From in situ observations to gridded fields using DIVAnd and the cloud

Charles Troupin (1), Alexander Barth (1), Sylvain Watelet (1), Jean-Marie Beckers (1), Merret Buurman (2), Yohann Legrand (3), Mikael Karlsson (4), Sebastian Mieruch (5), Giorgio Santinelli (6), and Peter Thijsse (7)

(1) University of Liège, GeoHydrodynamics and Environment Research, Astrophysics, Geophysics and Oceanography, Liège, Belgium (ctroupin@uliege.be), (2) Deutsches Klimarechenzentrum GmbH, Hamburg, Germany, (3) Altran, France, (4) CSC, Finland, (5) Alfred Wegener Institut, Bremerhaven, Germany, (6) Deltares, Delft, The Netherlands, (7) Mariene Informatie Service 'MARIS' Bv, Voorburg, The Netherlands

Interpolating sparse in-situ observations is a frequent practice in oceanography: the resulting gridded fields are used in a wide range of applications including model initialisation, quality control, data visualisation. After two decades of development of the DIVA (Data-Interpolating Variational Analysis) software tool, we created a brand new tool called DIVAnd ($n-$dimensional DIVA).

In both tools the method consists of the minimization of a cost function that penalizes the deviation from the observations, the deviation from a first guess and abruptly varying fields (in comparison with a given correlation length), with the capability to add physical constraints representing advection, diffusion or source terms. One major advantage of the method is that it naturally decouples basins that are not connected.

DIVAnd extends the 2D capabilities of the DIVA and allows the interpolation of observations on curvilinear orthogonal grids in an arbitrary high dimensional space. DIVA was written in Fortran and called via shell scripts; DIVAnd has been rewritten from scratch in Julia (http://julialang.org/) which allows a high-level programming style along with good performance thanks to the '*just in time*' compilation.

DIVAnd in jupyter-notebooks

Jupyter notebooks are web-based interactive computational environment made of code fragments, markdown text and images. They are becoming increasingly popular in the research community as they allow one to provide, in a single document, the code, the documentation and the results. These notebooks are considered to be the ideal solution to document the way DIVAnd has to be used and in particular how the climatologies can be generated. The use of a notebook also results in a smaller number of input files, as most of the analysis parameters are defined within the notebook. DIVAnd only requires:

- a bathymetry dataset covering the region of interest and

- one (or more) data file(s), different format among which SeaDataNet netCDF (one of the standards developed in the projects) or ODV spreadsheet, are accepted. These formats ensure interoperability with other tools processing and visualisation tools.

DIVAnd on the cloud

The main hurdles for the use of the tool may be the installation and the need of relatively high computing resources if large regions are to be analysed. For these reasons, a Virtual Research Environment (VRE) is being set up in the frame of SeaDataCloud project (https://www.seadatanet.org). The principle is that a user logs in and enters the VRE and has direct access to SeaDataCloud data and to different tools designed to process and visualise data and data products. A typical workflow for the creation of a climatology would consist of the following steps:

1. Login using MarineID (https://www.marine-id.org).

2. Upload of "private" dataset.

3. Use of ODV-on-Web tool to perform a pre-processing (quality control, duplicates removal, . . .) and export the resulting data in a netCDF file.

4. Open a Jupyter notebook and adapt it for their region of interest: datafile name, spatial coverage, bathymetry, parameters. . .

5. Run of the notebook to generate a 4D gridded field in netCDF.

6. Use of the visualisation tools developed by Deltares.

We will present a realistic application using temperature and salinity data in the Atlantic Ocean.