Generating ocean climatologies from in situ observations

Alexander Barth, Université de Liège (Belgium), a.barth@uliege.be
Charles Troupin, Université de Liège (Belgium), c.troupin@uliege.be
Sylvain Watelet, Université de Liège (Belgium), swatelet@uliege.be
Aida Alvera-Azcárate, Université de Liège (Belgium), a.alvera@uliege.be
Jean-Marie Beckers, Université de Liège (Belgium), jm.beckers@uliege.be

Methodology

DIVAnd (Data-Interpolating Variational Analysis - n-dimensional) is an analysis tool that allows one to generate gridded data products from in situ observations. DIVAnd extends the 2D capabilities of the DIVA tool and allows the interpolation of observations on curvilinear orthogonal grids in an arbitrary high dimensional space by minimizing a cost function. This cost function penalizes the deviation from the observations, the deviation from a first guess and abruptly varying fields based on a given correlation length (potentially varying in space and time). Physical constraints can be added to this cost function such as an advection constraint, diffusion or source terms. One major advantage of the method is that it naturally decouples basins that are not connected and where water masses often have very different properties.

While the classical DIVA was written in Fortran and called via shell scripts, DIVAnd was rewritten from scratch using generalized mathematical formulation, in the programming language Julia which allows a high-level programming style but it compiled to machine code for better performance. The quality of the analysis depends strongly on the two main analysis parameters:

1. the correlation length, which represents over which space-time distance two location are significantly correlated) and
2. the observational error variance, i.e. how accurate and representative the observations are. Several tools have been written to estimate the correlation and observational error variance to guide the user to choose suitable values for these parameters.

Observational data

DIVAnd can load observations in NetCDF or ODV spreadsheet formats. There are also functions implemented to directly query online the following databases: World Ocean Database, Copernicus Marine Environment Monitoring Service (CMEMS) and EMODnet Physics. The bathymetries, needed to delimit the domains of interpolation, based on GEBCO and EMODnet Bathymetry have been adapted to work with DIVAnd.

Additional tools around DIVAnd have been developed for automatic outlier detection and duplicates check. The outlier detection is based on a DIVAnd analysis to evaluate how consistent a data value is with other data values in its surroundings. The duplicated check implements an n-dimension Quadtree to quickly searches for observations nearby a given location.

Applications

An outlook on new developpements of DIVAnd will also be given, including the combination of DIVAnd with non-linear analysis techniques from machine learning (neural networks). New practical application in the context of SeaDataCloud for surface currents high-frequency radar data, EMODnet Physics data and EMODnet Biology data with initial test of DIVAnd coupled with a neural networks will be given.