V5	4	1413
P05		International Standards Organisation ISO19115 Topic Categories	ISO Topic Categories	0	19
P06		BODC data storage units	BODC units	55	288

P02

DOCU	Dissolved organic carbon uptake and production in the water column	
DGPW	Dissolved oxygen concentration parameters in sediment pore waters	
DOXY	Dissolved oxygen parameters in the water column	

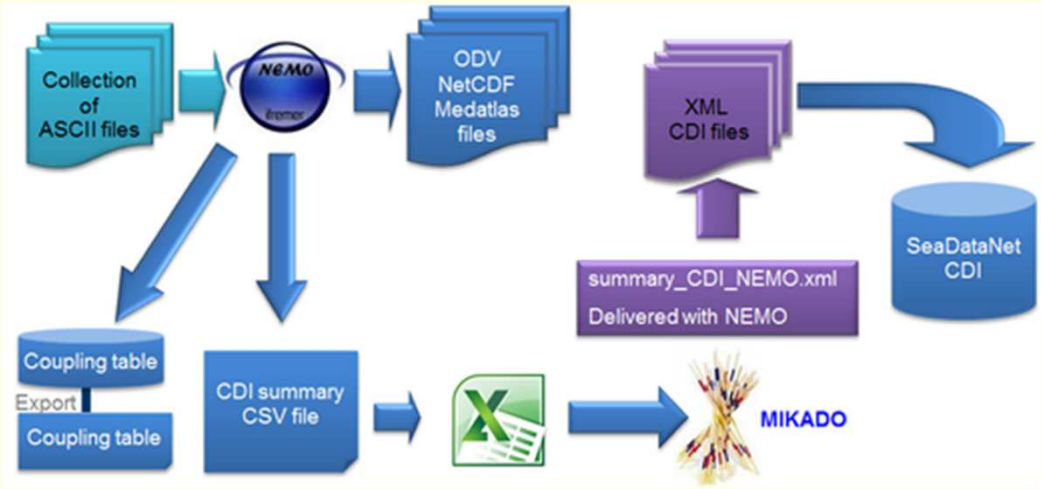
p01	Conceptid	Pref label
	CODZZZZZ	Chemical oxygen demand {COD} per unit volume of the water body
	DOKGWITX	Concentration of oxygen {O2} per unit mass of the water body [dissolved plus reactive particulate phase] by Winkler titration
	DOSDPR01	Concentration standard deviation of oxygen {O2} per unit volume of the water body [dissolved plus reactive particulate phase] by in-situ Beckmann probe
	DOXMSDXX	Concentration standard deviation of oxygen {O2} per unit mass of the water body [dissolved plus reactive particulate

PERSEUS CAST DB

- The DB conforms to SeaDataNet standards such as the use the Common Vocabularies and adaptation of ODV format including latest SDN modifications.
- Contains a bulk of historical oceanographic observations in the Mediterranean and the Black Seas.
- Submission of the data can be done by any partners using an on line interface **within a short time after the observations** (facilitates the data transfer within the project).
- A data quality control, based on a standalone application with easy manipulated criteria.

How to convert Sea Bird cast data to ODV format

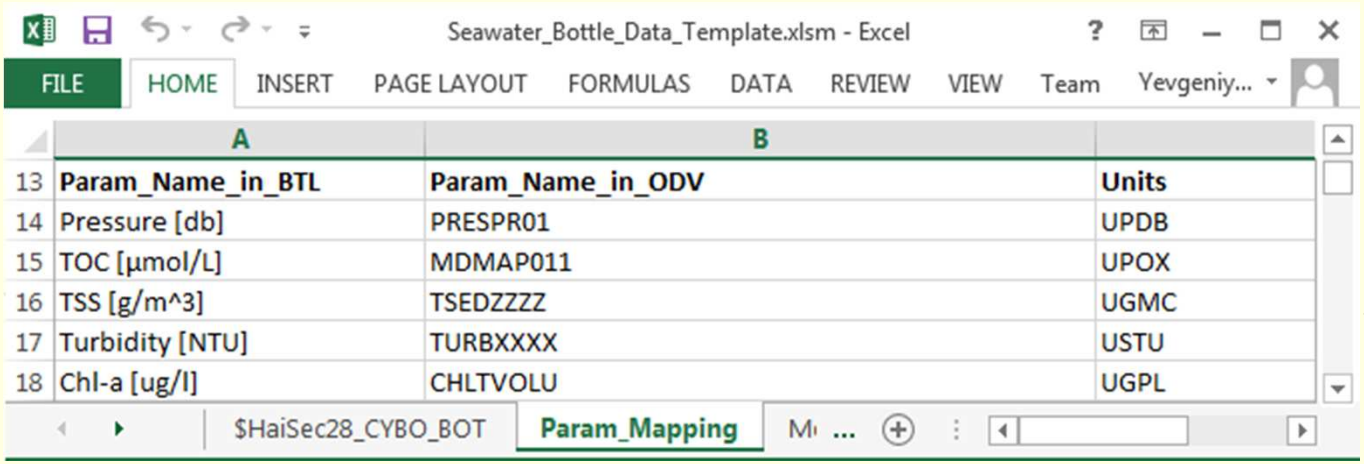
- After processing according to Sea Bird software instructions one has two files with data in engineering units: xxx.cnv and xxx.btl
- These files can be converted in ODV format, which is the European standard for presentation of cast data.
- The SeaDataNet recommended converter is NEMO software



Conversion of Sea Bird cast data to ODV format:
EXCEL table to ODV format

- Generating .cnv and .btl files by SeaBird
- Import .cnv / .btl files into Excel template
- Export from excel to ODV

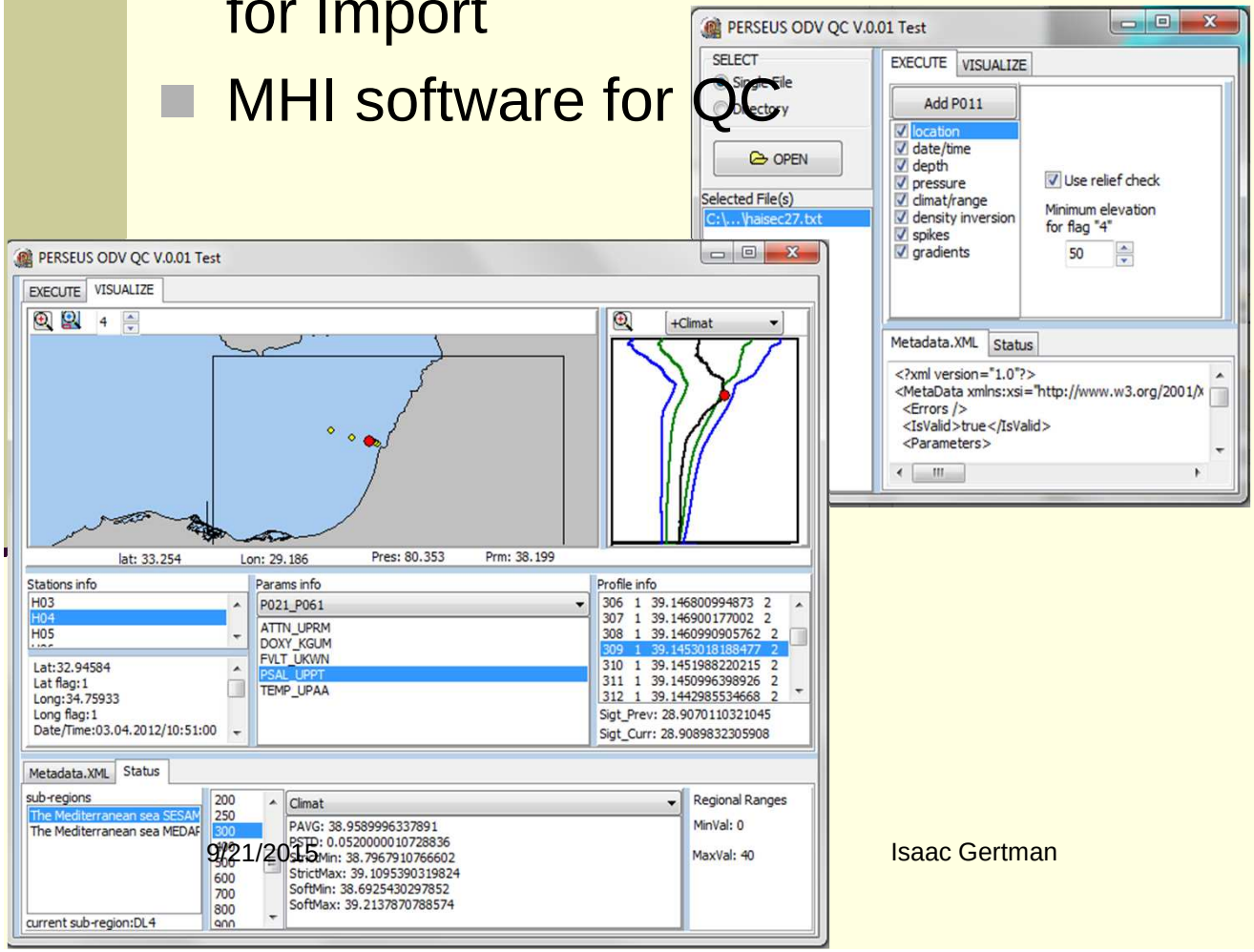
Download link:
http://isramar.ocean.org.il/isramar_data/TextTemplates/Convertors/Data_Template.zip

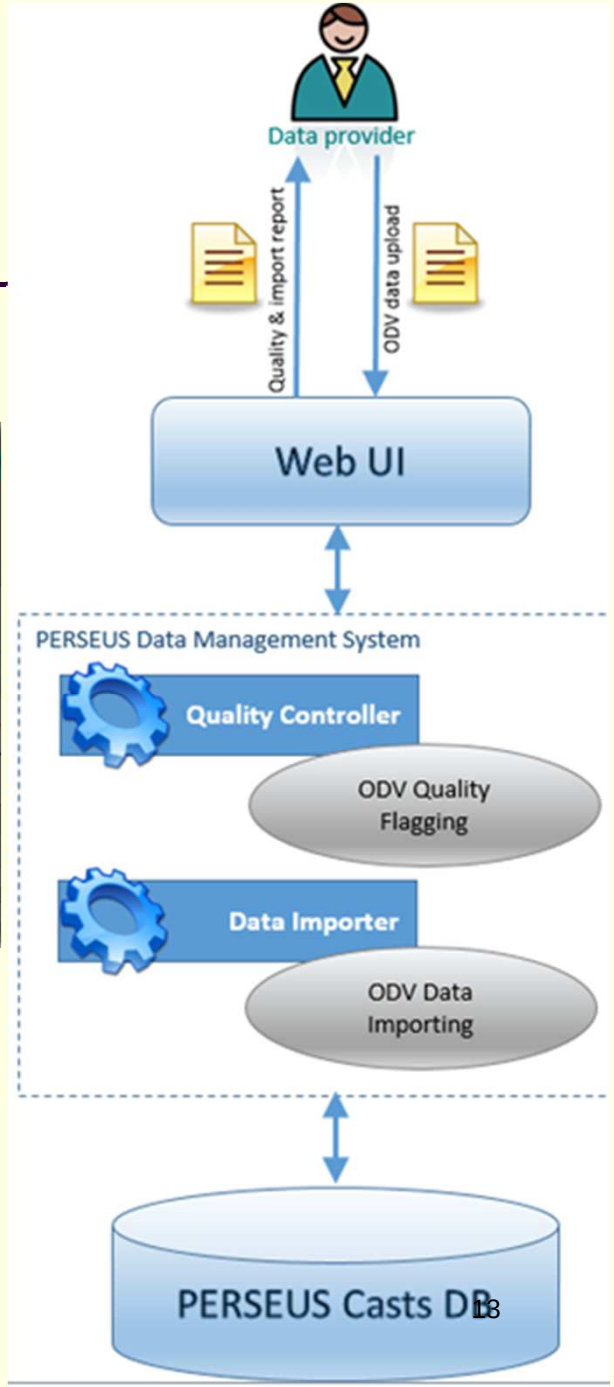


Add mapping to P01

On-line Import & Quality Control

- ODV format (short and full versions) for Import
- MHI software for QC





ODV generic format ASCII, Tab delimited

First Reiner Shlitzer version

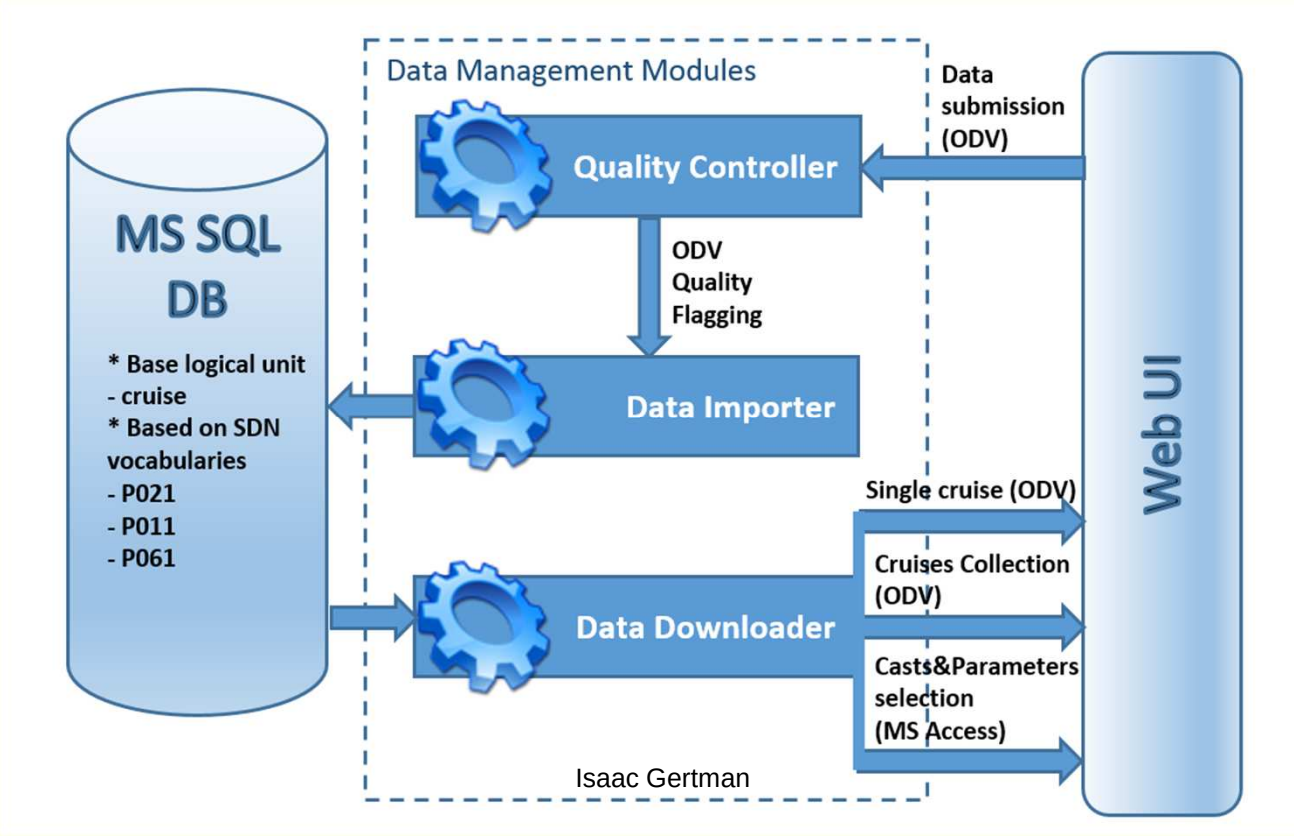
Cruise	Station	Type	mon/day/yr	hh:mm	Lon (°E)	Lat (°N)	Bot. Depth [m]	Pres[db]	QF	Temp.[deg]	QF	PSAL[psu]	QF	DOC2
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	0	1	29.3181	1	39.4764	1	186.731	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	1	1	29.2624	1	39.4756	1	186.852	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	2	1	29.2722	1	39.4757	1	186.892	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	3	1	29.1578	1	39.4691	1	186.995	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	4	1	28.9984	1	39.4716	1	186.569	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	5	1	28.9801	1	39.4691	1	186.245	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	6	1	28.9717	1	39.468	1	187.253	1	0
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	7	1	28.9703	1	39.4679	1	187.426	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	8	1	28.9673	1	39.4675	1	187.248	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	9	1	28.9626	1	39.467	1	187.055	1	0
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	10	1	28.9603	1	39.4664	1	187.115	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	11	1	28.9586	1	39.466	1	187.011	1	
SESIL02	h150/1426	C	09/07/2008	08:57	34.86717	32.893	162	12	1	28.9559	1	39.4658	1	187.361	1	

SDN version

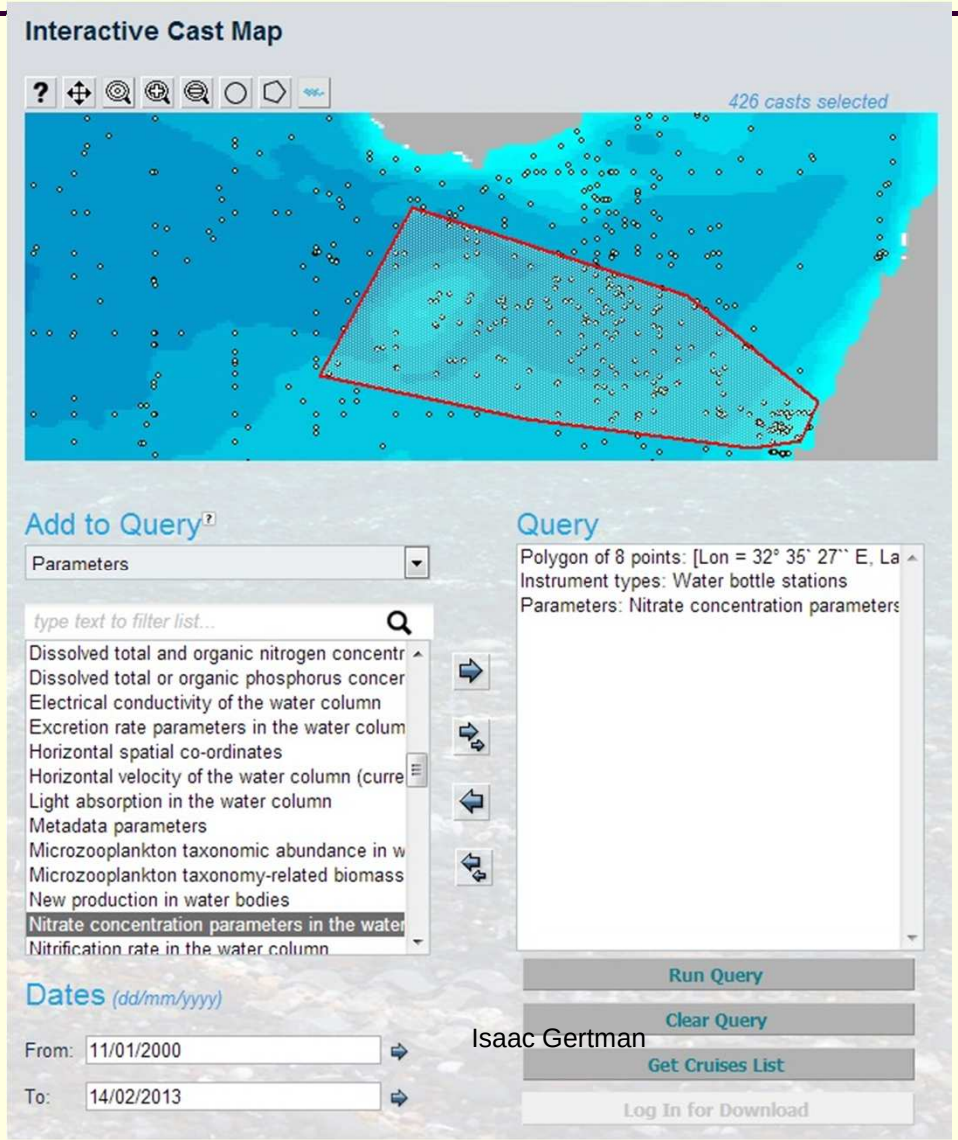
```
//<Version>ODV Spreadsheet V4.0</Version>
//<CreateTime>6/8/2014 11:59:40 AM</CreateTime>
//
//SDN_parameter_mapping
//<subject>SDN:LOCAL:PRES</subject><object>SDN:P01::PRESPR01</object><units>SDN:P06::UPDB</units>
//<subject>SDN:LOCAL:ATTNZS02_UPRM</subject><object>SDN:P01::ATTNZS02</object><units>SDN:P06::UPRM</units>
//<subject>SDN:LOCAL:DOXYSU02_KGUM</subject><object>SDN:P01::DOXYSU02</object><units>SDN:P06::KGUM</units>
//<subject>SDN:LOCAL:FLUOZZZZ_UKWM</subject><object>SDN:P01::FLUOZZZZ</object><units>SDN:P06::UKWM</units>
//<subject>SDN:LOCAL:PSALST01_UPPT</subject><object>SDN:P01::PSALST01</object><units>SDN:P06::UPPT</units>
//<subject>SDN:LOCAL:TEMPS901_UPAA</subject><object>SDN:P01::TEMPS901</object><units>SDN:P06::UPAA</units>
//
Cruise    Station    Type    yyyy-mm-ddThh:mm:ss.sss    Longitude [degrees_east]    Latitude [degrees_north]    LOCAL_CDI_ID    EDMO_code    Bot. Depth [m]
HaiSec25    H01    C    2011-03-27T08:30:00.00    34.92433    32.898    963    51    2    1    0.1011    1    224.834    1    0.063258    1    38.9819
                                3    1    0.1008    1    224.477    1    0.059116    1    38.9804    1    18.4472    1
                                4    1    0.1005    1    224.748    1    0.054633    1    38.9707    1    18.3085    1
                                5    1    0.1017    1    225.149    1    0.053765    1    38.973    1    18.1455    1
                                6    1    0.1046    1    224.717    1    0.054145    1    38.9738    1    18.1233    1
                                7    1    0.107    1    224.199    1    0.056015    1    38.9741    1    18.1066    1
                                8    1    0.1092    1    224.336    1    0.059007    1    38.9744    1    18.0953    1
                                9    1    0.1104    1    224.095    1    0.062631    1    38.9747    1    18.0907    1
                                10    1    0.111    1    224.414    1    0.062222    1    38.9749    1    18.0881    1
                                11    1    0.1115    1    224.801    1    0.063943    1    38.9751    1    18.0847    1
                                12    1    0.1113    1    224.724    1    0.066946    1    38.9745    1    18.0828    1
```

Cast DB management system

- Casts are grouped both by physical cruise and by instruments (CTD, Bottles, Floats etc.,).
- One group forms a base logical unit.



On-line query builder for data selection and download



Query result

Selected Cruises List

Back to cruises selection on map

Cruise Name	Start Date	End Date	Country	Ship Name	Availability	Download
POEM05-AS87 (IBM-I)_BOT	31/08/1987	18/09/1987	Italy	Bannock		
POEM05-AS87 (ITT-I)_BOT	31/08/1987	18/09/1987	Italy	Bannock		
POEM05-AS87 (IRPEM-I)	31/08/1987	18/09/1987	Italy	Bannock		
03906	31/08/1987	17/09/1987	Italy	Unknown		

0 cruises selected for download of 250 allowed.

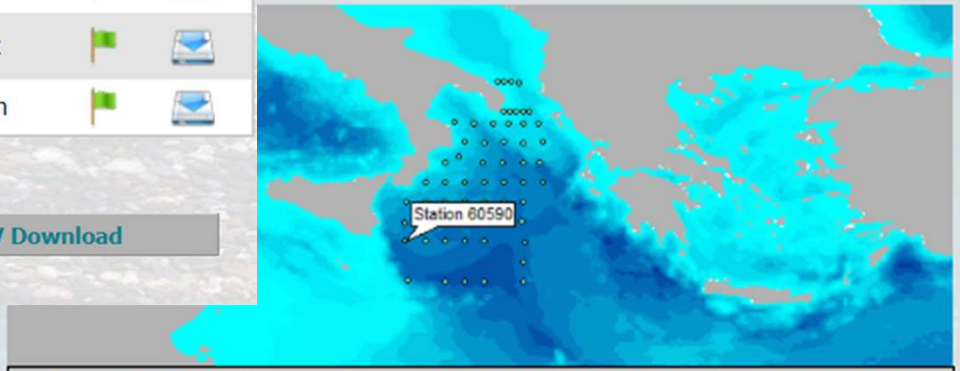
Include SDN data columns in ODV.

Agregate ODV Download

POEM05-AS87 (ITT-I)_BOT

3, 1987

POEM 05 AS87 (ITT/IBM/IRPEM)



Start: Aug 31, 1987 | Sep 18, 1987 14:10:00, Station: 60590 | End: Sep 18, 1987

Cruise Metadata

Instrument type:	Water bottle stations	**Project:**	POEM
Country:	Italy	**Ship:**	Bannock
Data provider:	*unknown*	**Contact:**	*unknown*
Data accessibility:	unrestricted	**Download:**	

Cruise Measured Parameters

Code P021	Parameter	Casts
AHGT	Vertical spatial coordinates	59
AMON	Ammonium concentration parameters in the water column	33
DOXY	Dissolved oxygen parameters in the water column	106
NTRA	Nitrate concentration parameters in the water column	35
NTRI	Nitrite concentration parameters in the water column	34
PHOS	Phosphate concentration parameters in the water column	33
PSAL	Salinity of the water column	59
SLCA	Silicate concentration parameters in the water column	35
TEMP	Temperature of the water column	59

Data Availability

According to PERSEUS Data Policy and PERSEUS Publication Strategy, the access to each dataset is defined by the data provider during the data submission procedure. The following data availability flags are implemented according to SeaDataNet Data Access Restriction Policies (L081) vocabulary.

Unrestricted *(free)*

The data are freely available to anybody and may be used for any purpose. Usage acknowledgement may be required.

Most of the cruises are exported from public available databases MEDATLAS 2002; MATER; WODB0; CORIOLIS; ICES.

By negotiation *(restricted)*

The metadata regarding the cruises is open but data can be obtained on a case-by-case basis through negotiation with data provider.

Organization *(for partners)*

The datasets are available to PERSEUS partners only.

Data export after selection

- Single ODV file with data from one cruise and one instrument (It can be loaded into ODV by drag and drops)
- Aggregated ODV files (Up to 250 cruises as zip file. It can be loaded into ODV by Import SDN spreadsheet)
- Casts with parameters and units homogenization in form of MS ACCESS DB, up to 100,000 casts.

MS ACCESS DB

MS Access Database Download

- **To finish download request submit, click 'Submit Download Request' button.**
- To change the selection or cancel the download process, please return to the interactive cast map.
- To find used codes definition, download Standards Vocabularies MS ACCESS file or use online SDN Common Vocabularies interface.

Query

Cruises groups	METEOR 1966
Cruises groups	METEOR 1967
Cruises groups	METEOR 1970
Cruises groups	METEOR 1971
Cruises groups	METEOR 1973
Cruises groups	METEOR 2001
Cruises groups	METEOR 25C 1993
Cruises groups	Meteor 84
Cruises groups	METEOR-1974 1974
Ships	METEOR
Ships	Meteor (built 1964)
Ships	Meteor (built 1986)

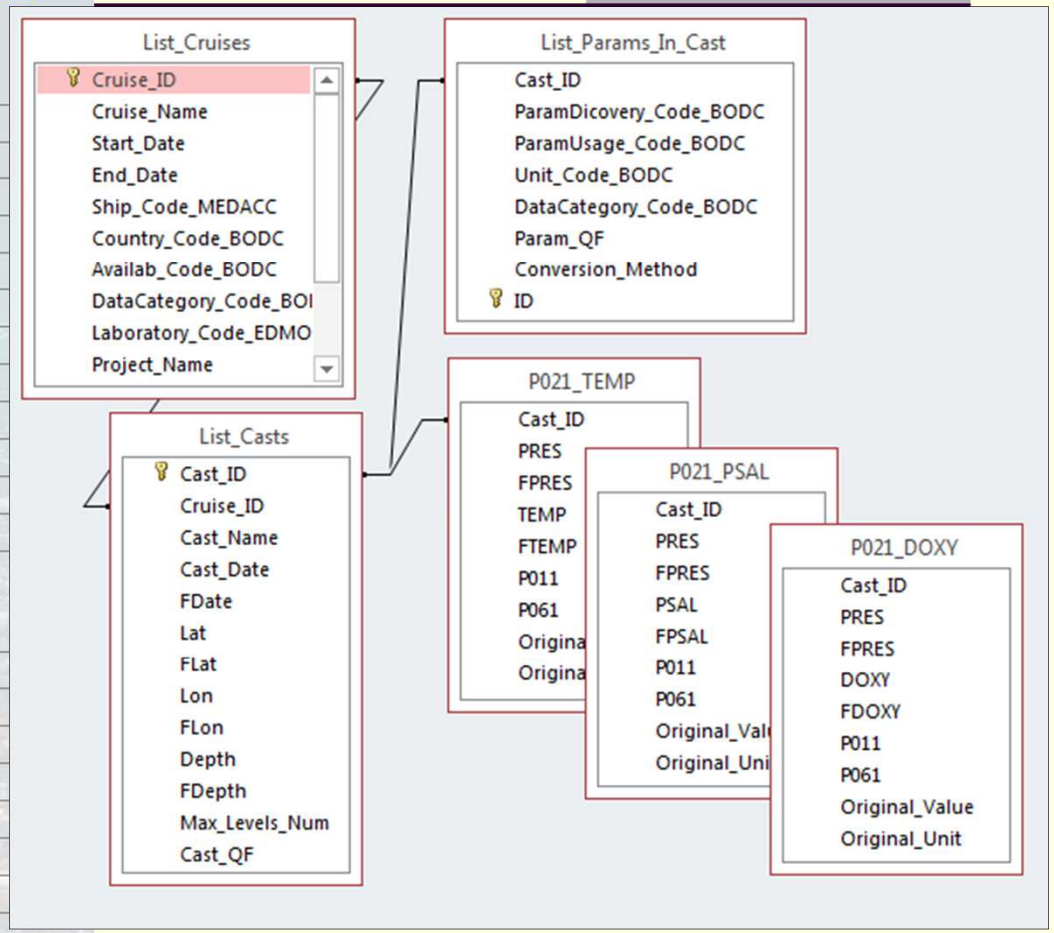
Download Request Analysis

Submission Result	Your request was successfully submitted! You will receive an email.
User Permission	Any data download
User Active Requests	Current request is 1 of 3 maximum allowed
Query Text Validation	Valid for download
Casts Number	253
Parameters Number	33
Approximate File Size	00.478 MB

After the request is processed, you will receive an information email with link for file download.
The download link will be also avaliable from your downloads management area.

Maximum number of active (not processed yet) download requests per user is 3. You can submit 2 more download requests.

Relationships in the exported MS ACCESS DB



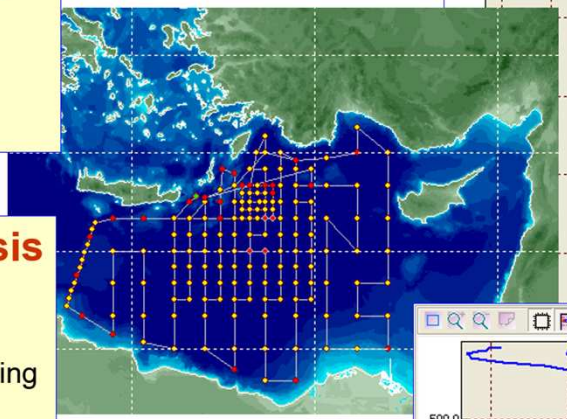
Data Visualizer for the MS ACCES exported data

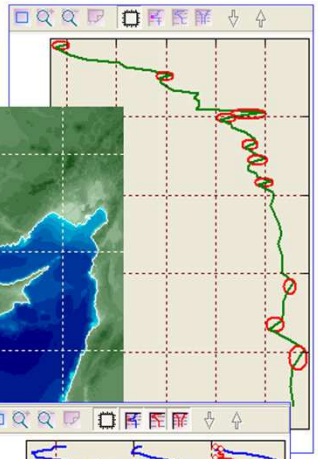
External Quality Control Wizard

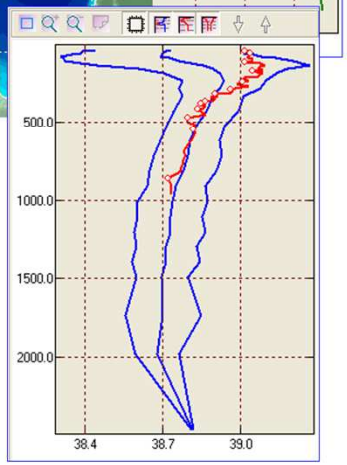
- Metadata inspection
- Comparison with available climatology
- Hydrostatic stability control
- Spikes control

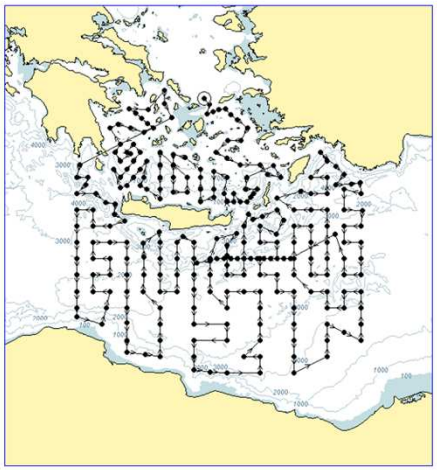
External Oceanographic Analysis Tools

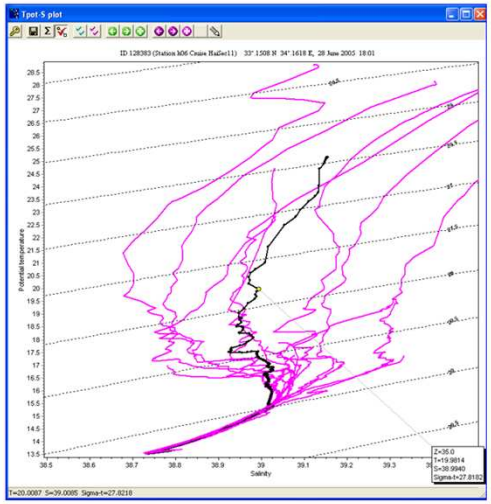
- Sophisticated data profile charting
- SURFER based vertical and horizontal mapping
- Calculation oceanographic parameters



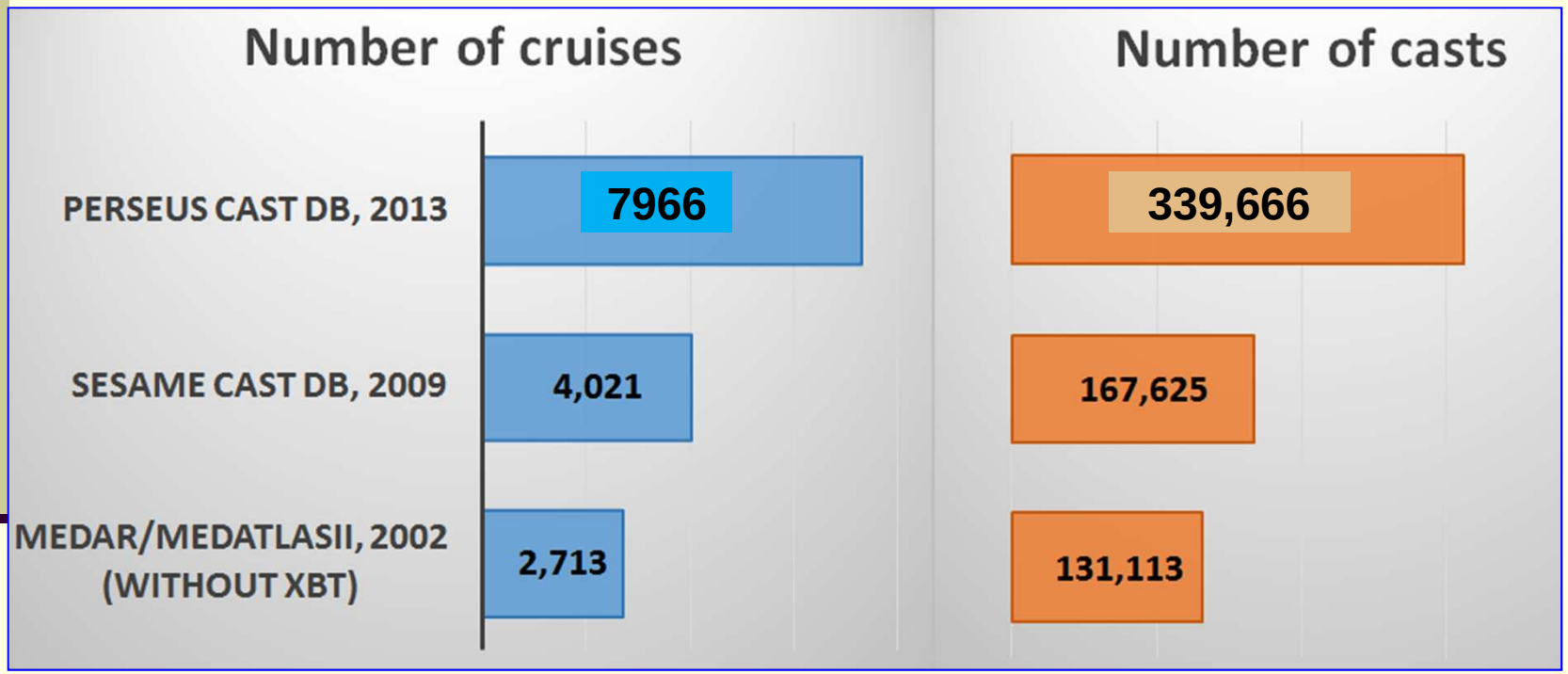




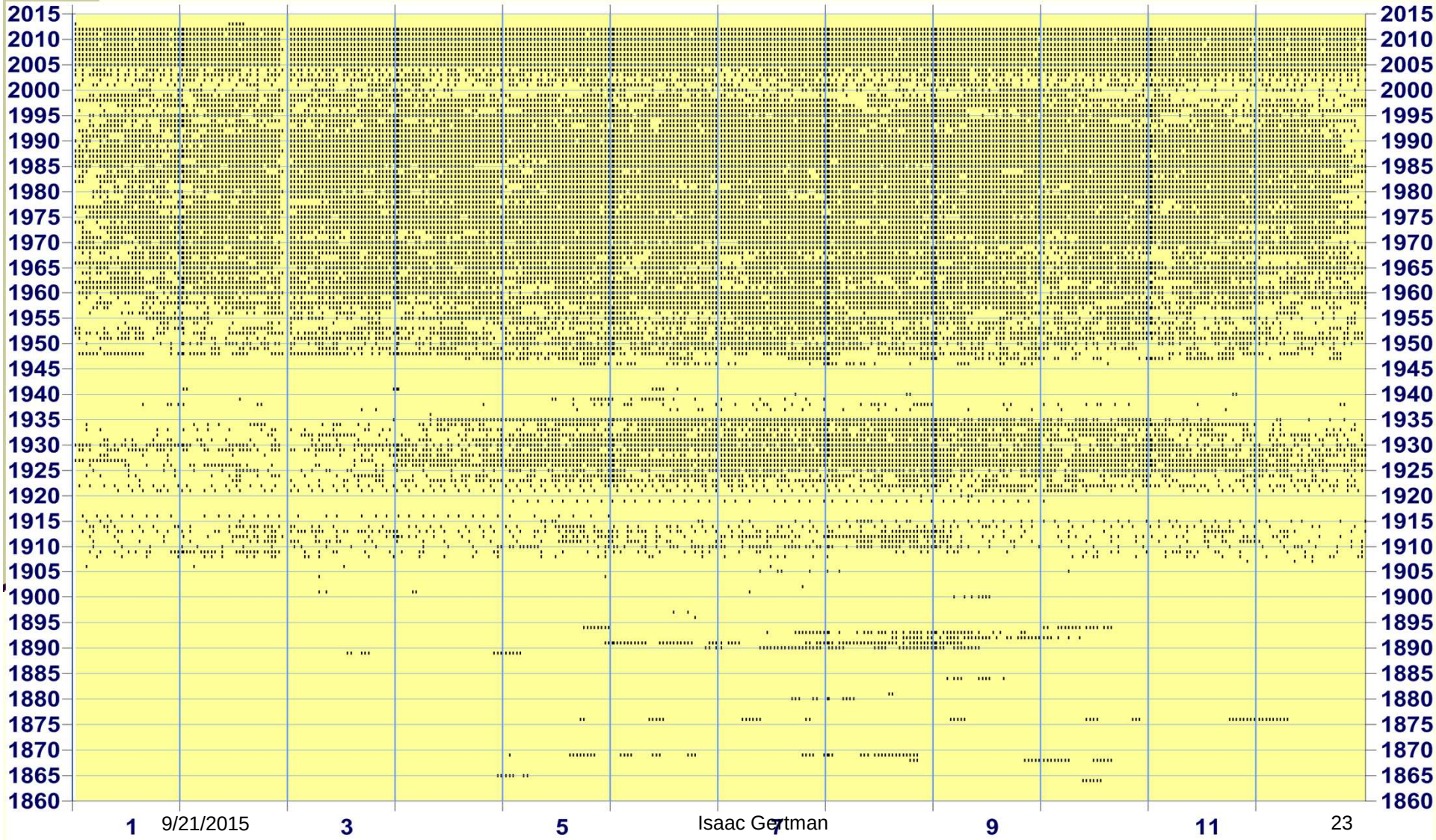




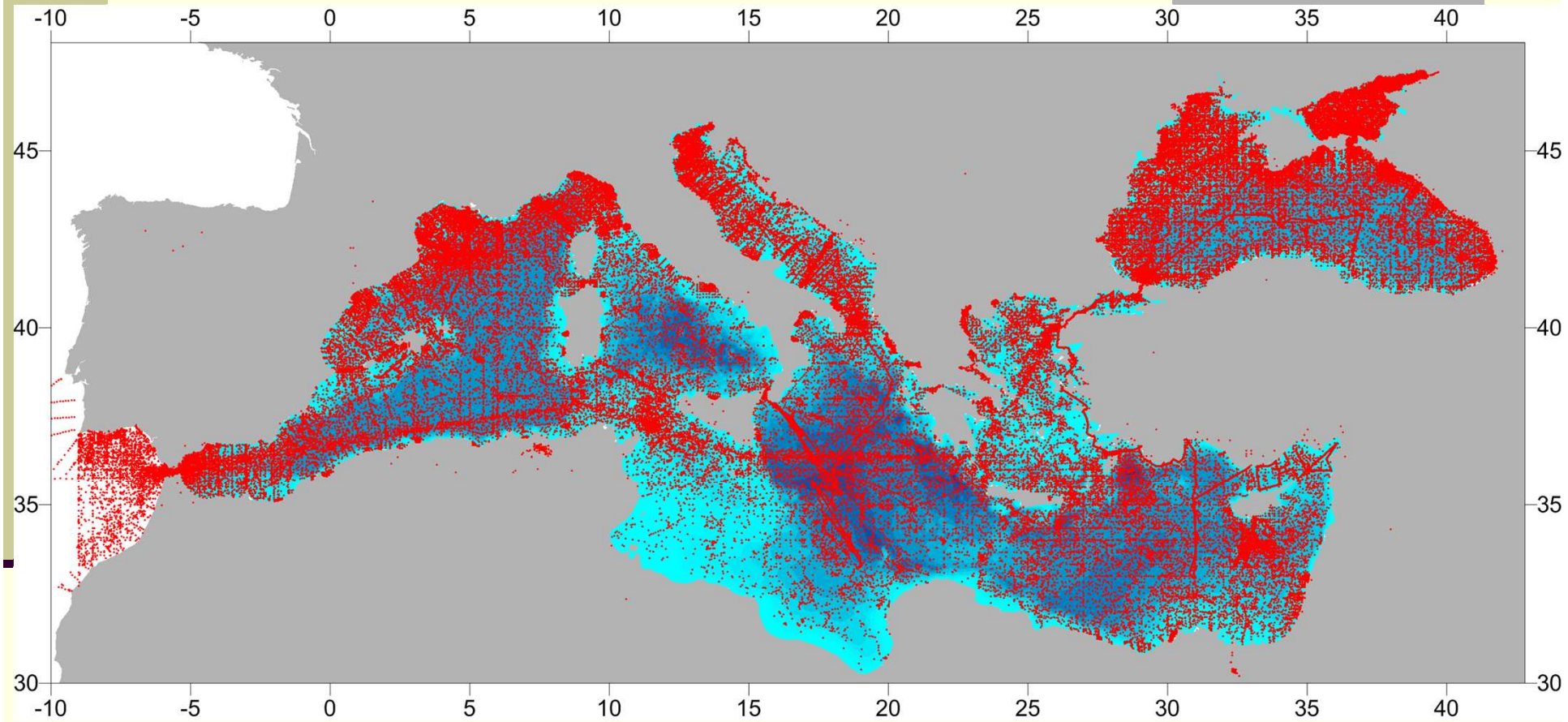
Current volume of data in PERSEUS Cast DB



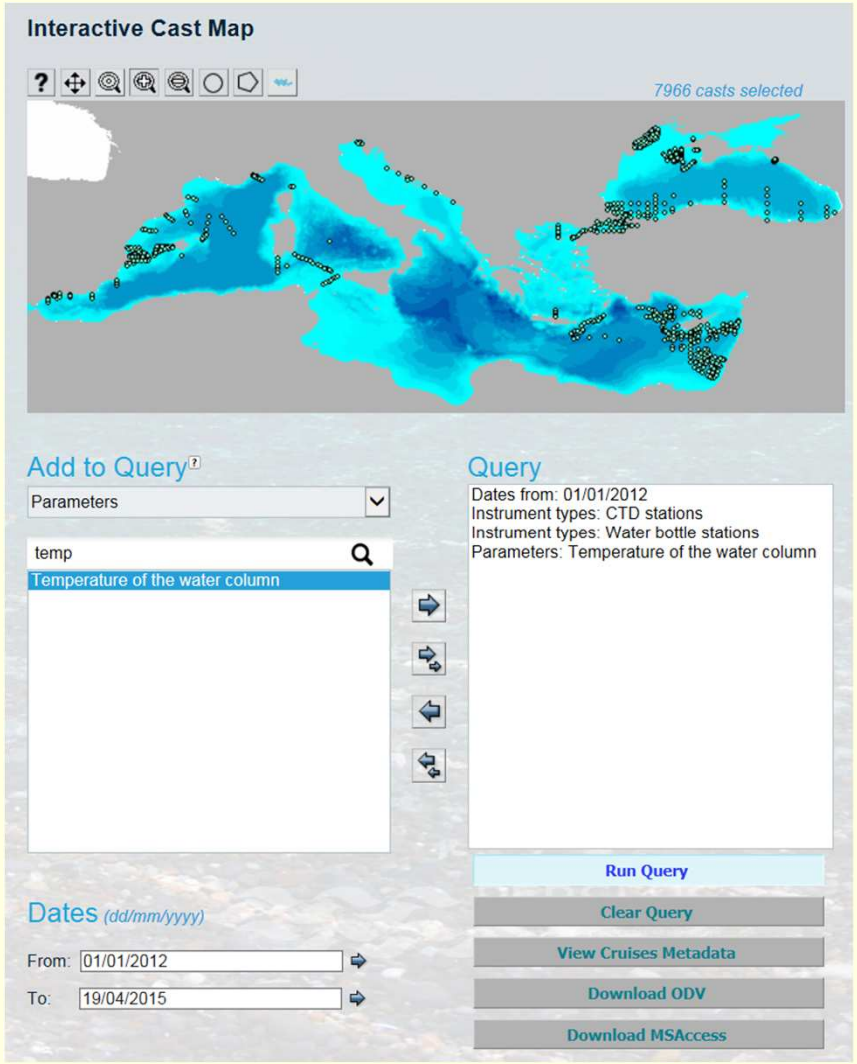
Time distribution of casts



Space distribution of casts

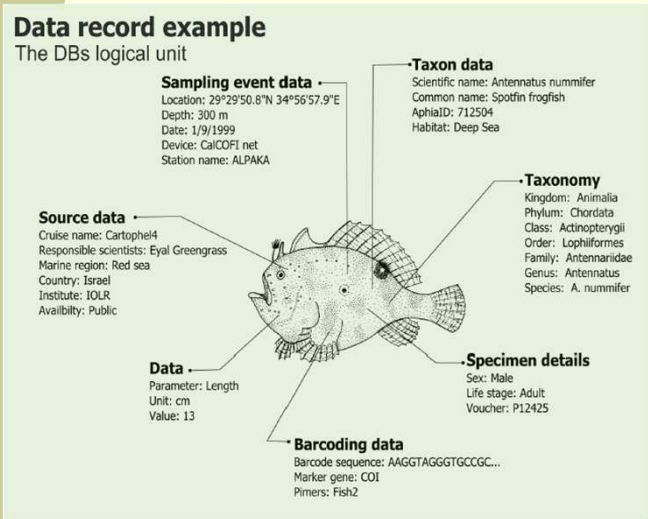


About 100 recent cruises (about 8,000 CTD casts) which were carried out during the last three years (2012-2014).



Israмar-Bio

The taxon-specific data collected in Israel is homogenized according to SeaDataNet standards and imported in the MSSQL DB. An EXCEL based format for online data submission was developed. An on line access to the DB is available and further development continues

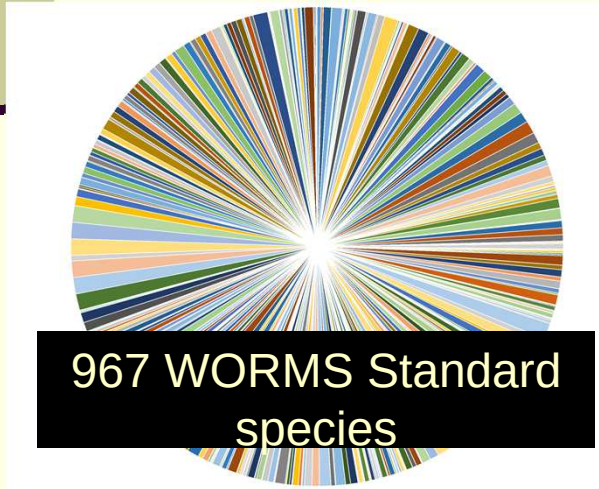


100,000 Observations

A	B	C	D	E
AphiaID	Scientific_Name	Parameter	Units	Value
712504	Antennatus nummifer	Length of unspecified biological entity	Centimetr	13

Ash-free dry weight biomass of unspecified biological entity per unit area of the bed
Count (in assayed sample) of unspecified biological entity
Length of unspecified biological entity
Specimen age of unspecified biological entity
Wet weight biomass of unspecified biological entity per unit area of the bed
Wet weight biomass of unspecified biological entity per unit volume of the water body
Coverage of unspecified
Temperature of the wate

Downloadable submission format
(including drop down lists)



967 WORMS Standard species

Species search WEB page