Advanced Data Interpolating Variational Analysis.
Application to climatological data.

C. Troupin[a] (ctroupin@ulg.ac.be), D. Sirjacobs[b], M. Rixen[c], P. Brasseur[d], J.-M. Brankart[d],
A. Barth[a], A. Alvera-Azcárate[a], A. Capet[a], M. Ouberdous[a], F. Lenartz[a], M.-E. Toussaint[a] & J.-M. Beckers[a]

[a] GeoHydrodynamics and Environment Research, MARE, University of Liège, BELGIUM
[b] Algology, Mycology and Experimental Systematics, Department of Life Sciences, University of Liège, BELGIUM
[c] NURC - NATO Undersea Research Centre, La Spezia, ITALY
[d] LEGI - CNRS - Université de Grenoble, FRANCE

1. Objectives

Diva is a method applied to grid sparse and noisy data. It has been successfully used to create regional climatologies (North Sea, Black Sea, Baltic Sea, Mediterranean Sea, ...). Here we will:

✓ Present the new features of Diva software.

✓ Demonstrate its capabilities on a realistic application.

2. Data and method

Data set = combination of SeaDataNet (http://www.seadatanet.org) + World Ocean Database 2009 (WOD09, Boyer et al., 2009).
Period: 1950-2010.   Month: September.   Depth: 30 m.
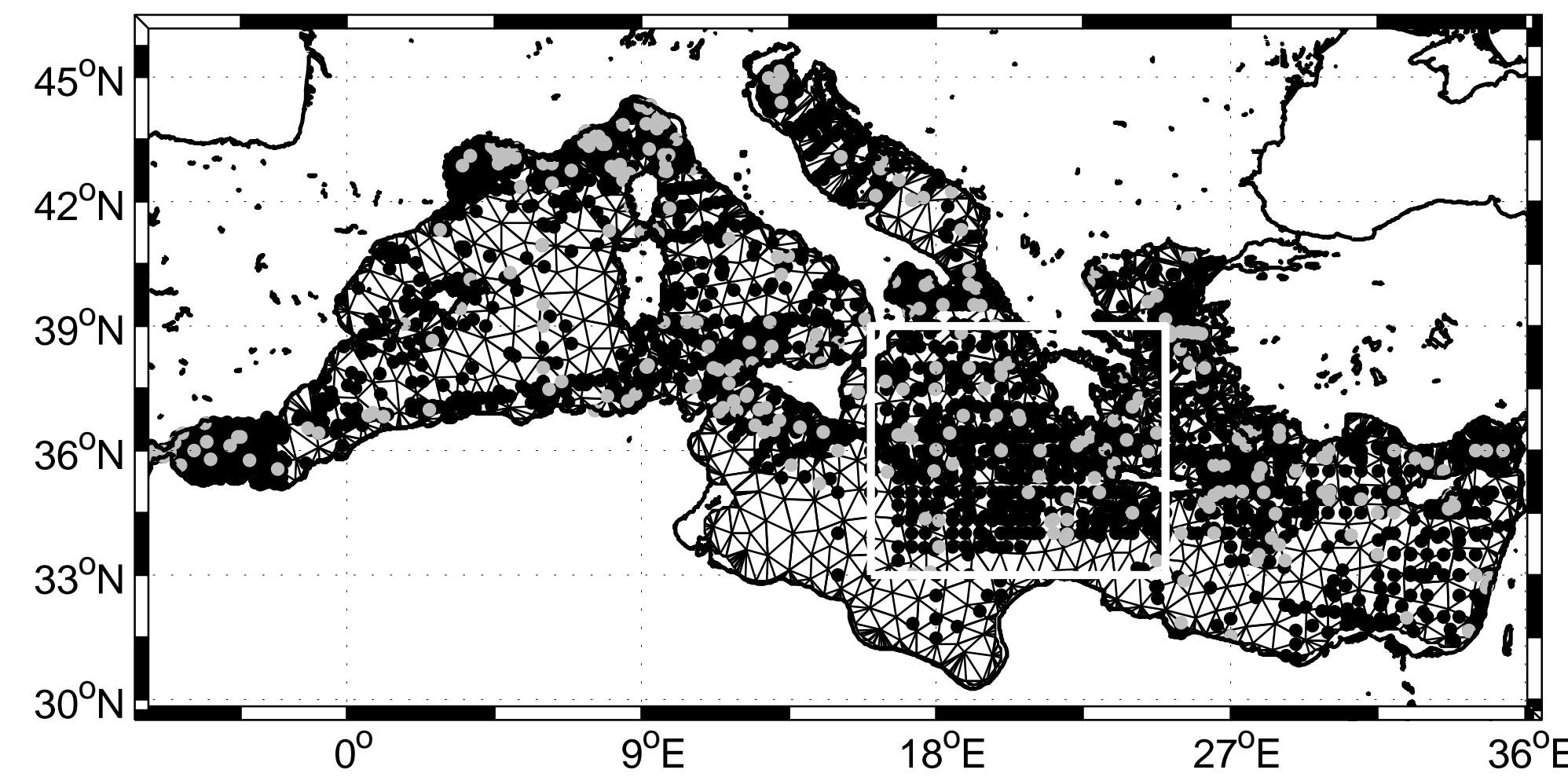Total: 4699 data points, 352 randomly chosen for validation (Fig. 1)

FIGURE 1: Finite-element mesh and data locations. Black dots indicate data used for the analysis, while grey ones are those discarded for validation.

Diva = Data Interpolating Variational Analysis Brasseur et al., 1996; Brankart and Brasseur, 1998
= particular methodology (Eq. 1)
+ efficient finite-element solver (Fig. 1)

Formulation: for N data d_i located at (x_i, y_i), construct field $\varphi(x, y)$ by minimizing the functional:

$$\min J[\varphi] = \sum_{i=1}^{N} \mu_i \overbrace{[d_i - \varphi(x_i, y_i)]^2}^{\text{data-analysis misfit}}$$
$$+ \underbrace{\int_D \left(\boldsymbol{\nabla\nabla}\varphi : \boldsymbol{\nabla\nabla}\varphi + \alpha_1 \boldsymbol{\nabla}\varphi \cdot \boldsymbol{\nabla}\varphi + \alpha_0 \varphi^2 \right) dD}_{\text{field regularity}}. \quad (1)$$

where μ_i are the data weights, $\boldsymbol{\nabla}$ is the gradient operator.
In non-dimensional form:

$$\tilde{J}[\varphi] = \sum_{i=1}^{N} \mu_i L^2 [d_i - \varphi(x_i, y_i)]^2 \quad (2)$$
$$+ \int_{\tilde{D}} \left(\boldsymbol{\tilde{\nabla}\tilde{\nabla}}\varphi : \boldsymbol{\tilde{\nabla}\tilde{\nabla}}\varphi + \alpha_1 L^2 \boldsymbol{\tilde{\nabla}}\varphi \cdot \boldsymbol{\tilde{\nabla}}\varphi + \alpha_0 L^4 \varphi^2 \right) d\tilde{D}.$$

3. Parameter estimation

Coefficients μ_i, α_1 and α_0 related to data through:

● the correlation length L: estimated by fitting of the data covariance to an analytical function → $L = 150\ km$

● the signal-to-noise ratio λ: estimated ordinary or generalized cross-validation (Fig. 2) → $\lambda = 4.02$.
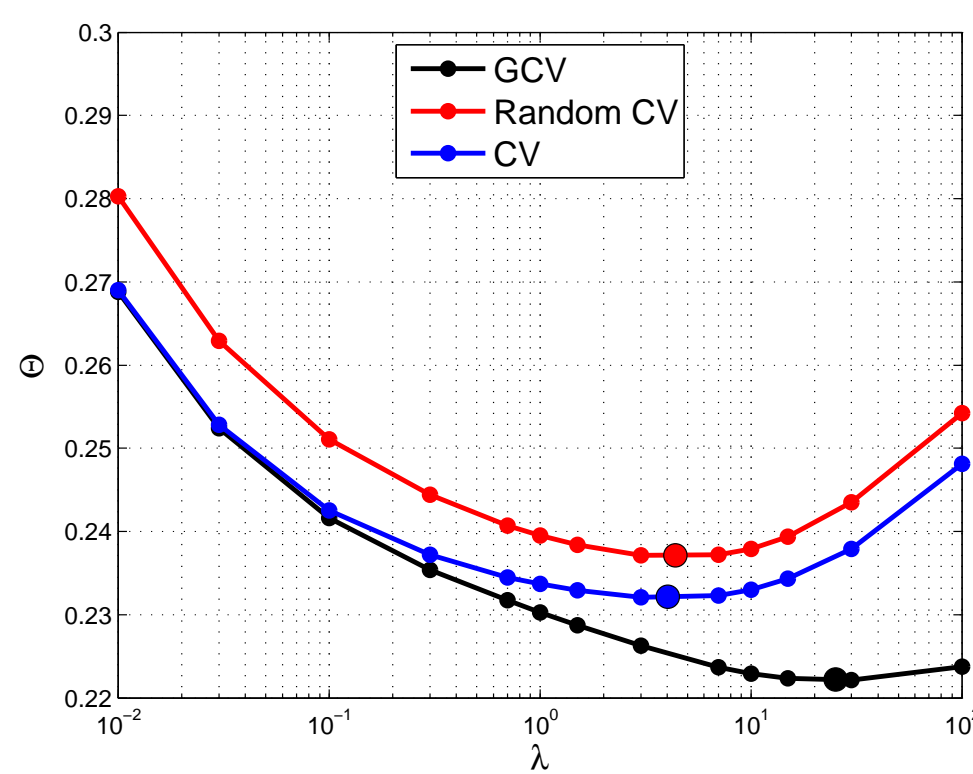

FIGURE 2: Generalised cross-validator for a set of given λ values. GCV = generalised cross-validation, CV = ordinary cross-validation, Random CV = CV performed on a random subset of data points.

The corresponding analyse (Fig. 3) shows:

✓ penetration of Atlantic water through the Strait of Gibraltar,

✓ decrease of salinity in the northern Adriatic Sea due to river input,

✓ signal of the anticyclonic eddy south of Cyprus ,

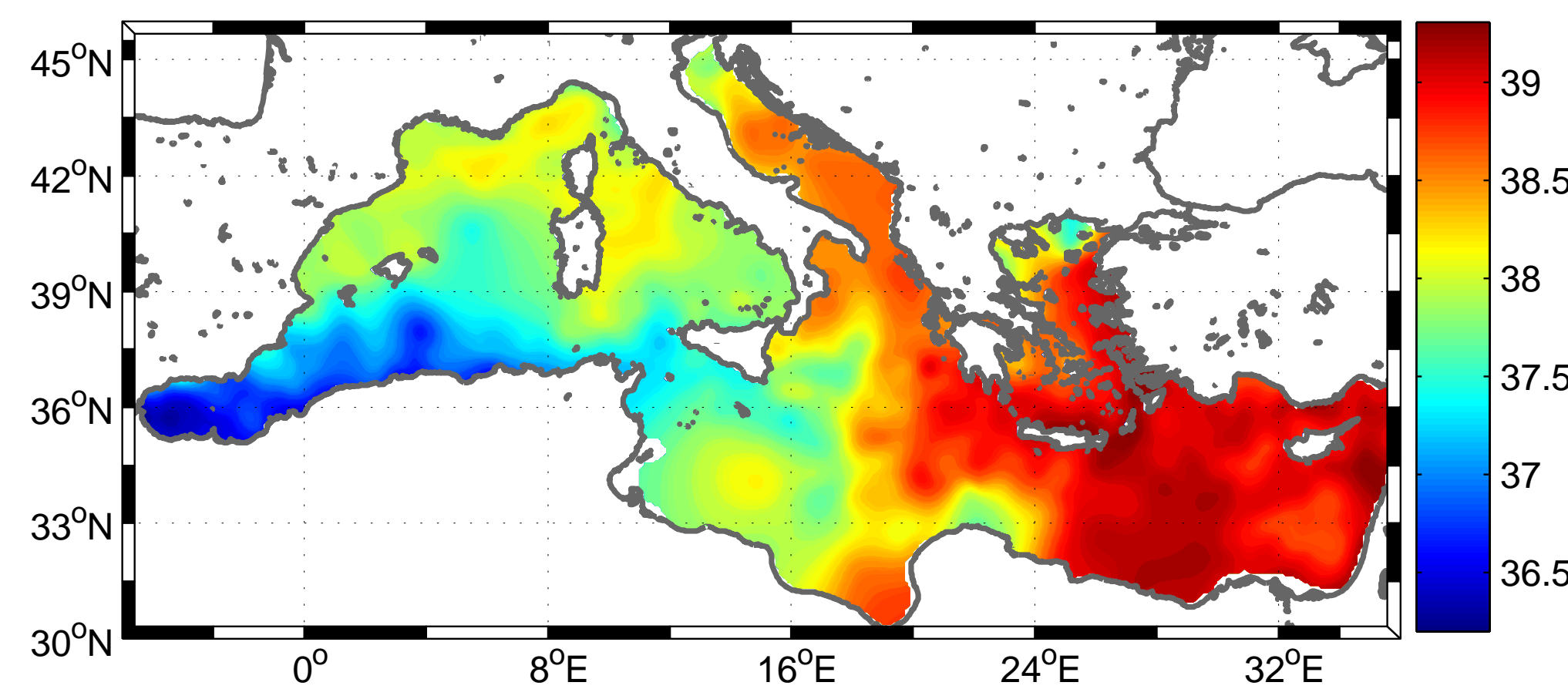✓ sharp separation of eastern and western Mediterranean waters.

FIGURE 3: Analysed field of salinity obtained with $L = 1.346°$ and $\lambda = 4.02$.

4. Advection constraint

Advection is taken into account in the functional by adding a term in (2):

$$\tilde{J}_a = \tilde{J}(\varphi) + \frac{\theta}{\mathcal{U}^2} \int_{\tilde{D}} \left[\vec{u} \cdot \boldsymbol{\tilde{\nabla}}\varphi - \frac{\mathcal{A}}{L} \boldsymbol{\tilde{\nabla}} \cdot \boldsymbol{\tilde{\nabla}}\varphi \right]^2 d\tilde{D}. \quad (3)$$

where $\mathcal{U}$ is the velocity scale and θ measures the strength of the advection.

Experiment:

● Data located in a determined zone are discarded (Fig.5).

● Velocity field: 1/16°-resolution model (F. Lenartz et al., in prep.).

● Two analyses: 1. without advection, 2. with advection constraint.

The advection constraint improves the quality of the analysed field by transporting the information along the streamlines.
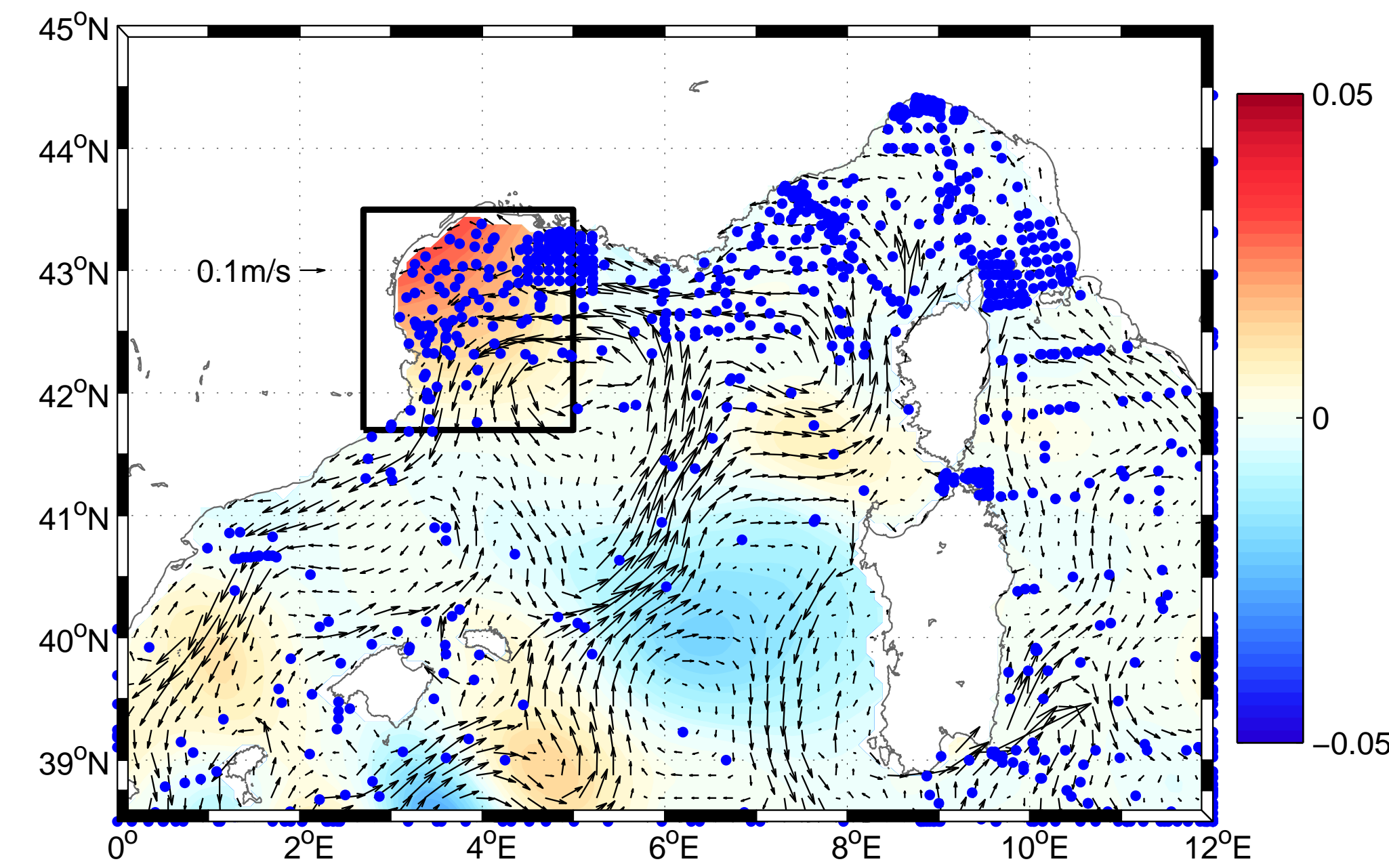

FIGURE 4: Difference between the analysed salinity fields in September at 30 m without and with advection constraint. The black rectangle represents the area where the data have been removed, while the blue dots indicate the data positions.

5. Error fields

New developments allows the estimation of the real covariance function (Troupin et al., 2011). Errors are compared in Fig. 5. Poor man's method underestimates the error. The three other methods provides similar results, with the largest errors along the coasts of Libya.
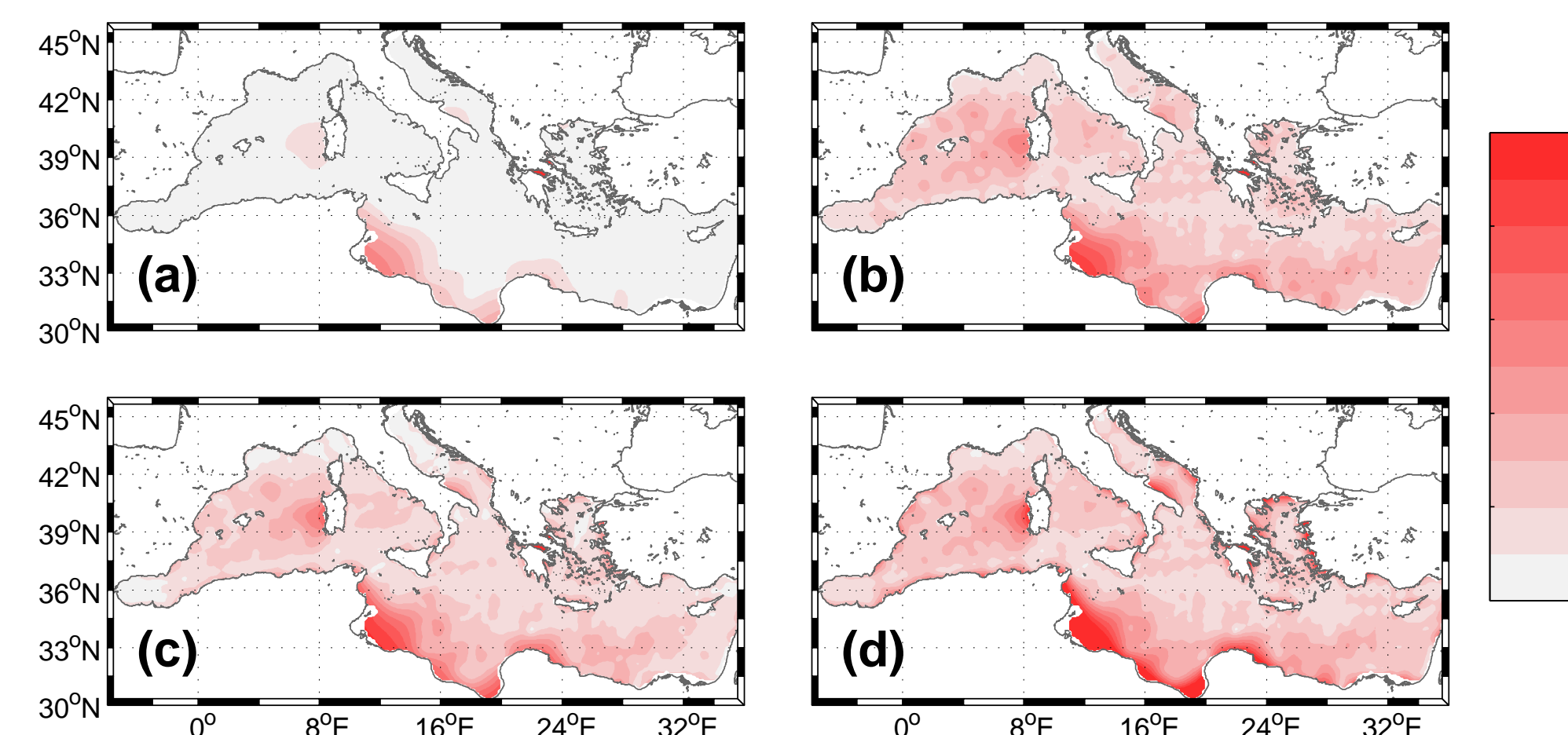

FIGURE 5: Error fields computed using four different methods: (a) poor man's estimate, (b) hybrid method, (c) real covariance and (d) real covariance with boundary effects.

6. Outliers detection

Random data points are displaced or have their salinity artificially changed. The outlier detection criterion implemented in Diva allows us to catch them (Fig. 6).
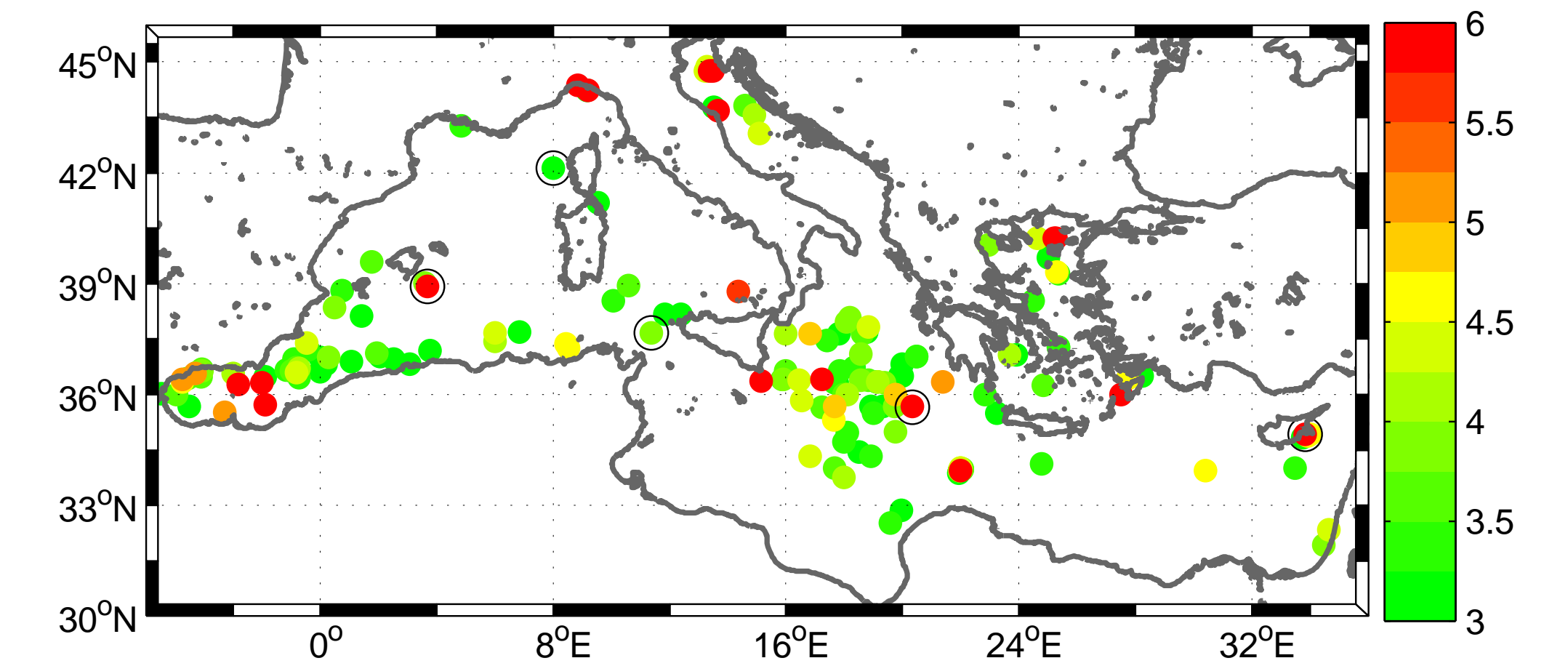

FIGURE 6: Outliers detected using a criterion based on the misfit between the data and the reconstruction. Black circles indicate data points of which the position/value has been changed.

7. Comparison with OI

The example show the influence of the boundaries on the analyzed field. The difference of two fields are weak in the open ocean, but larger close to the coast.
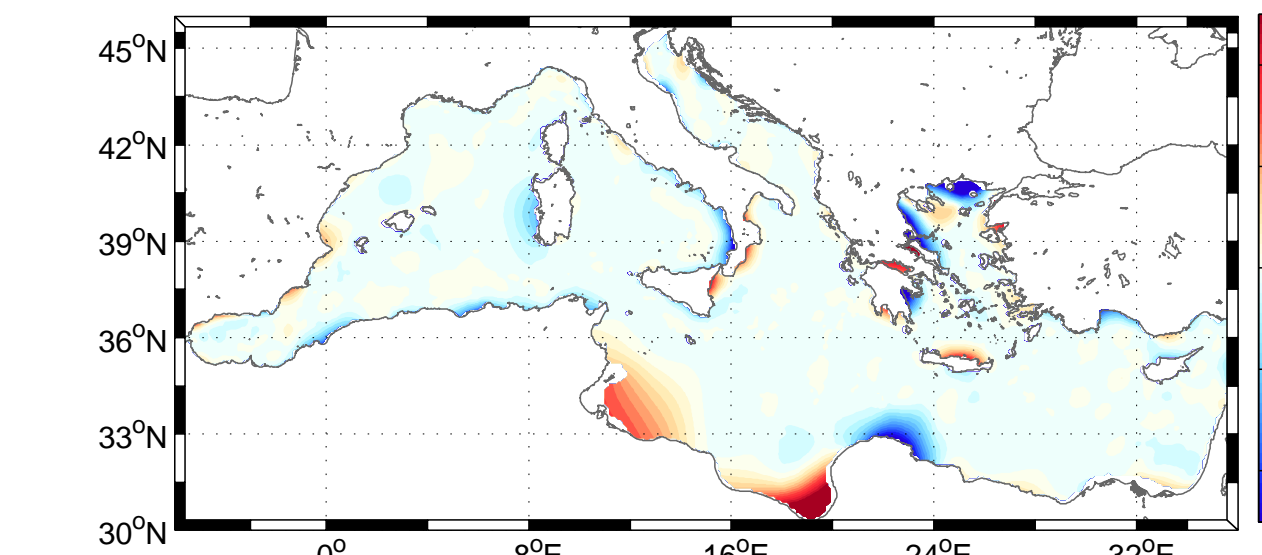

FIGURE 7: Difference between the salinity fields obtained by Diva and by OI.

Main references

Brankart, J.-M. and Brasseur, P.: The general circulation in the Mediterranean Sea: a climatological approach, J. Mar. Syst., 18, 41–70, doi:10.1016/S0924-7963(98)00005-0, 1998.

Brasseur, P., Beckers, J.-M., Brankart, J.-M., and Schoenauen, R.: Seasonal temperature and salinity fields in the Mediterranean Sea: Climatological analyses of a historical data set, Deep-Sea Res. I, 43, 159–192, doi:10.1016/0967-0637(96)00012-X, 1996.

Troupin, C., Sirjacobs, D., Rixen, M., Brasseur, P., Brankart J.-M., Barth, A., Alvera-Azcárate, A., Capet, A., Ouberdous, M. , Lenartz, F., Toussaint, M.-E., and J.-M. Beckers: Advanced Data Interpolating Variational Analysis. Application to climatological data, submitted to Ocean Model.