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

Alexander Barth[1], Sylvain Watelet[1], Charles Troupin[2], Aida Alvera Azcarate[1], Giorgio Santinelli[3], Gerrit Hendriksen[3], Alessandra Giorgetti[4] and Jean-Marie Beckers[1]

[1]GHER, University of Liège, Belgium

[2]SOCIB, Spain

[3]Deltares, Netherlands

[4]OGS, Italy

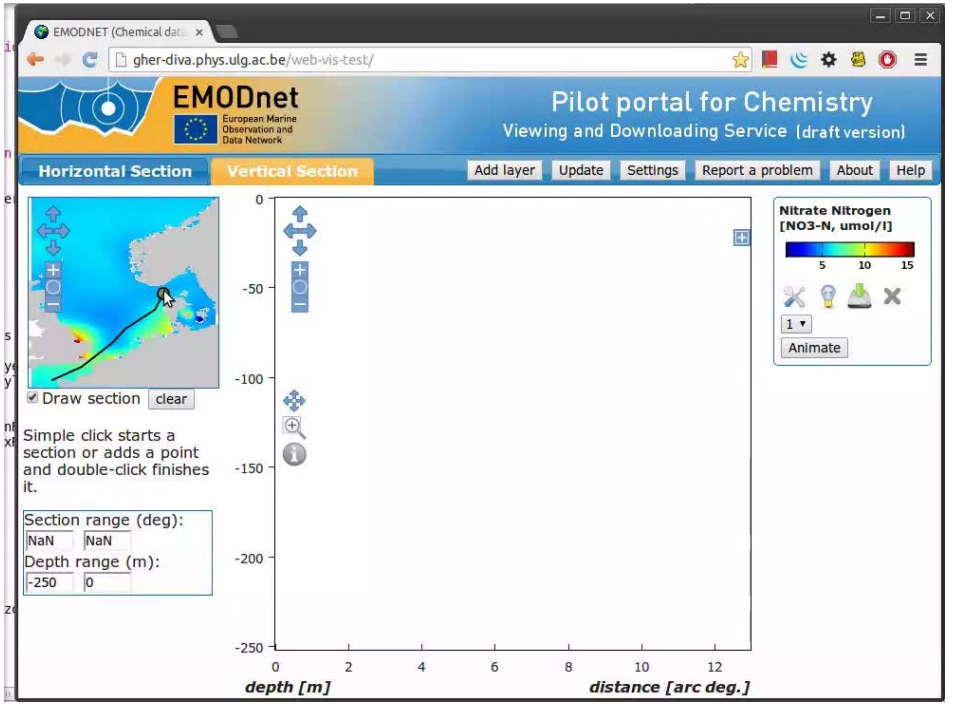
What is OceanBrowser?

- Web-interface to **visualize gridded** data sets in NetCDF
- Implements the Web Map Service protocol
- Horizontal and vertical sections
- Scalar and vector fields
- OceanBrowser is used in
  - **SeaDataNet**
  - **EMODNET Chemistry**
- In those projects it is used to visualize gridded data sets generated by the tool DIVA (Data-Interpolating Variational Analysis)

http://ec.oceanbrowser.net/emodnet/
http://sdn.oceanbrowser.net/web-vis/

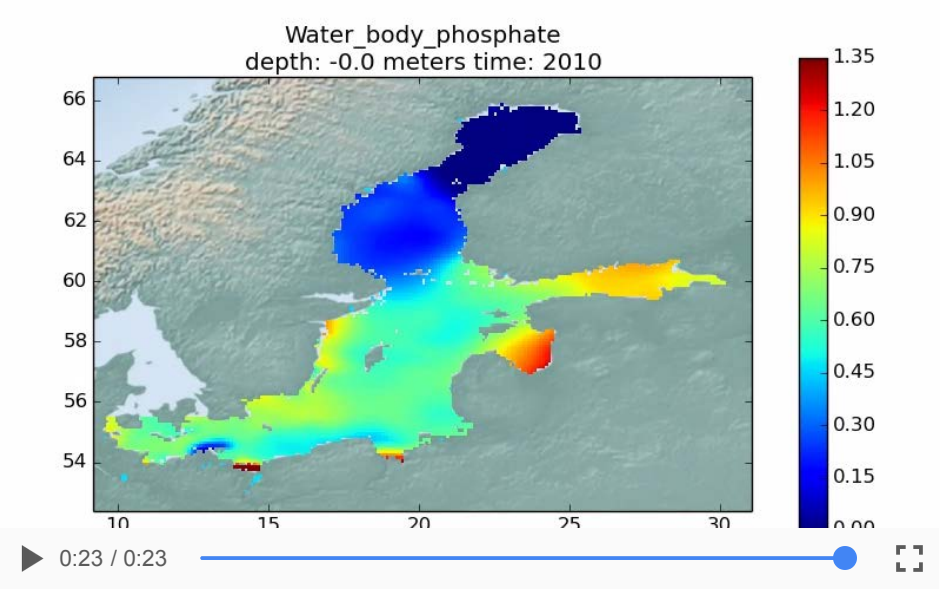
Vertical section

- Vertical section can be drawn with the mouse
- Data product will be extracted along this section (x-axis: distance from starting point and y-axis depth)
- Section coordinates can be saved (to visualize two parameters along exactly the same section)
- The path of a vertical section can be generated automatically by:
  - fixed distance from coast
  - or fixed ocean depth



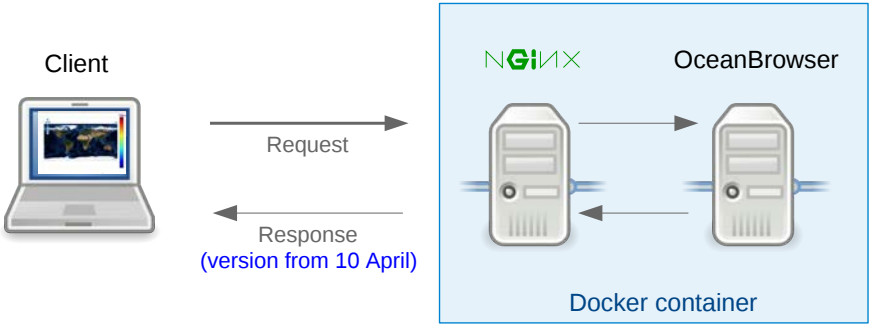
Export animations

- Winter distribution of phosphate (produced by SMHI)
- Centred 10-year average of all winter months
- OceanBrowser: export of animation by choosing MP4 or WebM animation.



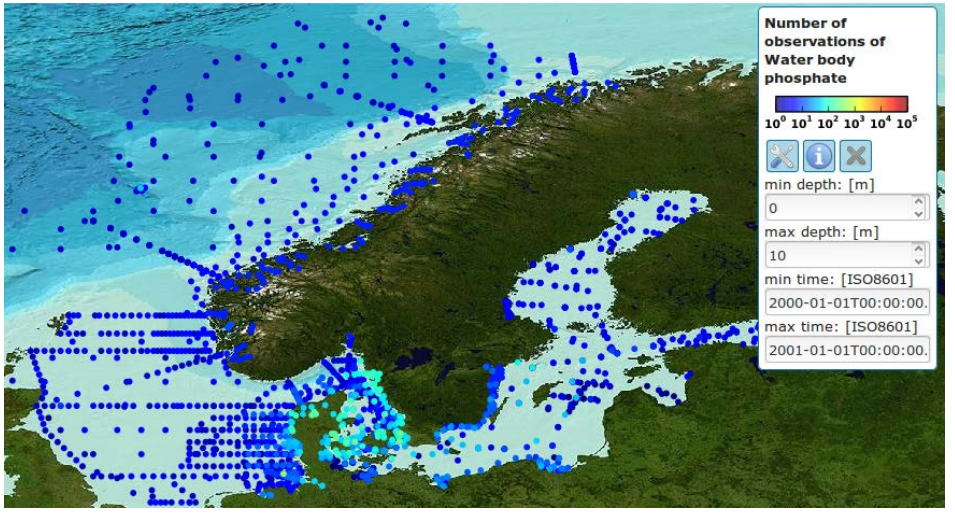
Speed optimization

- OceanBrowser implements **cache control headers**
- Significant **improvement of the responsiveness**
- Minimizing the risk to using an out-of-date content.
- Web browser must **check with the server if a newer version exists**
- The server can:
  - confirm that the cached version is the current version (cache revalidation)
  - respond with a latest version of the corresponding request
- Potential cache revalidation is fast because (only a comparison of the time-stamps)



Observation location

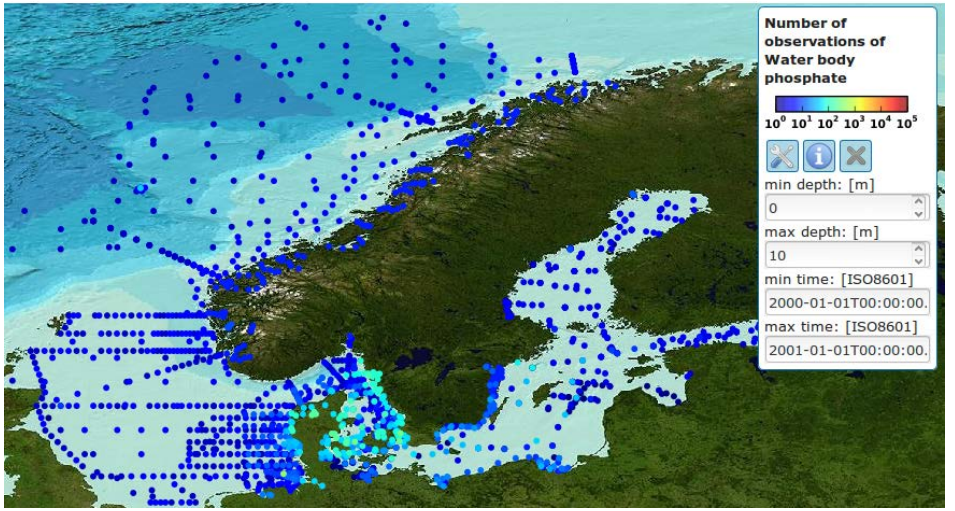
- Web Feature/ Processing Service by Deltares
- Web Feature Service → List of all available parameters



- Requirements
  - Data location (within depth and time range)
  - Color shows the number of observations
- In the past: Web Feature Service → Location of every observation

Observation location

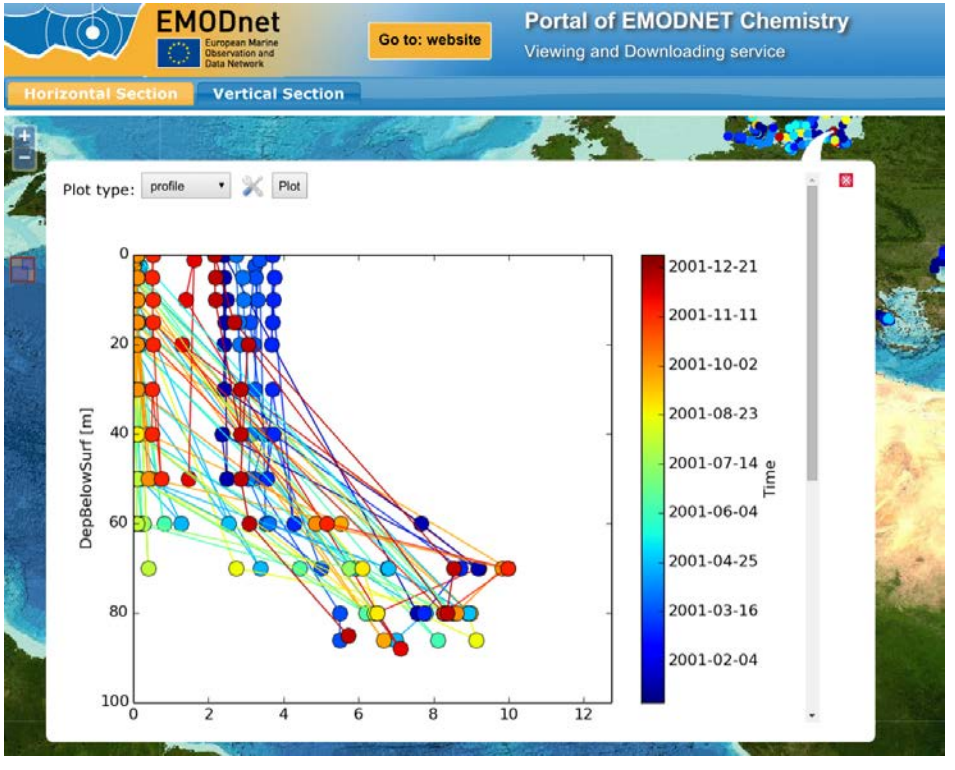
- Web Feature/ Processing Service by Deltares
- Web Feature Service → List of all available parameters



- Requirements
  - Data location (within depth and time range)
  - Color shows the number of observations
- In the past: Web Feature Service → Location of every observation
- However:
  - About 10000-100000 data points: **to much data for a web browser**
  - Web Feature Service: only filtering, **no aggregation**
- Web Processing Service → Image with the observation location

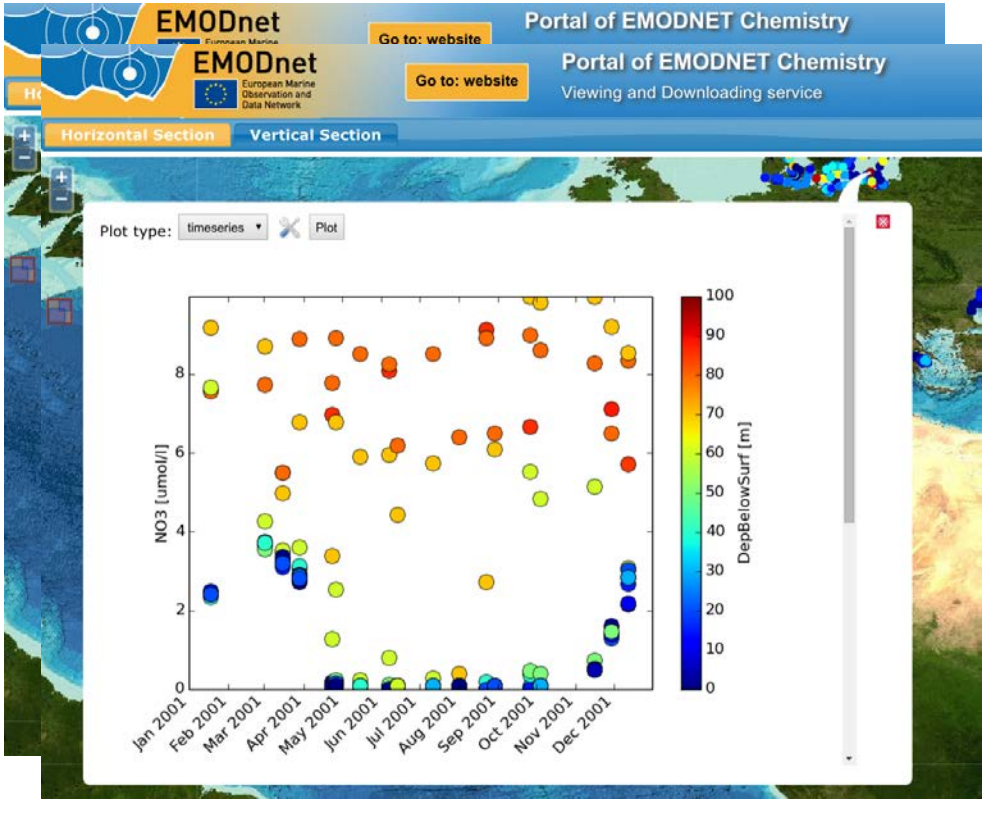
Profile and time series plots

- Plots can be changed dynamically
- Varying parameters: depth, time and measured value
- Profile (x: value, y: depth, color: time)



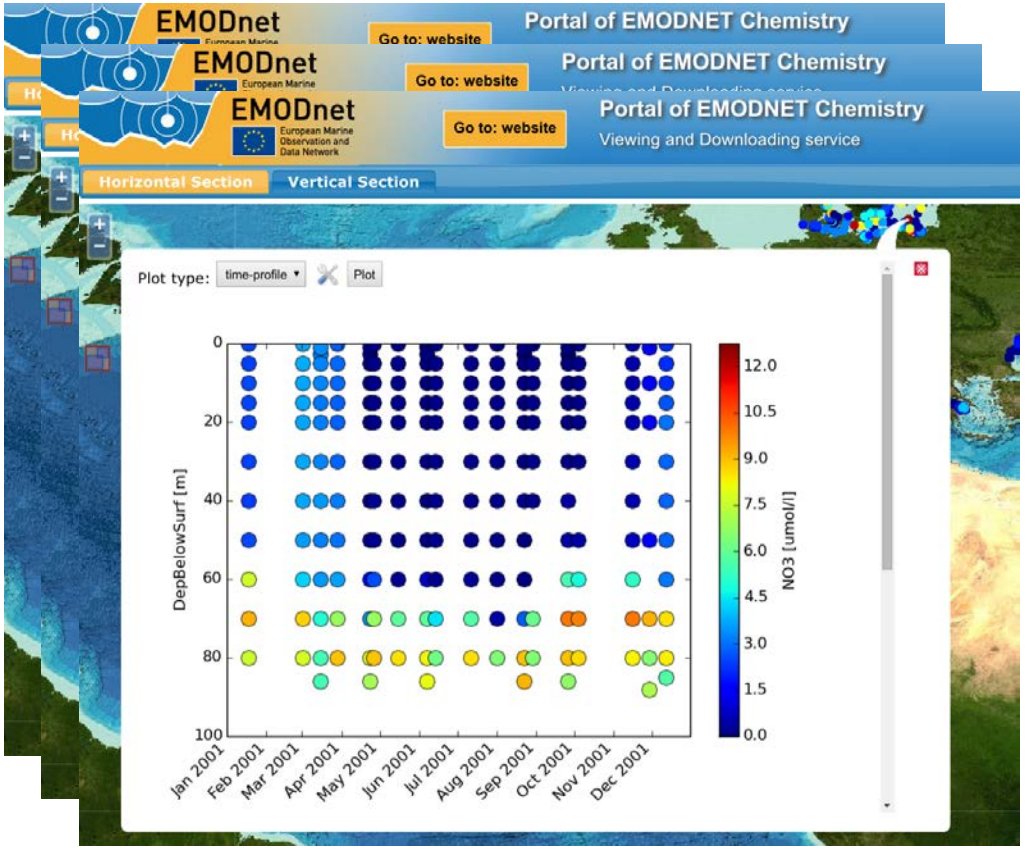
Profile and time series plots

- Plots can be changed dynamically
- Varying parameters: depth, time and measured value
- Profile (x: value, y: depth, color: time)
- time series (x: time, y: value, color: depth)



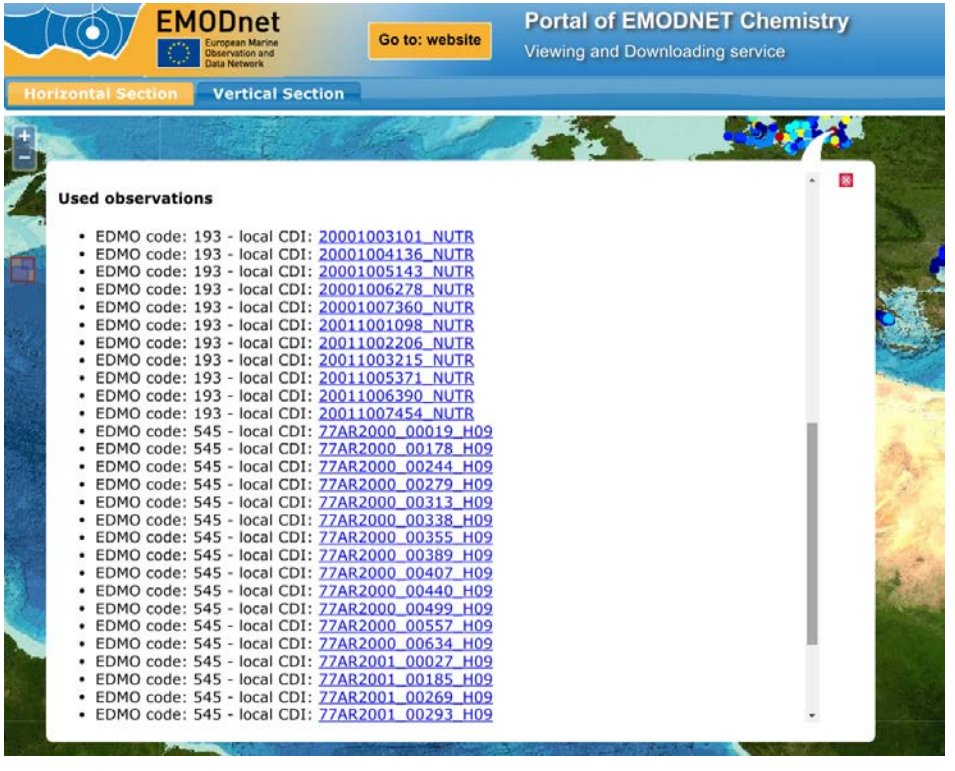
Profile and time series plots

- Plots can be changed dynamically
- Varying parameters: depth, time and measured value
- Profile (x: value, y: depth, color: time)
- time series (x: time, y: value, color: depth)
- time section (x: time, y: depth, color: value)



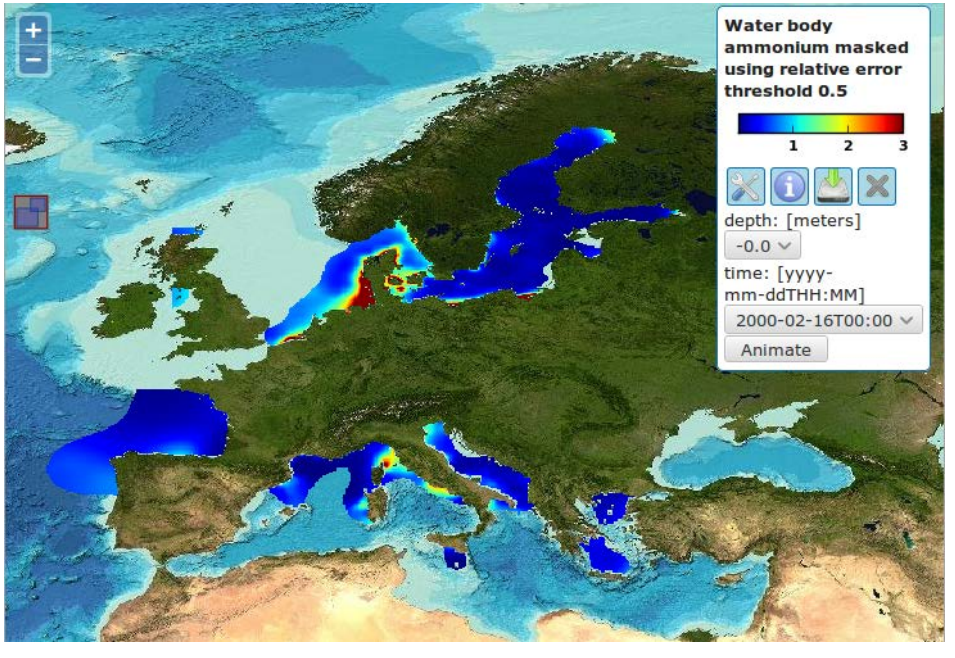
List of all observations

- Observation in SeaDataNet and EMODNET Chemistry are identified by an:
  - EDMO code: institution
  - CDI (Common Data Index) identifier
- For each plot: the **list of all used observation** included with a link to the central repository



Combined EMODNET data products

- Currently: one data product per domain and season
- EMODNET products represent 10-year average using all observation of the same season
- One data file per parameter
- Combine all seasons and domains



NetCDF compression

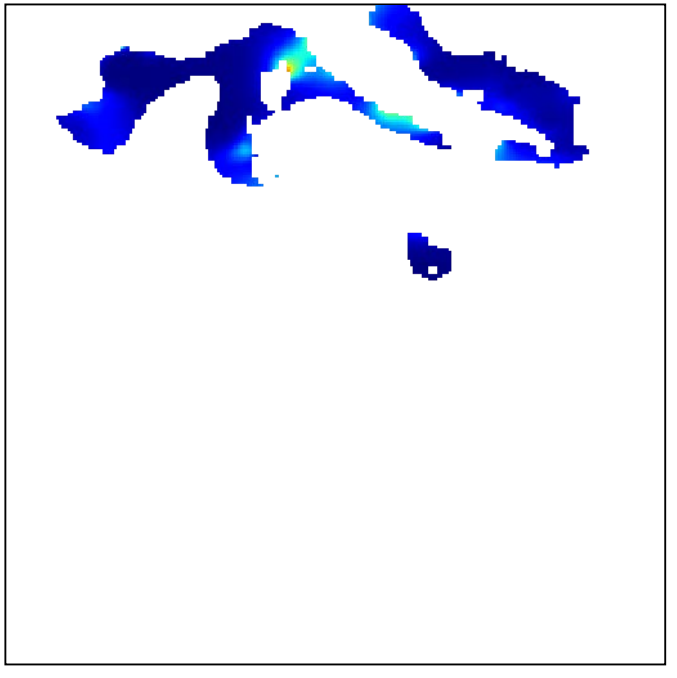
- In NetCDF, data is stored as a **multi-dimensional array** (e.g. longitude, latitude, depth and time)
- NetCDF 4 supports **compression** (based on zlib)
- Not the whole file is compressed, bot only **chunks** of data (i.e. blocks of the multi-dimensional array)
- Metadata is never compressed
- When reading data, only the **chunk** to be read have to be decompressed
- **shuffling** as an option (byte interlacing: store first byte of all values, then the second byte of all values, ...)
- Example of shuffling for decimals:

```
10.3, 10.4, 11.2  -> 111, 001, 342
```

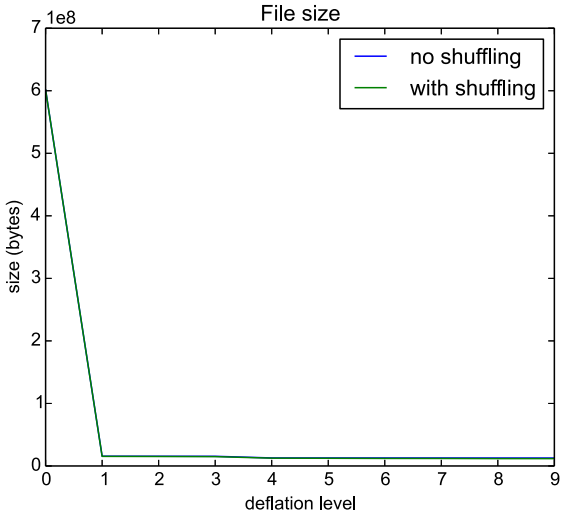
- Deflation levels: 0 (no compression) to 9 (highest compression)
- With compression:
  - smaller file size (ease storage requirements),
  - less data has to be read from the disk
  - but the CPU has to decompress the data
- Particularity of the EMODNET data set: many grid points are equal to the fill value (either land points or masked because of insufficient observations nearby)

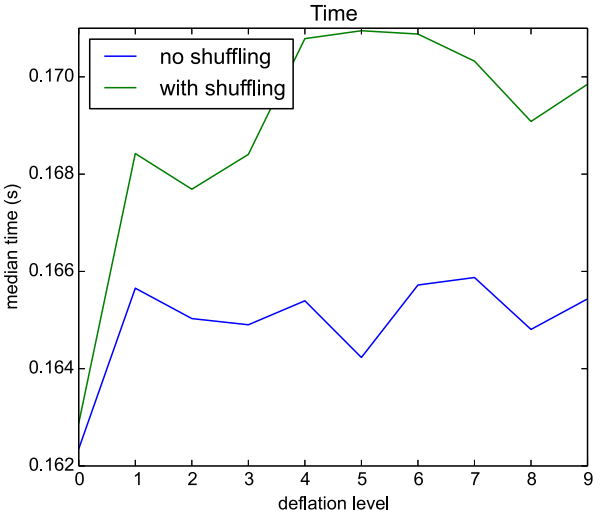
Benchmark

- Data set: surface ammonium, chunked over time
- i.e. **every time frame is compressed independently**
- Generate a 512 x 512 PNG image using a the WMS `GetMap` request
- The tile corresponds to the Mediterranean Sea
- The image is generated 1000 times and the median time is shown
- WMS tile **cache is deactivated**



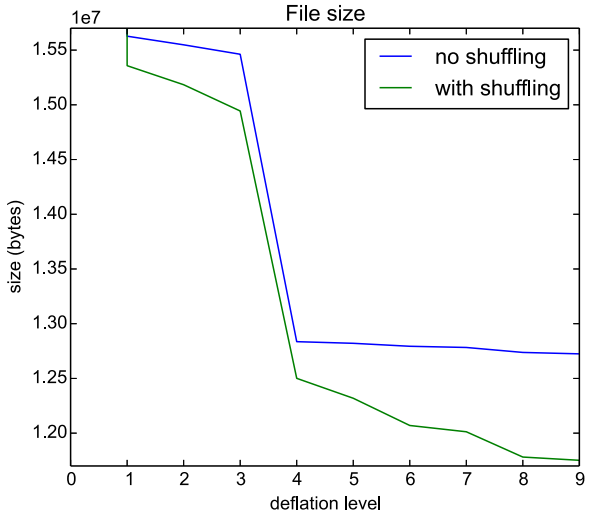
File size vs access time

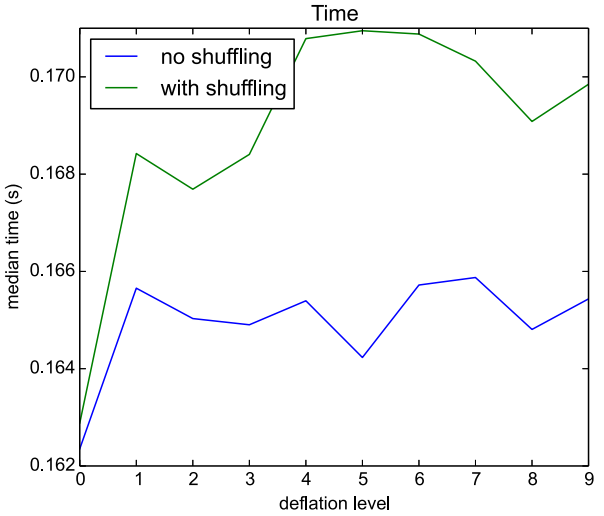




- dramatic decrease of file size even with lowest compression by a factor of 38 (574M to 15M)
- A significant portion of the data set is indeed land or masked

File size vs access time





- Another significant file size **decrease at deflation level 4 by 20%**
- Shuffling reduces the file size even more
- The WMS map generation time is **slightly increased using compression**
  - with shuffling, only by 5% (at most)
  - without shuffling, only by 2% (at most)
- Reasonable tread-off: use compression level 5 without shuffling
- However: user downloading directly the NetCDF file, need to have the the NetCDF4 (and HDF5) libraries with compression enabled.

Installation

- OceanBrowser is open source and freely available
- The hard way: install 13 packages, configure Apache and OceanBrowser and set file permissions
- The easy way:

```
docker run -p 8080:80 --name my-oceanbrowser-container \
-v /some/netcdf/files/:/var/www/data:ro abarth/oceanbrowser
```

  - Docker automatically downloads and runs OceanBrowser in a Linux container
  - OceanBrowser is available at port 8080 on localhost and servers files in the directory `/some/netcdf/files/`
  - Configuration using environment variables (public URL, name,...)
  - More info at the docker http://registry.hub.docker.com and search for OceanBrowser

Summary

- OceanBrowser allows the visualization of gridded data sets:
  - along a **horizontal section** (at given time and depth)
  - along a **vertical section** (e.g. at a fixed distance from coast)
- Various download options (full **NetCDF file**, subset via **OPeNDAP**, **Image** (PNG, EPS, SVG, ...) and **Animation** (webm, mp4))
- HTTP cache control headers work well with the Web Map Service standard
- Installation simplified using **Linux containers** (Docker)
- **Open source** (AGPL) and based on python and matplotlib
- Using on **OGC standards** (WMS, WFS and NetCDF)
- **Density of observations** (for a specified depth and time range)
- Ability to show **profile** and **time series plots**
- **NetCDF 4 compression is very beneficial** in the context of serving ocean climatologies by WMS
  - significant file size reduction
  - only small overhead when creating image tiles

More information

- Barth, A., Watelet, S., Troupin, C., Alvera-Azcárate, A., & Beckers, J. (2017). Analysis of Ocean in Situ Observations and Web-Based Visualization: From Individual Measurements to an Integrated View. In P. Diviacco, A. Leadbetter, & H. Glaves (Eds.) **Oceanographic and Marine Cross-Domain Data Management for Sustainable Development** (pp. 345-371). Hershey, PA: . doi:10.4018/978-1-5225-0700-0.ch015
- or ask me directly (a.barth@ulg.ac.be).

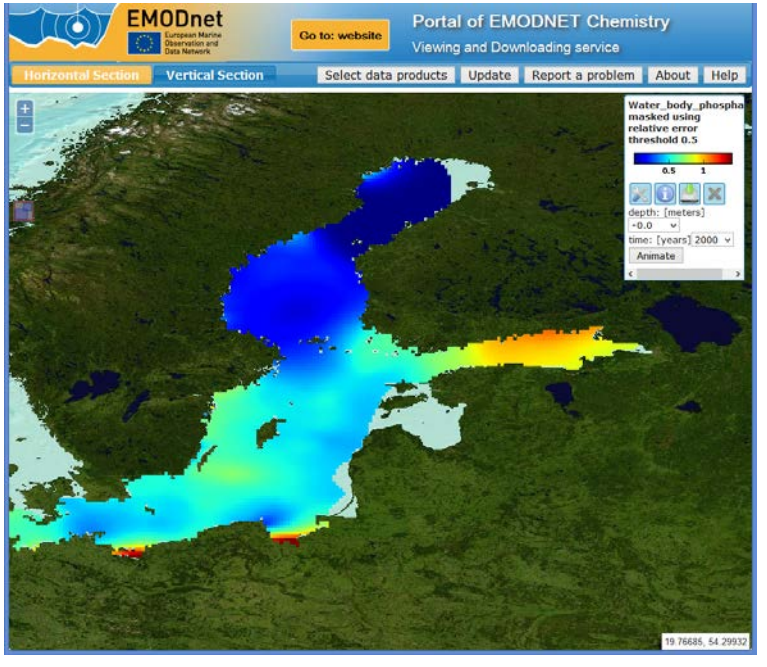
What is DIVA?

- DIVA: Data Interpolating Variational Analysis
- Objective: **derive a gridded climatology from in situ observations**
- The variational inverse methods aim to derive a continuous field which is:
  - **close to the observations** (it should not necessarily pass through all observations because observations have errors)
  - "**smooth**"
- DIVA works internally on a finite element mesh:
  - decouples basins based on **topography**
  - can take **ocean currents** into account
  - can detect **trends** in your data
  - can detect and remove **outliers**
  - consistent **error variance estimation**

Horizontal section

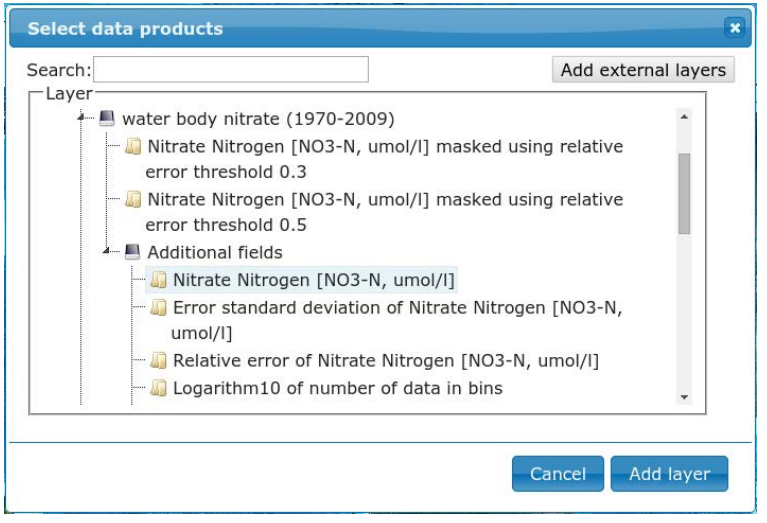
The right panel controls the current layer:

- **Select depth and time**
- Plot style
- Metadata
- Download of data product



Layer selection

- Simple directory **structure on the server** is mapped into a hierarchical list of layers
- NetCDF files can be added on-the-fly (without a server restart)
- **Virtual sub-folder** can be added to show some NetCDF variables more/less prominently



For DIVA field:

- 1st level: analysis masked by an error threshold
- 2nd level: Full field available under “Additional fields”