

**On line generation of PERSEUS standalone  
databases. Deliverable Nr. 9.7**



Project Full title		Policy-oriented marine Environmental Research in the Southern EUropean Seas	
Project Acronym		PERSEUS	
Grant Agreement No.		287600	
Coordinator		Dr. E. Papathanassiou	
Project start date and duration		1 <sup>st</sup> January 2012, 48 months	
Project website		www.perseus-net.eu	
Deliverable Nr.	9.7	Deliverable Date	30/NOV/2014
Work Package No		9	
Work Package Title		Communication, Outreach & Information Management	
Responsible			
Authors & Institutes Acronyms		Dr. Isaac Gertman and Mrs. Yevgeniya Krivenko (IOLR), Dr. Vladimir Belokopytov (MHI)	
Status:		Final (F)	
		Draft (D)	●
		Revised draft (RV)	
Dissemination level:		Public (PU)	●
		Restricted to other program participants (PP)	
		Restricted to a group specified by the consortium (RE)	
		Confidential, only for members of the consortium (CO)	



## CONTENTS

Executive summary / Abstract .....	3
Scope .....	3
1. Introduction .....	4
2. Data Download.....	5
2.1 Download selected data in ODV format (after the “Get Cruises List” button click). .....	6
2.2 Download selected data in form of standalone MS ACCESS Database (after the “Download” button click). .....	7
3. Standalone MS ACCESS Database.....	7
3.1 Harmonization of parameters within the standalone DB.....	7
3.2 Organization of data within the standalone DB. ....	8
4. Visualization of data exported in the standalone DB .....	9
5. References:.....	10

## EXECUTIVE SUMMARY / ABSTRACT

This deliverable describes procedures for selection and online downloading of a subset of historical oceanographic casts data, from the project oriented database developed in the framework of PERSEUS project. The data can be exported in ODV format and in a form of standalone MS ACCESS Database. ODV data can be visualized through freely available ODV software. Volume of the data which can be exported during one session is relative limited (up to 250 cruises). Export volume into standalone DB is significantly larger. Moreover, in parallel to data export, a process of data harmonization (conversion to a standard units) can be carried out. Direct visualization of data in the standalone DB is possible with help of autonomous freely available software.

## SCOPE

The deliverable explains how to easily extract and manipulate a desired portion of historical oceanographic casts data, from the largest (currently) collection of Mediterranean and Black Sea oceanographic cruises data – the PERSEUS Cast Database.



## ON LINE GENERATION OF PERSEUS STANDALONE DATA BASES

### 1. Introduction

A project oriented oceanographic cast data management system ([PERSEUS Cast DB](#)) was developed in framework of task 9.1. The database (DB) located on a dedicated server of a company provided internet services (Netvision). The DB contains Mediterranean and Black Sea multiparametric historical data observed in a form of vertical profiles by Bottle, CTD, Floats, Sea Gliders and Ship Underway upper layer observations. As the primary source for the historical data, the SESAME cast data base (Gertman et al., 2010) was used. In contrast to the old SESAME system, where the primary DB was MS ACCESS DB, the new system has a primary MS SQL DB which includes direct [online interface for data import](#) of new data in [ODV](#) format and QC. The identification of parameters in the PERSEUS system is based completely on the [Common Vocabularies](#) which were introduced as standard terms by [SeaDataNet](#) : [P011](#) (parameter usage vocabulary), [P061](#) (data storage units) and [P021](#) (parameters discovery vocabulary). GIS-like [on line user interface](#) provides a flexible data selection and analysis of metadata (Gertman et al., 2013).

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

- A single cruise or group of cruises in ODV format. Parameters and units within downloaded set of data are the same as in the original submission.
- A group of cruises in form of standalone MS ACCESS DB. All identical parameters (having the same P021 code) with different units within the standalone DB are aggregated to a single parameter and converted to a single unit recommended by SeDataNet as a standard unit for this parameter.

An autonomous relocatable application, developed by MHI partners, allows visualization and other analysis of data downloaded in the standalone DB.



## 2. Data Download

Data download is available for registered users only after log in procedure. Download starts from selection on interactive cast map.

**Interactive Cast Map**

6799 casts selected

**Add to Query**

Instrument types

type text to filter list...

- CTD stations
- Current meters
- Neutrally buoyant floats
- Subsurface temperature and salinity measurements
- Surface measurements underway (T,S)
- Water bottle stations

**Query**

Dates from: 01/01/2012  
Instrument types: CTD stations

**Dates** (dd/mm/yyyy)

From: 01/01/2012  
To: 03/11/2014

Run Query  
Clear Query  
Get Cruises List  
Download

Fig. 1. Primary query definition window



When the primary selection query is defined in “Query” window, and the query is fulfilled (after clicking “Run Query” button, only selected casts are shown on the Interactive cast map), a user should define a way of the selected data download. Two export formats are available:

1. Clicking “Get Cruises List” button allows export of all or part of selected cruises in generic ODV format.
2. Clicking “Download” button allows export all selected cruises in form of standalone MS ACCESS DB.

### 2.1 Download selected data in ODV format (after the “Get Cruises List” button click).

Below there is an example of selected cruises list (Fig.2). Clicking a download button in a cruise row brings to a generation and export of the cruise data in ODV format. To group several cruises in a single ODV collection one has to check several desired check-boxes and click “Aggregate ODW Download” button. The system will aggregate all checked cruises, and will generate ODV file ready for download with all submitted parameters. A typical time for the file generation is less than 10 minutes, but the number of cruises is limited to 250.

Selected Cruises List						
<a href="#">Back to cruises selection on map</a>						
<input type="checkbox"/> Cruise Name	Start Date	End Date	Country	Ship Name	Availi ability	Down load
<input type="checkbox"/> <a href="#">Cruise_BOT^23</a>	14/08/2012	14/08/2012	Georgia	Unknown		
<input type="checkbox"/> <a href="#">Cruise_BOT^24</a>	14/08/2012	14/08/2012	Georgia	Unknown		
<input type="checkbox"/> <a href="#">Cruise_BOT^25</a>	14/08/2012	14/08/2012	Georgia	Unknown		
<input type="checkbox"/> <a href="#">Cruise_BOT^26</a>	14/08/2012	14/08/2012	Georgia	Unknown		
<input type="checkbox"/> <a href="#">Monitoring_2012_08</a>	22/08/2012	27/08/2012	Croatia	Unknown		
<input type="checkbox"/> <a href="#">FGTO^11</a>	16/09/2012	20/09/2012	France	Tethys II		
<input type="checkbox"/> <a href="#">BOUSOLE_2012 (127)</a>	19/09/2012	26/09/2012	France	L'Europe		
<input type="checkbox"/> <a href="#">MOOSE (DYFAMED)^5</a>	21/09/2012	21/09/2012	France	L'Europe		
<input type="checkbox"/> <a href="#">Monitoring_2012_09</a>	26/09/2012	28/09/2012	Croatia	Unknown		
<input type="checkbox"/> <a href="#">Monitoring_2012_10</a>	22/10/2012	29/10/2012	Croatia	Unknown		
<input type="checkbox"/> <a href="#">Monitoring_2012_11</a>	15/11/2012	15/11/2012	Croatia	Unknown		
<input type="checkbox"/> <a href="#">MOOSE (DYFAMED)^8</a>	24/11/2012	24/11/2012	France	Tethys II		
<input type="checkbox"/> <a href="#">BOUSOLE_2012 (130)</a>	07/12/2012	10/12/2012	France	L'Europe		
<input type="checkbox"/> <a href="#">MOOSE (DYFAMED)^6</a>	11/12/2012	11/12/2012	France	Tethys II		
<input type="checkbox"/> <a href="#">Monitoring_2012_12</a>	12/12/2012	20/12/2012	Croatia	Unknown		
<input type="checkbox"/> <a href="#">CYBO_JOLR</a>	23/12/2012	27/12/2012	Cyprus	Shikmona		
<input type="checkbox"/> <a href="#">SHODB_2013</a>	02/01/2013	03/11/2013	Turkey	Unknown		
<input type="checkbox"/> <a href="#">NIB_EWN_2013</a>	10/01/2013	19/06/2013	Slovenia	Sagita		
<input type="checkbox"/> <a href="#">SOCIB-RADMED0213 3 Z</a>	14/02/2013	14/02/2013	Spain	Unknown		
<input type="checkbox"/> <a href="#">SOCIB-RADMED0213</a>	14/02/2013	18/02/2013	Spain	Unknown		
1 2 3 4 5						
0 cruises selected for download of 250 allowed.						
<input type="checkbox"/> Include additional metadata columns in ODV export.						
<a href="#">Aggregate ODV Download</a>						

Fig. 2. Example of Selected Cruises List available for registered user after “Run Query” button clicking.





## 2.2 Download selected data in form of standalone MS ACCESS Database (after the “Download” button click).

Alternatively to ODV data download, **all cruises, selected by customer’s primary query, will be grouped and exported in the standalone MS ACCESS DB. In addition, if the customer’s primary query contains list of parameters, only cruises containing these parameters will be grouped and exported and only these parameters will be included** in the standalone DB.

When generation of the standalone DB is finished on the server, an e-mail message with download address is sent to the customer. The standalone DB generation time can take 1-3 hours depending of the subset volume and the server load. The maximum number of casts, which can be included in the standalone DB, is 100,000. The download procedure can be automatically interrupted by server in two cases:

- The number of selected casts is more than 100,000.
- The size of standalone DB (resulted mdb file) exceed 2 GB.

After the interruption, the customer will get e-mail suggesting to decrease selected number of cast (the selected number of casts appears during selection process above the interactive casts map after clicking “Run Query” button).

## 3. Standalone MS ACCESS Database

### 3.1 Harmonization of parameters within the standalone DB.

Parameters identification in the PERSEUS system is done in terms of [Common Vocabularies](#) which were introduced as standard terms by [SeaDataNet](#) : [P011](#) (parameter usage vocabulary), [P061](#)(data storage units) and [P021](#) (parameters discovery vocabulary). One entry term in P021 vocabulary corresponds to different combinations of P011 terms and P061 terms. Oceanographic analysis of data and specially estimation of interannual variability of water mass parameters, can be done after harmonization of data and conversion of several combination of P011 and P061 terms to a single couple of P021 and P061 terms. Some combinations of P011 and P061 terms cannot be easily converted to desired combination of P021 and P061.

**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**

Parameters	Transmittance and attenuation of the water column
Cruises names contain	hai
Instrument types	CTD stations

☐ Convert parameters data to a single default unit where possible. [Units conversion formulas](#)

[Submit Download Request](#)

Fig. 3. Final submission of standalone DB generation form. To turn on the data harmonization option by units conversation one have to check the “Convert parameters...” checkbox.



Details regarding conversion procedure one can find on the “MS ACCESS Database Download” window (Fig.3). Units conversion formulas and conversion coefficients are presented in “Units conversion formulas table”. Despite the efforts to elaborate optimal conversion procedures one can find that some procedures are not commonly accepted. Therefore, an option to export unharmonized (unconverted) data (in form as it was submitted by data provider) is also available (check box – “Convert parameters”, Fig. 3).

### 3.2 Organization of data within the standalone DB.

A standalone DB can be downloaded using reference in the e-mail message or using “Download Requests Manager” window. The last one can be opened by clicking “MS Access downloads” button on the main page of PERSEUS system window. The download should be unzipped and the MS ACCESS mdb file can be used offline.

Standalone DB contains following tables:

1. **List\_casts:** Metadata about each cast (Cast ID, Cruise ID, station name, coordinates, date/time, bottom depth, number of observation levels, quality flags for metadata, L20).
2. **List\_cruises:** Metadata about the cruise (Cruise ID, Cruise name, Time limits; Ship code, Country code, Data availability, Data Category Code - C077, Observation Laboratory code, Project name, Scientist name, Number of casts, Submission center code).
3. **List\_Params\_In\_Cast:** List of parameters included in the standalone DB for this cast (Cast ID, Parameter Discovery Code - P021, Parameter usage code - P011, Unit Code - P061, Data Category Code - C077, Integral quality code - L20, Conversation Method - free text).
4. **P021\_TEMP:** Harmonized data of in situ temperature of the water column in units recommended for TEMP by SDN standard (Cast ID; PRES - pressure in decibars of the observation level; FPRES - quality flag code for PRES; TEMP - in situ water temperature on the level PRES; FTEMP quality flag code for TEMP; P011 - parameter usage code as it was submitted by data originator. P061 - parameters unit code recommended for harmonization of TEMP parameter; Original\_Val - value of parameters as it was submitted before conversion; Original\_Unit - original unit code as it was submitted by data originator).
5. **P021\_Psal:** Harmonized data of salinity in units recommended by SDN standard (the table structure is identical to the P021\_TEMP table).
6. **P000\_PTS:** Contains united TEMP and PSAL harmonized data. The table can be used instead of P021\_TEMP and P021\_Psal for acceleration of data manipulation.

In addition to above listed tables, a standalone DB contains tables with data for parameters which were selected by primary query during the DB download process. Structure of the harmonized data tables is identical to P021\_TEMP table. If data harmonization was not requested during download process, all data tables contain only original values and units as it was submitted by data originator. A standalone DB

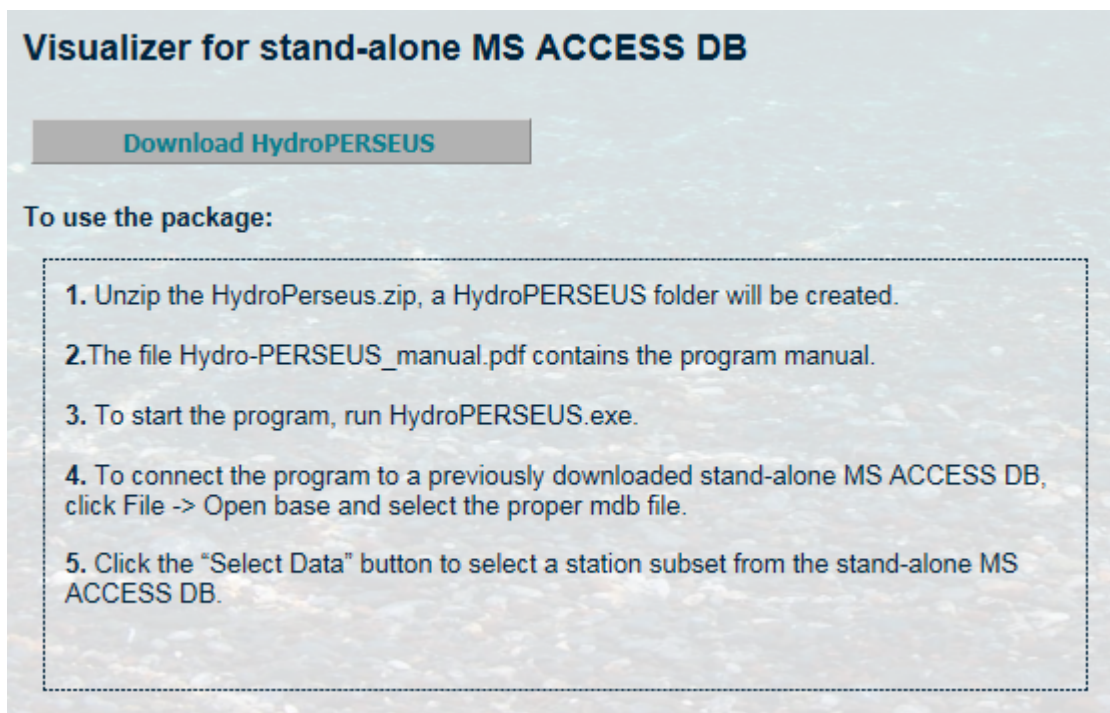




contains also all necessary SDN vocabulary tables with entry terms, which are relevant to the DB.

#### 4. Visualization of data exported in the standalone DB

On the [Oceanographic Data Management Start Page](#) one can find “Visualization Software” button and open window with reference to HydroPERSEUS software developed in MHI by Dr. Belokopytov V.

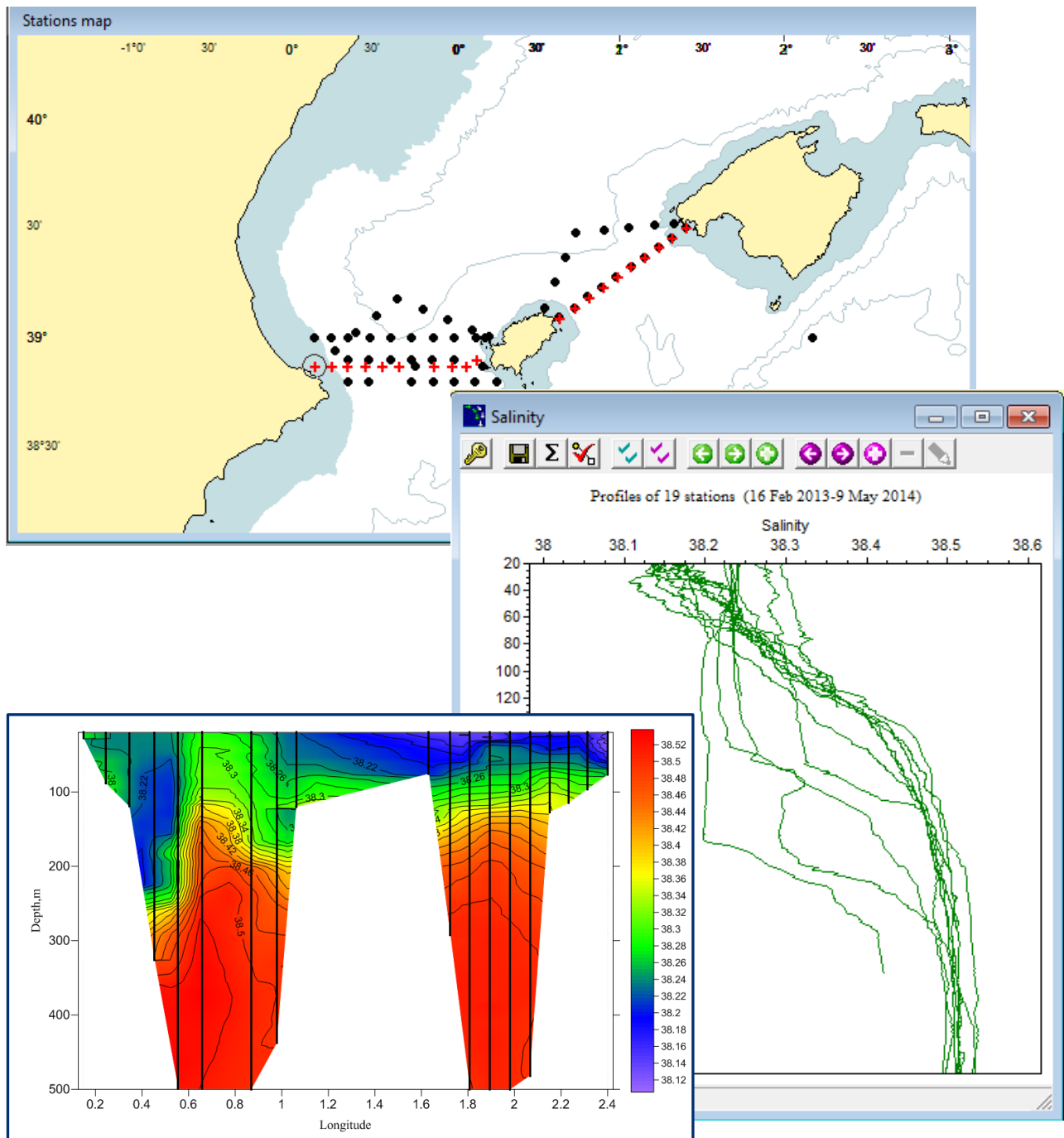


*Fig. 4 Instruction for download and use HydroPERSEUS software for visualization of standalone PERSEUS DB.*

HydroPERSEUS is autonomous and relocatable application. The software allows:

- Select data by cruise name, ship name, time of execution, co-ordinates, parameters.
- Building interactive maps with station position
- Navigation between stations with parallel visualization of data profiles and parameter to parameter diagrams
- Building vertical and horizontal sections

Full list of available data manipulation can be found in the software “User Manuel” downloaded together with the software. Below one can find example of HydroPERSEUS software visualization of data, exported from PERSEUS Cast DB in a standalone MS ACCESS DB.



*Fig. 4 Example of example of HydroPERSEUS software visualization of data, exported from PERSEUS Cast DB in a standalone MS ACCESS DB.*

## 5. References:

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



Gertman I., Krivenko Y., Tal O., Belokopytov V., Ingerov A., Khaliulin A. (2013). Regional project (PERSEUS) oriented system for storage and operative exploration of Mediterranean and Black Sea cast data. Inter. Conf.on Marine Data and Information Systems (IMDIS 2013), 23-25 September 2013, Lucca, Italy, 136-137.

