

Final report

25 p science, 15 p impact

Science:

Tools

(that have been developed during the project): Analytical data base & portal (WP2&4)

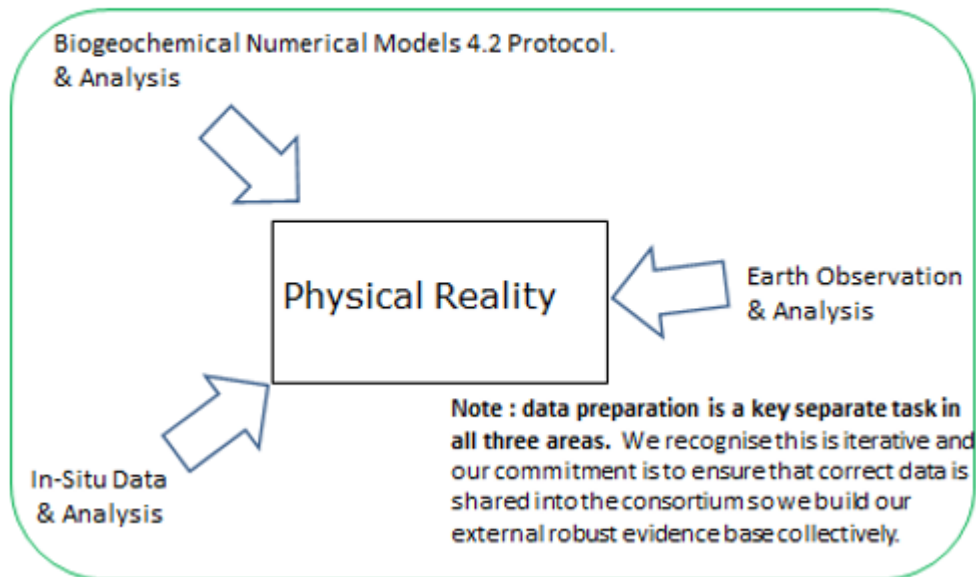
In-situ plankton data are required for a variety of purposes. GreenSeas data has been accessed from various sources and combined into an analytical database. Such a database not only contains the records, but allows data from different sources to be combined and analyzed as though they were from a single source. The successfulness of this approach relied to a large extent on the development of the database structure with great emphasis on data inter-comparability, metadata, quality assurance and taxonomy. After populating the database statistical analytical methods have been used to examine spatio-temporal variation on a range of scales. To create interoperability across in situ plankton data, model output and indicators the GreenSeas Analysis Framework (GSAF) is established, Fig xy, as a protocol to ensure a responsible data sharing policy that ensures scientific robustness of the resulting joint analyses.

In parallel, a data delivery system based on standard open source web-GIS protocols has been developed. The partners and other sources have provided historical and new plankton data, as well as simulation output and indicators in predefined formats, also providing metadata, some making their databases available through OGC compliant servers. The system is designed to comprise a set of distributed data nodes offering the various products. Users access the system through a common web browser, where they can search for plankton data from all providers and display the retrieved datasets on the same map. Selected datasets can also be downloaded for processing in the user's own information systems.

Figure

GreenSeas Analysis Framework

x.



GreenSeas Analysis Framework. To integrate the information together in a combined way the equivalent spatio-temporal data are linked together into one analysis framework by mapping all three observed signals - in-situ, earth observation and numerical models - onto the equivalent in-situ sampling locations.

Analytical Database

In developing the combined analytical database for historical plankton data task has involved integrated work across the work packages, especially with - and building on - Task 4.2,

the data protocol (**D4.2**, WP4). Based on this protocol, the WP2 GreenSeas Data collection template was collectively designed and also the BODC (British Oceanographic Data Center). In addition, a seasonal Atlantic Observatory Discussion Forum, hosted by GreenSeas participants at PML has aided the developing of the exchange of information and ideas across the whole science community (cf. WP7).

Home GreenSeas Data Database Client

Greenport > GreenSeas Data

Navigation

GreenSeas Data

- Numerical Model data
- In Situ Data
- Earth Observation Data
- Cruises/Stations

GreenSeas Data

Overview

Location (Partner)	Total # Records	Time Series Information						
		1900-1930	1930-1950	1950-1980	1980-1990	1990-2000	2000-2010	2010+
Arctic (MMBI)	13,000							
Nordic (Uni-Res)	13,000							
Atlantic (AMT : PML/ NERC)	19,000							
Southern Ocean (UCT, CSIR, Uni-Res)	20,000							
Earth Observation	Extremely large							
Numerical Models	Extremely large							

GreenSeas Thredds Server
GreenSeas Arctic: <http://thredds.nersc.no:8080/thredds/greenseas/catalog.html>

GreenSeas Numerical Model Data
GreenSeas HYCOM-NORWECOM Model data: <http://thredds.nersc.no:8080/thredds/wms/greenpath/Model/topaz?service=WMS&version=1.3.0&request=GetCapabilities>

GreenSeas CMCC Phosphate Model data: http://thredds.nersc.no:8080/thredds/wms/greenpath/Model/cmcc_phosphate?service=WMS&version=1.3.0&request=GetCapabilities

GreenSeas CMCC Chl-a Model data: http://thredds.nersc.no:8080/thredds/greenseas/greenseas-Model-CMCC-Chl-a.html?dataset=CMCC_Chla_Model

GreenSeas CMCC Sea Ice Fraction Model data: http://thredds.nersc.no:8080/thredds/wms/greenpath/Model/cmcc_sea_ice?service=WMS&version=1.3.0&request=GetCapabilities

GreenSeas EO Data
GreenSeas EO data: <http://thredds.nersc.no:8080/thredds/greenseas/greenseas-EO.html>

GreenSeas In Situ Data
GreenSeas In Situ data: <http://tomcat.nersc.no:8080/geoserver/web/?wicket.bookmarkablePage=org.geoserver.web.demo.MapPreviewPage>

The joint approach and harmonizing of data from all the data supplying GreenSeas partners allowed the building of a comprehensive capture (**D2.2, D2.3**) of data from the Arctic (MMBI), Atlantic (PML, NERC, key example shown in Task 2.3) and Southern Ocean (UCT, CSIR, FRUG, Uni-Research). [Figure 2](#) shows the geographic extent of the historical data. This *analytical database* contains 65,000+ in-situ based sampling records, covering a latitudinal gradient ranging from the Arctic (13,000+ records), Nordic seas (13,000+) through the Atlantic (19,000+) and to the Southern Ocean (20,000+), with a wide combination of physical, planktonic based, environmental and chemically based information ([Table 1](#)).

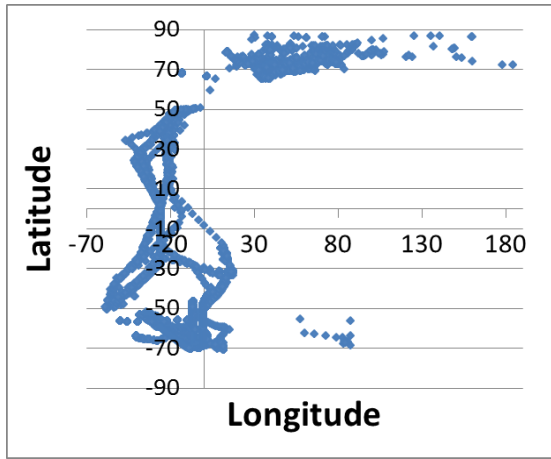


Figure 2. Geographical distribution of measurements within the analytical database.

	Total #	Time Series Information					
		1900-1930-	1930-1950-	1950-1980-	1980-1990-	1990-2000-	2000-
	65,000						
	13,000						
	13,000						
Atlantic	19,000						
Southern	20,000						
Ocean							
Earth							
Observation	Extreeml						
Numerical	Extreeml						

Table 1. Number of records and temporal distribution of measurements in the analytical database.

The definition of the data protocol has also implied that

1. The required plankton taxonomy has been agreed.
2. The key variables needed to address the project science themes have been defined and collected.
3. The outputs of the merged and analysable data are provided in a common format to facilitate exchanges between the different activities in the project.

The datasets can be obtained from the GreenSeas dataportal <http://www.greenseas.eu/content/greenseas-dataportal> as well as from the internal web-site by the GreenSeas consortium, with access also to working / preliminary datasets. They comprise the historical component of the over-all data base building in WP 4 (Task 4.3), especially containing a digitized set of Arctic plankton data providing time series data back to 1913.

D2.3 and 2.4

To create interoperability across in situ plankton data, model output and indicators we have jointly established the GreenSeas Analysis Framework (GSAF), Fig xy, as a protocol to ensure a responsible data sharing policy that ensures scientific robustness of the resulting joint analyses. For instance, statistical analysis of long term time series signals has been carried out with integrated data from the numerical models, earth observation and the in-situ data in the analytical database. This data base provides outputs of manipulated analysable data to the users, providing a consistent plankton dataset by which analyses from various angles (process studies, modelling etc) can provide an updated cross-disciplinary & integrated understanding of the state and function of the Atlantic sector plankton ecosystem in the light of future climate change. The analyses may be extended to global, as the data base is augmented in the future.

...

...

Data delivery system

D4.4 and 4.5

In order to advance our understanding and predictive capacities of how marine ecosystems will respond to global change, GreenSeas has employed a combination of observation data, numerical simulations and a cross-disciplinary synthesis to develop a high quality, harmonized and standardized plankton and plankton ecology long time-series, data inventory and information service. In so doing GreenSeas will contribute to monitoring and assessment of environmental pressures and risks, including tensions and conflicts related to the depletion of natural resources and environmental services (Worm et al 2006) which may arise due to rapid environmental changes and/or natural and man-made hazards.

The essence of the GreenSeas concept is the establishment of a 'core' service following the open and free data access policy implemented in the Global Monitoring for Environment and Security (GMES) programme. The GreenSeas core service delivers both new and historical plankton data and information products.

A specific task has been to define an architecture of the data delivery system based on a distributed client-server model, where a set of data nodes will offer plankton data and/or derived products from a WMS, WFS or WCS compliant data server. The development builds on previous achievements (e.g. DISMAR, InterRisk, NETMAR, MyOcean, ICEDS and ESA-SSE), using established standards for web-GIS and web services, including standardization of data and metadata. DISMAR and InterRisk developed a first version of a web-GIS system integrating remote sensing images, in situ observations and model simulations for marine pollution monitoring and forecasting. NETMAR is currently developing this system further, incorporating new product types as well as new functionality such as processing and chaining services for generating products dynamically and use of ontologies and semantic frameworks for flexible product and service discovery and access. Within the present achievements the capabilities of web-GIS with support for handling of global plankton data have been extended.

For guiding the development, the consortium defined the following generic features and requirements of the data delivery system:

Ease of access

Clear inventory

Visualization – web map

Gridding / Binning

Model- data matchup

If possible: Statistics

Taylor diagram

Main users:

- Modelers
- Earth Observation community
- Plankton Biologist
- Oceanographer
- Policy makers

For the present implementation, modellers have been the main target. The main features include ingestion, processing, data base and viewing.

[System architecture](#)

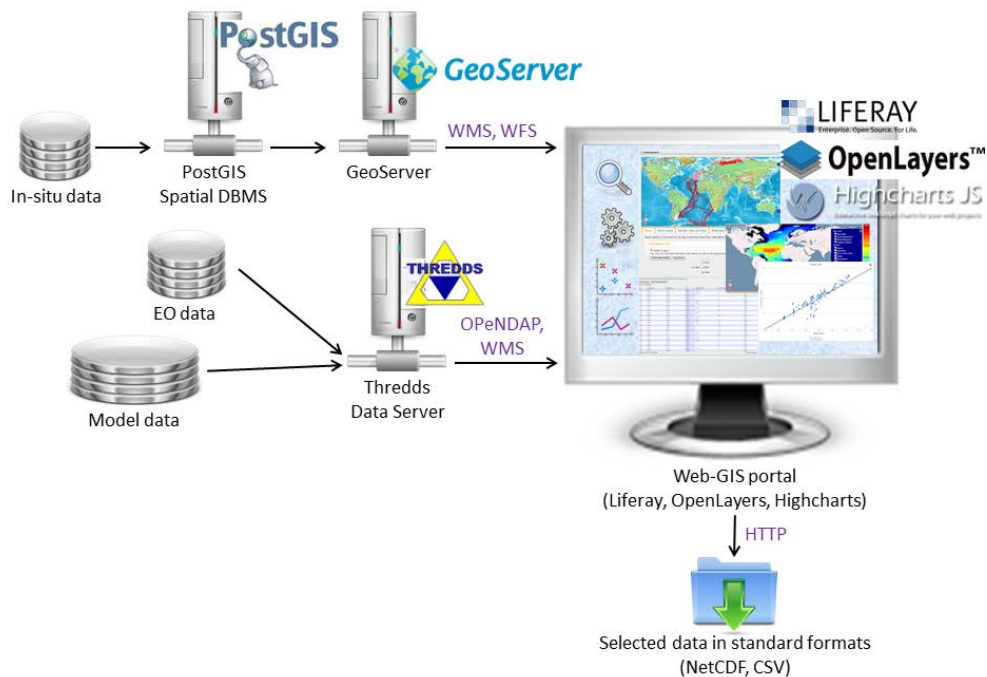


Figure x. Overview of the system architecture.

The GreenSeas DDS is a distributed system where multiple data providers can be connected to the web GIS portal and client through standard protocols. The analytical database of in situ data is connected to the portal through OGC WMS and WFS by means of GeoServer (Figure x). The database of EO and model data is connected through WMS and OPeNDAP, another community standard allowing data streaming, i.e. the client can download selected parts rather than entire files. Adherence to standards for data exchange makes it easy to connect to other external data sources. The DDS will not see any difference between local or external data servers.

The web GIS client, offering query, metadata and data download, presentation and export, as well as simple statistics and different forms of in situ-EO/model intercomparison tools, is also implemented using open source tools and libraries. The main components include Liferay as a web platform on which the client is created, OpenLayers is used for map display including management and colouring of layers, jQuery is used for menus, input fields and the “parameter tree” used to select parameters to include in a search, and Highcharts is used for plotting. Detailed overview of technologies used by the DDS and other specifications are provided in D4.4.

1 System architecture

1.1 Technologies

TO BECOME DIMINISHED BY SEVERAL PAGES:

Geo-spatial databases for in situ, EO and model data

INCLUDE IN DATA BAE CHAOTER Figure 2 shows the information about the GreenSeas databases, comprised of in situ, EO and model data. The different databases are described in sections below.

Figure 1: The web pages contain information about the database.

1.2 In situ data

The GreenSeas in situ database has been compiled by the partners from different areas and time periods, as illustrated in Figure 2. Data compilation was done using a template defined in WP2. The template defined the metadata to be collected for each measurement, including:

- Location: latitude, longitude, depth (of sample and sea), date and time
- Ownership: 1: PI/ Originator, 2: Institute, 3: Research group, 4: Data collection details

For each parameter the following was specified (D2.3: Collated data sets in database and outputs available for analysis):

- Data title and up to 3 subtitles
- Units used and BODC data quality flag¹
- Field description

During the compilation, fields for GS (GreenSeas) originator (including contact information) and cruise/station (fixed) ID were added.

The template was used to collect:

- Physical variables, such as temperature, salinity, optical depth, mixed layer depth, and dissolved Oxygen
- Plankton data, such as diatoms and flagellates (full list in Deliverable D4.2)
- Nutrients, such as N, P, Nitrification N₂, NO₃, NO₂, NH₄, Urea, Si, Calcite

¹ <http://greenport.nersc.no/web/guest/quality-flags>

Table 2 summarises the in situ data compiled in the GreenSeas project. In addition to these cruises, there are data (~11000 points) from the Arctic that has not been identified with a cruise or station id.

Name	Year	Parameters recorded
Iceland Sea GNS	1991-2006	Nutrients, inorganic carbon, salinity and temperature
PS22 - ANT X/6	1992	Nutrients, inorganic carbon, POC, primary production, chl-a via fluorometry, salinity, temperature and various bacteria, zooplankton and pythoplankton
MB95-01	1995	Salinity and temperature
AMT 1-20	1995-2011	Nutrients, oxygen, inorganic carbon, POC, rate measurements, chlorophyll, salinity, temperature and various bacteria, zooplankton and pythoplankton
SR04	1998	Nutrients, inorganic carbon, salinity, temperature
Polarfront	2001-2009	Nutrients, inorganic carbon, chl-a, phaeopigments, salinity and temperature
Gough 2008	2008	Salinity, temperature, DO and fluorescence
SANAE 48	2008-2009	DO, salinity, nutrients, primary production, chl-a via fluorometry and HPLC
SANAE 49	2009-2010	Nutrients, primary production, chl-a via fluorometry and HPLC
SANAE 50	2010-2011	DO, nutrients and chl-a via fluorometry
Winter Cruise	2012	DO, nutrients, chl-a via fluorometry and POC
Expedition Cruise	2013	DO, nutrients and chl-a via fluorometry
SOSCEX	2013	Nutrients, chl-a via fluorometry and POC
SANAE 52	2014	DO and chl-a via fluorometry

Table 1: Cruises and fixed stations in the database.

1.3 EO data

EO data were used to prepare five indicator datasets for intercomparison with the in situ data:

- Chlorophyll-a concentration
- Primary production
- Phytoplankton phenology

- Phytoplankton size fractions
- Ratio of euphotic depth to mixed layer depth

All EO data sets were prepared by PML as NetCDF/CD 1.0 files, and made available by the Thredds Data Server at NERSC through OPeNDAP.

EO source	Creator	Variable	unit	Temporal span / resolution
SeaWiFS	PML	Chlorophyll-a	mg m-3	1997-2010 / monthly
SeaWiFS	PML	Fraction of Microphytoplankton	fraction	1997-2010 / monthly
SeaWiFS	PML	Fraction of Nanophytoplankton	fraction	1997-2010 / monthly
SeaWiFS	PML	Fraction of Picophytoplankton	fraction	1997-2010 / monthly
SeaWiFS	PML	Ratio euphotic depth to mixed layer depth	ratio	1997-2007 / monthly
SeaWiFS	PML	Phenology start	week of year	1998-2007 / annual
SeaWiFS	PML	Phenology max	week of year	1998-2007 / annual
SeaWiFS	PML	Phenology end	week of year	1998-2007 / annual
SeaWiFS	PML	Phenology duration	weeks	1998-2007 / annual

Table 2: EO based indicators available in the database

1.4 Model data

Table 4 summarises the model data found in the GreenSeas database. These model datasets are then briefly described below.

Model name	Variable	unit
HYCOM-NORWECOM	gross nitrogen primary production	mg N m ⁻³ s ⁻¹
HYCOM-NORWECOM	concentration of nitrogen in flagellates	mg m ⁻³
HYCOM-NORWECOM	concentration of silicon in sea water	mg m ⁻³
HYCOM-NORWECOM	sea water salinity	PSU
HYCOM-NORWECOM	integrated chlorophyll one optical depth	mg m ⁻²
HYCOM-NORWECOM	chlorophyll maximum depth	m
HYCOM-NORWECOM	concentration of nitrogen in sea water	mg m ⁻³
HYCOM-NORWECOM	sea water potential temperature	Celsius
HYCOM-NORWECOM	gross primary productivity of carbon euphotic depth	g m ⁻² day ⁻¹
HYCOM-NORWECOM	net primary productivity of carbon euphotic depth	g m ⁻² day ⁻¹
HYCOM-NORWECOM	sea ice thickness	m
HYCOM-NORWECOM	net nitrogen primary production	mg N m ⁻³ s ⁻¹
HYCOM-NORWECOM	chlorophyll maximum chlorophyll concentration	mg m ⁻³
HYCOM-NORWECOM	sea floor depth below sea level	meter
HYCOM-NORWECOM	concentration of nitrogen in diatoms	mg m ⁻³
HYCOM-NORWECOM	concentration of phosphorus in sea water	mg m ⁻³
HYCOM-NORWECOM	concentration of chlorophyll a in sea water	kg m ⁻³
HYCOM-NORWECOM	averaged chlorophyll one optical depth	mg m ⁻³
HYCOM-NORWECOM	sea ice area fraction	fraction
HYCOM-NORWECOM	concentration of oxygen in sea water	mg m ⁻³
PELAGOS	Phosphate	mmol P m ⁻³
PELAGOS	Chlorophyll-a	mg m ⁻³
PELAGOS	Sea ice	fraction

Table 3: Model data available in the database

1.4.1 GreenSeas HYCOM-NORWECOM model

HYCOM-NORWECOM is a coupled physical biological modelling system. HYCOM (v 2.2.12), the HYbrid Coordinate Ocean Model (Bleck, 2002), is an ocean model using hybrid coordinates; that is isopycnal coordinates in the deep stratified waters, and z-level coordinated in the upper mixed layer. A description of this setup of HYCOM can be found in (Sakov et al., 2012). NORWECOM (Skogen and S iland, 1998) is currently run with 11 variables: nitrate, phosphate, silicate, diatoms, flagellates, micro- and meso-zooplankton, nitrogen detritus, phosphorous detritus, biogenic silica and oxygen.

1.4.2 GreenSeas PELAGOS model

The PELagic biogeochemistry for Global Ocean Simulations (PELAGOS²) model is a coupling between the Biogeochemical Flux Model (BFM) and NEMO, the general ocean circulation model of the Euro-Mediterranean Center on Climate Change (CMCC). The PELAGOS model is used in GreenSeas to predict phosphate, as well as many other plankton

² <http://www.cmcc.it/models/pelagos-pelagic-biogeochemistry-for-global-ocean-simulations>

parameters at 2 degree resolution for the whole ERA-interim period (1980-2010) with standard parameters set from Vichi and Masina (2009). The output from the model for phosphate, chlorophyll-a and sea ice parameters has been interpolated to a regular 1 degree grid for the GreenSeas database.

Search and retrieval features

The data inventory is available through the graphical interface from which a search is initiated (Figure 3). The user can customize the query to his/her needs, through a suite of search criteria.

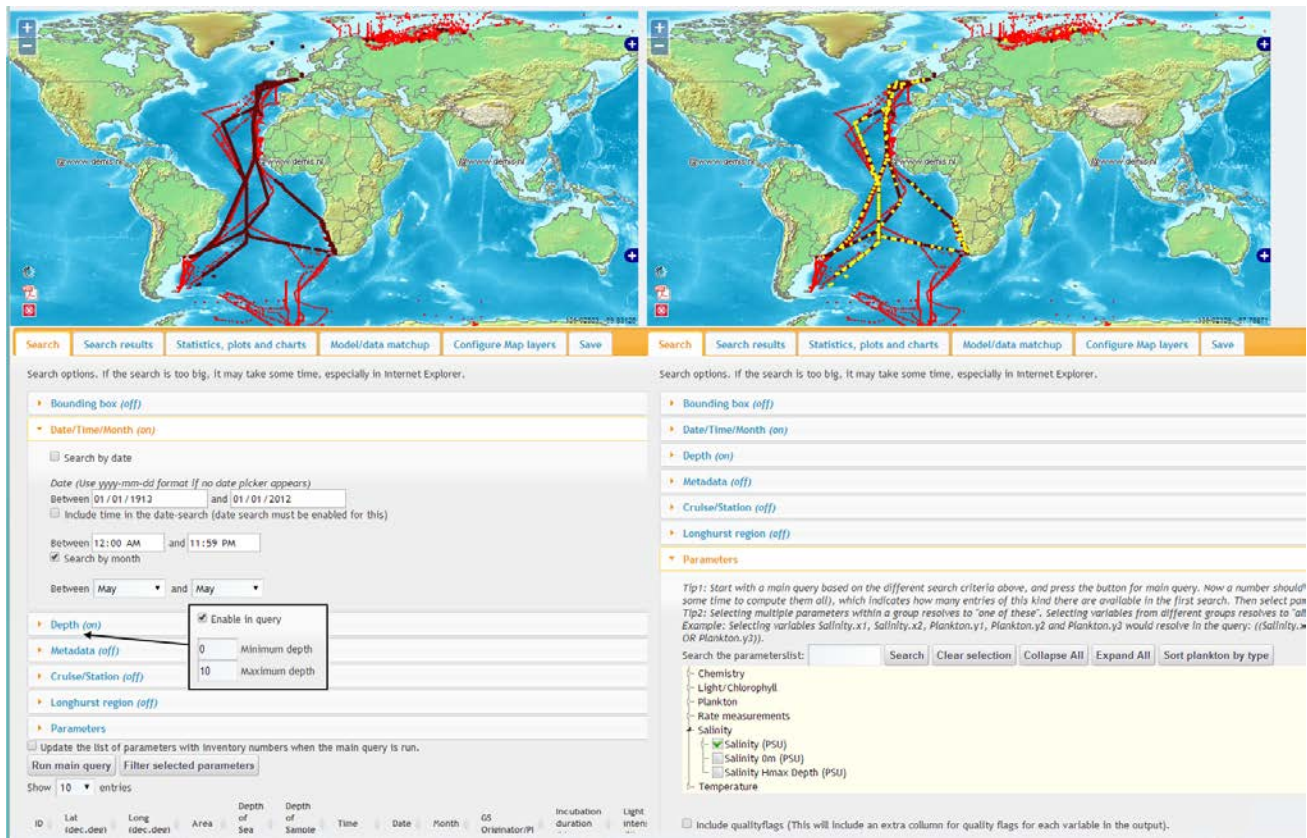


Figure 0: Example of a search via the graphical interface. All data shown as red dots. Left: A main query has been run with the options: month: May and depth: 0-10 meters. Results are displayed on the map as brown dots and also listed at the bottom of the screen. Right: The parameter salinity has been selected, and the map is updated with yellow dots where salinity has been found. The search results can be displayed or exported as available from the menu.

Search by area

1.4.3 Bounding box

The most basic search option is to use a bounding box to limit the search results. A user can either type in the minimum/maximum of latitude/longitude manually or directly on the

map, and optionally zoom to the selected area when the main query is run. Further options for limiting the search include:

- Longhurst provinces or biomes - a set of predefined areas explained in [section 3.4](#), can be selected.
- Search by time – date or month range – user selects a date range (and optionally a time-of-day interval) and/or a month range.
- Search by depth - The user can specify a depth range to search by.
- Search by cruise or fixed station - The cruises and fixed stations is presented to the user as a dropdown menu from which he/she can select a specific cruise or fixed station to search by.
- Select metadata - By default, all metadata will be output by a query. There are 15 metadata variables, of which all may not needed by the user.
- Filter for specific parameters
-

Filtering for specific parameters

The main feature of the DDS is to be able to retrieve plankton data. A tree structure (Figure 8) is accommodated from which the user can easily select the desired parameters from a list of currently 240. The parameters are pre-categorized and grouped to ease overview. At the highest level, there are 6 nodes or categories: Chemistry, Light/Chlorophyll, Plankton, Rate measurements, Salinity and Temperature. The groups are easily adaptable to change if new parameters are added or a new structure needed.

To further enhance the selection of parameters, there are four important functions for the search tree:

- Search for any text string in the tree. The results will be highlighted in red and appropriate nodes opened to display all the matches.
- Toggle between sorting the tree by plankton type and plankton size. (Default: size) (Figure 8).
- Download information. (Short name, long name, unit, description).
- Display number of occurrences of each parameter in the search result.

Parameters

Search the parameterslist: Search Clear selection Collapse All Expand All Sort plankton by type Download info on selected

Loading of inventory numbers is complete

- Chemistry [365]
- Light/Chlorophyll [411]
- Plankton [44]
 - Picoplankton [0]
 - Nanoplankton [44]
 - Open Ocean Diatoms (20-200µm) [0]
 - Microplankton [44]
 - Ciliata [44]
 - Abundance - P500M17Z (# ml-1) [44]
 - Carbon biomass - S7338A30 (mg m-3) [0]
 - Dinophyceae [Size: >20µm Subgroup: heterotrophic] [0]
 - Dinophyceae [Size: >20µm] [0]
 - Microplankton 20-30 µm [0]
 - Microplankton 30-50 µm [0]
 - Microplankton 50-100 µm [0]
 - Mesoplankton [0]
 - Others [44]
- Rate measurements [0]
- Salinity [192]
 - Salinity (PSU) [192]
 - Salinity 0m (PSU) [0]
 - Salinity Hmax Depth (PSU) [0]
- Temperature [216]

Figure 2: List of parameters, updated with inventory numbers for the AMT 1 cruise.

2 Statistics, Plotting and intercomparison features

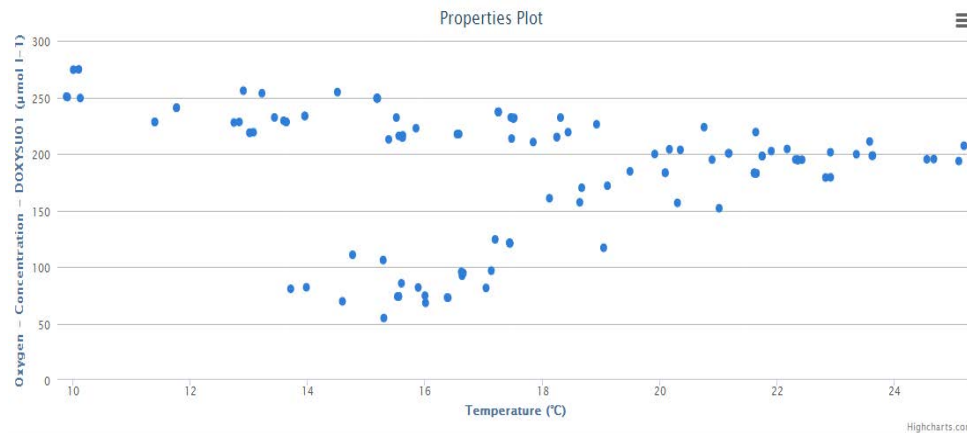


Figure 3: Create property-property plots shows a property-property plot comparing temperature and oxygen concentration measurements from the AMT 15 cruise, in September-October 2004, at 90-110 meters depth.

-
- Statistics - can optionally be calculated for each parameter in the search: number of samples, min, max, average, median, lower quartile, upper quartile, sample standard deviation, variance.
- Time series - A time series graph of a selection of parameters can be generated. The user can zoom on the graph to take a closer look at a specific time period. Information on each point is displayed when hovering over it with the mouse. The graphic can be exported (png, jpeg, pdf or svg format).
- Property-property plot - allows plotting of any two parameters from the search against each other, matched on station and depth. Optionally, one can bin depth in case of multiple hits. Figure 13
- Match-ups - Optionall the average value (over depth) can be used in stead of all data points.
- The user has the option to match on month and/or year, and the option to choose exact match, or nearest match.
- Scatter plot - User selects a pair of parameters from any source to create a scatter plot as an easy way of comparing the data.
- a data source and a variable, a form for choosing time and elevation interpolation is displayed. The time resolution of the model and EO data in the database is monthly or yearly, so an interpolation is sensible in order to match the data. By hovering over the map with the mouse, more information on nearest point is displayed: time, latitude, longitude, depth and value from both sources, and ID (for in situ data).
-
-

The depth resolution varies, but it is never high enough to be able to match the data at exact depths. The user could either pick a specific depth in the raster and match all in situ points to that depth or choose to compare each point to the closest level.

Figure 15 shows an example of how this would be used in the application. This example is following the sequence from Figure 10, where salinity for the month of May on the depths of 0-10 meters was selected. This example would compare the salinity from the situ data results with the salinity for the month of May of the nearest year in the HYCOM-NORWECOM model. The model has the following depth layers: 0, 5, 10, 25, 50, 75, 100, 125, 150 and 200 meters. The in situ data would be compared to the nearest match, which would be 0, 5 or 10 meters depth.

2.1 Match in situ data with raster data

2.1.1 Data source

User may select from two lists for data sources in addition to in situ data: by OPeNDAP or as netCDF files. In the latter case, the user may save the file instead of re-uploading it on next visits at portal.

2.1.2 Add to data output

You can choose to add the data to output, so that you can calculate the statistics of the matched data (Figure 11) or export the data in different formats (chapter 6).

2.2 Add layers on the map

In situ data can be compared with the raster data by displaying them as separate layers on the map, e.g. model results overlain by insitu data to evaluate model performance.

3 Data export features

An important feature of the DDS is the function that allows the user to download the data from the search, currently in CSV or netCDF format.

3.1 CSV

A CSV file is a plain text file, where tabular data is separated by a separator. The separator should be a character that is not used in the data, hence semicolon is chosen for our CSV file. The first line in the file is always “*sep=;*” which tells anyone who uses this file what the separator is.

The CSV file can be opened in any text-editor and also a several other programs, like Microsoft Excel. In Figure 18 the export from Figure 17 is opened and displayed in Microsoft Excel

3.2 NetCDF

The second type of file is netCDF (Network Common Data Form)³ using the CF-conventions 1.7 feature type: point⁴.

The netCDF-file can be read through a number of netCDF applications. In Figure 19 the exported file is opened in the ToolsUI tool from Unidata⁵.

3.3 Export map

The map can at any time be exported as PDF, with the current visible map layers, by clicking the PDF-icon in the bottom left corner of the map. A human readable description of the query that is performed and the name of the layers is included in the PDF. This also helps the user to recommence work if a session is interrupted for various reasons.

3.4 Export plots

The different plots created by the application, time series, property-property plots and scatter plot, can all be exported in several formats: PNG, JPEG, PDF or SVG.

IMPROVEMENTS Match parameters by station ID:

- Add new parameters to the GreenSeas database.

³ <http://www.unidata.ucar.edu/software/netcdf/>

⁴ <http://cfconventions.org/Data/cf-conventions/cf-conventions-1.7/build/cf-conventions.html>

⁵ <http://www.unidata.ucar.edu/downloads/netcdf/netcdf-java-4/index.jsp>

- Integrate more external data sources such as GBIF, MAREDAT, OBIS, and EMODNET to the system, to increase the amount of data available to the user.
 - Try to integrate also local marine data collection projects and initiatives, by providing them with instructions on how they can upload and share their data.
 - Develop services that enhance the functionality of the system for scientists wanting to use it for storage of their data, as well as for scientists that want to use it to find relevant data for their task at hand.
 - A function to export the map with details about the search layers that was applied. the user can now export the map and information about the query as a PDF.
 - Changes to the statistics presentation – round the numbers to be of same precision as the input.
 - Include a colour legend for the raster layers.
 - Link the application to also use the in situ data in PANGAEA.
 - Quality control is the responsibility of the data provider, e.g. check for outliers and missing data.
 - A unique cruise ID and station ID is important to be able to identify the data. An Expocode⁶ is recommended identifier for a cruise. A DOI⁷ is also important to identify a dataset.
 - A search by data collection method could be useful for advanced users.
 - The lowest level of grouping in the parameters tree could be based on unit.
 - Version control of the data is necessary to be able to extract identical output from the same query in order to be able to reproduce any experiment that follows the data extraction.
-
- Simplify the parameter tree.
 - Include more data, specifically in the Norwegian Sea.
 - Link the application to other data sources, not only the GreenSeas database.

All of these requests have been mentioned earlier, and only the first request has been adopted.

Summing up, the potential impact of the GreenSeas DDS includes

- Offering a new tool for validation of ecosystem models, allowing to upload of modeller's own results and comparison with in situ observations or EO derived parameters.
- Serving as a basis for maintaining an extensive database of in situ measurements, EO products and model simulations, where existing data sources can be integrated

⁶ <http://wiki.pangaea.de/wiki/EXPCODE>

⁷ <http://www.doi.org/>

dynamically through open and widely accepted standards for data exchange such as OPeNDAP and WFS.

- Providing a starting point for marine scientists who want to develop geo-spatial databases and/or services for another geographic region, by demonstrating a working web-GIS system comprised of a suite of open source technologies and tools that can be further developed to the benefit of the wider marine scientific community.

4 References

- Bleck, R. (2002), An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates, *Ocean Modelling*, 37, 55-88.
- Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., & Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecological Modelling*, 221(11), 1472–1483.
doi:10.1016/j.ecolmodel.2010.02.014
- Buitenhuis, E. T., M. Vogt, R. Moriarty, N. Bednarsek, S. C. Doney, K. Leblanc, C. Le Quere, Y.-W. Luo, C. O'Brien, T. O'Brien, J. Peloquin, R. Schiebel and C. Swan, 2013. MAREDAT: towards a world atlas of MARine Ecosystem DATA. *Earth Syst. Sci. Data*, 5, 227–239, 2013.
doi:10.5194/essd-5-227-2013
- De Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research*, 109(C12), C12003.
doi:10.1029/2004JC002378
- Hansen, C., Kvaleberg, E. and Samuelsen, 2010: Anticyclonic eddies in the Norwegian Sea; their generation, evolution and impact on primary production, *Deep Sea Res. Part I Oceanogr. Res. Pap.*, 57(9), 1079–1091, doi:10.1016/j.dsr.2010.05.013
- Koeller et al (2009). Basin-Scale Coherence in Phenology of Shrimps and Phytoplankton in the North Atlantic Ocean *Science* 324, 791-793.
- Longhurst, A. R. (2007), *Ecological Geography of the Sea*, 2nd ed., Academic, Burlington, Vt.
- Mélin, F. (2003). Potentiel de la télédétection pour l'analyse des propriétés optiques du système océan-atmosphère et application à l'estimation de la photosynthèse phytoplanktonique. PhD thesis, Université Toulouse III.
- Mélin, F., and Hoepffner, N.: Monitoring phytoplankton productivity from satellite. In "Handbook of Satellite Remote Sensing Image Interpretation: Applications for Marine Living Resources Conservation and Management", EU PRESPO and IOCCG, Dartmouth, Canada, 293 pp., Case study #6, pp79-93, 2011.
- O'Reilly, J.E., and 24 Coauthors, 2000: SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3. NASA Tech. Memo. 2000-206892, Vol. 11, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, 49 pp.
- Platt, T., and Sathyendranath, S. (1997). Modelling primary production III. *Aquabiology*. 19: 142-146.

Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R., and Korablev, A.: TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic, *Ocean Sci.*, 8, 633-656, doi:10.5194/os-8-633-2012, 2012.

Skogen, M. and Sjøiland, H.: A user's guide to NORWECOM v2. 0, a coupled 3 dimensional physical chemical biological ocean-model, *Nor. Ecol. Model Syst.* Bergen, 1998.

Vichi, M., and S. Masina (2009), Skill assessment of the pelagos global ocean biogeochemistry model over the period 1980-2000, *Biogeosciences*, 6(11), 2333–2353.

END OF DOCUMENT

New data & processes understanding

(by geographical region) (WP3) + analysis based on D2.4

Arctic D2.1 and D2.5

Nordic D3.5

Atlantic D3.3

SO D 3.4 and 3.6 and D3.1

Analysis D2.4

Biomes (WP2 & ???)

4.1 GreenSeas biomes

Biomes are regions which confirmed by measurements can be considered macro-scale ocean ecosystems that share consistent physical and biological conditions. Many different biomes have been defined by different scientific communities. According to

Deliverable D2.6 (Collated data and analyses provided to WP4), a widely accepted classification for the world's oceans is Longhurst's (2007) division into four major biomes: *polar, westerlies, trade winds, tropical*; each with variations described in total of approximately 50 ocean provinces.

In WP2, GreenSeas has analysed satellite ocean colour data from the ESA CCI Ocean Colour project, and using Harmonic Analysis of 12 and 6 month periods defined a new set of marine biomes:

1. Non Gyre 1: Asymmetric single peak early spring bloom – non light limited region
2. Non Gyre 2: Symmetric single spring peak, some light limitation
3. Non Gyre 3: Symmetric summer peaked bloom with secondary autumn peak – winter light limited region
4. Non Gyre 4: Asymmetric and prolonged single autumn peak – light limited region
5. Non Gyre 5 Equatorial: Symmetric double bloom due to Kelvin wave dynamic – non light limited region
6. Non Gyre 6 Polar: : Symmetric single summer peak - light limited in winter
7. Gyre 1: Symmetric double peak, major summer, minor winter bloom– non light limited region
8. Gyre 2: Symmetric single winter bloom and possible summer bloom – non light limited region
9. Gyre 3: Symmetric single peak early spring bloom – non light limited region
10. Gyre 4: Asymmetric prolonged single spring peak – non light limited region

The method used to define the GreenSeas biomes is described in detail in Deliverable D2.4 (Report of spatio-temporal patterns).

D2.4

[Application of data base to test models \(WP5\)](#)

D2.6 and D5.x

Impact: Indicator (WP6 stuff) D6.x; Next gen models (Richard/Ic) D5.6; CMIP runs (Marcello) D5.x; roadmap (Icarus); dissemination (web based dissemination tools), conference presentations, papers; IPCC; GEO; MSFD (biodiversity and food-webs) Icarus; WG on ecosystem health (CEFAS); Int ...Climate...Ecosystem.. SCAR (Richard was present)

Copernicus marine service (Johnny)
(Potential) impact: proof of concept (Richard)

Summer school, facebook, community is still communicating: a resource
Educational resources linking with ESA, application of remote sensing (indicators) for
graduate students.

Video!

Funding of students support, data base (quantifications if possible, no. of students)

Berths for students on AMTs (and other cruises?) in general.

Formed a robust community (sort of), memorandum of understanding.

References

D4.5 Data delivery system architecture. *GreenSeas report.*