1 Introduction

Diva is a software designed to create grid fields from sparse or noisy data and relies on a finite-element technique to solve a variational principle. We present an overview of the software capabilities as well as the latest modules developed in the frame of the SeaDataNet project:

- Semi-normed analysis:
- Ordinary cross validation;
- Advection constraint;
- Methods for error estimation.

2 Theory

2.1 Variational inverse method

The field $\psi$ reconstructed by Diva using $N_d$ data $d_i$ located at $(x_i, y_i)$ is the solution of the variational principle:

$$ J[f] = \sum_{j=1}^{N_d} |d_j - \psi(x_j, y_j)|^2 + \|\psi\|^2 $$

with

$$ \|\psi\|^2 = \int_{\Omega} \left( \frac{1}{2} \nabla \psi \cdot \nabla \psi - \frac{1}{2} \mu \psi^2 \right) d\Omega $$

where $\alpha$ and $\mu$ are determined from the data themselves, through their correlation length $L$ (tool divacross) and signal-to-noise ration $\lambda$ (tools divacross, divacross and divacrossr).

Resolution of Eq. (1) relies on a highly optimized finite-element technique, which permits computational efficiency independent on the data number and the consideration of real boundaries (constains and bottom).

2.2 Advection constraint

The advection constraint aims at modeling the effects of velocity on the reconstructed field. In theory, this constraint is activated by adding a term to the norm (2), leading to

$$ J = J[\psi] + \oint_{\Gamma} \left( \frac{1}{2} \nabla \psi \cdot \nabla \psi - \frac{1}{2} \mu \psi^2 \right) d\Gamma $$

where $U$ is a velocity scale deduced from the provided $(u, v)$ field.

$L$ is a characteristic length.

$\theta$ is a parameter that controls the weight of the additional term;

$\lambda$ is a diffusion coefficient.

Parameters $\theta$ and $\lambda$ as well as a velocity field on a regular grid (Fig. 5) have to be specified by the user.

3 Implementation

3.1 Data

Data were gathered from various databases in the region $0 - 60^\circ N \times 0 - 50^\circ W$. We processed them to remove duplicates, detect outliers and perform vertical interpolation with Weighted Pseudovals method [Renser and Ross (1963)]. We consider only surface temperatures during winter (Fig. 1).

3.2 Contours and mesh

Contours are created from Nodal Oceanographic Office Sما resolution topography on the standard depth levels (tool divasurface). Mesh is generated with a characteristic length $L = 3$ (command divasurface).

3.3 Parameters

Correlation length $L$ is determined through a fit of the data covariance function with the execution of divacross, yielding a value of $L = 1.5$, equivalent to 184×184 km. Signal-to-noise ration $\lambda$ is estimated with one of the following tools:

1. Ordinary Cross Validation (CV) applied on the whole dataset (divacross) or on a sample containing $n$ values (divacrossn).
2. Generalized Cross Validation (GCV, tool divacross).

Both methods are designed to provide the value of $\lambda$ that minimizes an estimator $\theta$ (Fig. 2). We get of value of $\lambda = 2.971$.

4 Analysis

4.1 Semi-normed analysis

Semi-normed analysis consists in four steps:

1. divacross → creates a so-called reference field (Fig. 4(a)), using large correlation length and small signal-to-noise ratio;
2. divacross → subtracts the reference field from the data values in order to work with anomalies (Fig. 4(b));
3. divacross → performs an analysis on the anomalies (Fig. 4(c));
4. divanorm → adding the analysed anomaly field to the background field (Fig. 4(d)).

The four steps are regrouped in script divasemiacc.

4.2 Advection constraint

The mean velocity field used for advection modeling is obtained from differ measurements. Several values of parameter $\theta$ were tried (Fig. 5) to observe the effect of weak, moderate or strong advection constraints.

5 Error field

One of Diva major asset is the possibility of an error field computation. Previously, the method was based on analogies with Optimal Interpolation and the error field was estimated as the difference of covariance fields. Now three methods are available, depending on the data considered and type of analysis performed:

1. A poor man’s error indicator, where the covariance to be analysed are replaced by 1; this method produces underestimated error field (Fig. 7(a)).
2. An hybrid error calculation [Rixen et al. (2000)], based on an additional analysis per point in which the error is to be calculated, which takes profit of the already performed LU decomposition (Fig. 7(b)).
3. An error calculation with the Diva covariance function: it demands two analyses per point in which the error is needed: one is done with the already existing LU decomposition of one Diva execution, the other with an existing LU decomposition of another Diva execution (Fig. 7(c)).

This method is recommended when working with variable correlation length or with advection (Fig. 7(d)).

6 Conclusion

We considered a large dataset covering the North Atlantic to illustrate the efficiency of various Diva software tools for producing realistic grid fields.

Analysis parameters (correlation length and signal-to-noise ratio) were determined in an objective way using tools provided with the software. Various error computation were tried their results underline the influence of data coverage.

Further work will be concentrated on multiple-level analysis in order to create a complete climatology for the region of interest.

Acknowledgments

Diva was developed during EU MODIS project and improved with the working package “Geostatistic analysis tools” (IRAT) of SeaDataNet project, an Integrated Infrastructure Initiative of the E.U. Sixth Framework Programme. We thank the participants to the Diva workshops in Liège (November 2006) and in Calvi (November 2007) for their numerous valuable comments which help us to improve the software.

* Contact: c.troquin@ulg.ac.be

http://modb.oece.ulg.ac.be/projects/1/diva