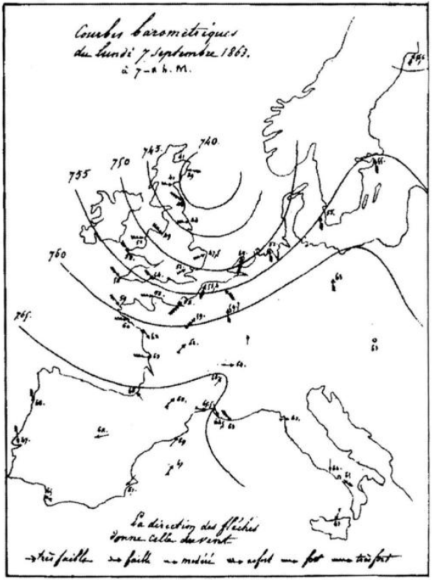
Diva Lecce 2016

Diva in 2 dimensions

Alexander Barth, Aida Alvera-Azcárate, Mohamed Ouberdous,
Charles Troupin, Sylvain Watelet & Jean-Marie Beckers

Acknowledgements: SeaDataNet, EMODnet Chemistry,
EMODnet Biology, STARESO

Interpolation 150 years ago…



What is Diva?

Data
Interpolating
Variational
Analysis

DIVA

What is Diva?

- a method to produce gridded fields
- a set of bash scripts and Fortran programs

What is not Diva?

- a plotting tool
- a *black-box*
- a numerical model

A little bit of history

Code development (1990-1996)

- Variational Inverse Method (VIM) (Brasseur, 1991, JMS, JGR)
- cross-validation (Brankart and Brasseur, 1996, JAOT)
- error computation (Brankart and Brasseur, 1998, JMS; Rixen et al., 2000, OM)

A little bit of history

Code development (1990-1996)

2D-analysis (2006-2007)

- set of bash scripts (`divamesh`, `divacalc`, ...)
- Fortran executables
- parameters optimization tools
- Matlab/Octave scripts for plotting

A little bit of history

Code development (1990-1996)

2D-analysis (2006-2007)

3D-analysis (2007-2008)

- superposition of 2D layers
- automated treatment and optimization
- stability constraint (Ouberdous et al.)

A little bit of history

Code development (1990-1996)

2D-analysis (2006-2007)

3D-analysis (2007-2008)

4D-analysis (2008-2009)

- start from ODV spreadsheet
- *detrending* (with J. Carstensen, DMU)
- NetCDF 4-D climatology files

A little bit of history

Code development (1990-1996)

2D-analysis (2006-2007)

3D-analysis (2007-2008)

4D-analysis (2008-2009)

Web tools

- On-line analysis (Barth et al., 2010, Adv. Geosci.) `http://gher-diva.phys.ulg.ac.be/web-vis/diva.html`
- Climatology viewer: `http://gher-diva.phys.ulg.ac.be/web-vis/clim.html`

A little bit of history

Code development (1990-1996)

2D-analysis (2006-2007)

3D-analysis (2007-2008)

4D-analysis (2008-2009)

Web tools

2011-2012

- multivariate approach
- data transformation tools
- 4-D graphical interface
- implementation of *source/decay* terms
- advanced error computation (Troupin et al., 2012, OM)

A little bit of history

Code development (1990-1996)

2D-analysis (2006-2007)

3D-analysis (2007-2008)

4D-analysis (2008-2009)

Web tools

2011-2012

2013-2015

- Modernisation of the code structure
- n-dimensional generalisation
- optimized and approximate error calculations (clever poor man)

A little bit of history

Code development (1990-1996)

2D-analysis (2006-2007)

3D-analysis (2007-2008)

4D-analysis (2008-2009)

Web tools

2011-2012

2013-2015

On-going:

- Analysis at a specific distance from the bottom
- Correlated observations errors (data weighting)
- ... ☕

A little bit of history

Code development (1990-1996)

2D-analysis (2006-2007)

3D-analysis (2007-2008)

4D-analysis (2008-2009)

Web tools

2011-2012

2013-2015

On-going:

General: user-driven developments

Diva history

4.6.7

Released in October 2014.

New features

- Transformation of user relative length or advection fields files (ascii format) into the gher binary format, via a run of Diva (new script "asctobin")

New bug fixes

- Correction of time axis and climatology_bounds in Netcdf output files (diva3Dwrt.F,diva4Dwrt.F,dv4DYRwrt.F,dv3DncYRw.F)
- Correction of some attributes in 4D netcdf (databins, snr, cl, varbak) (dv3DncYRw.F, diva3Dsub)
- Update of driver files (also in Example4D)

4.6.6

Released in September 2014.

New features

- Check for severe errors in DIVA 3D/4D (script "godiva") + simple errors and warnings
- Possibility of binning the data before the parameters optimization (script "divabin" + program "binning_lines.f90")
- Variable correlation length, depending on depth (script "divarlvardepth" + program "rlvardepth.f90")

New bug fixes

- Correction of the example in 4D (datasource)
- Correction of the script divaguessformODV4
- Exact match needed between variable name in "varlist" and its real name in the data file.

4.6.5

Released in April 2014.

New features

New bug fixes

- "end of line" problems under Windows (file "datasource")
- Portability of scripts using the "sort" command

- http://modb.oce.ulg.ac.be/mediawiki/index.php/New_Diva_Features

Diva related tools

Diva: base tool (command line), 2D analysis

Godiva: automatic repetition of 2D analysis

Diva-on-web: 2D analysis with your data on our server

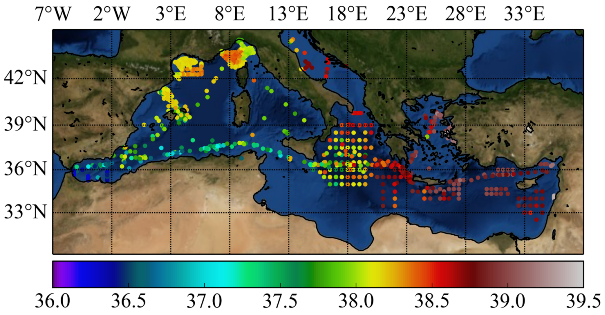
OceanBrowser: visualisation tool of 4D NetCDF files

divand: multi-dimension analysis (lon, lat, time, depth)

divaformatlab: wrapper to use in matlab

Clone-diva-x.x.x: virtual machine containing diva-x.x.x + other stuff (gfortran, netcdf,...)

Common problem



Appears when

- trying to produce maps
- calculate volume averages
- prepare initial conditions for models
- quality control of data
- ...

Solutions

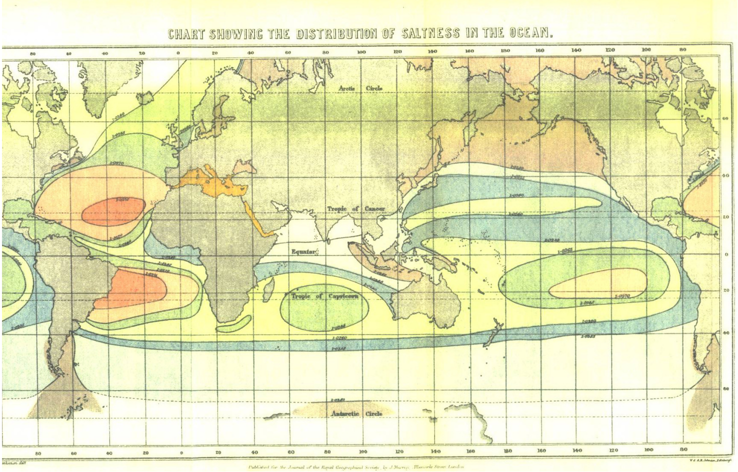
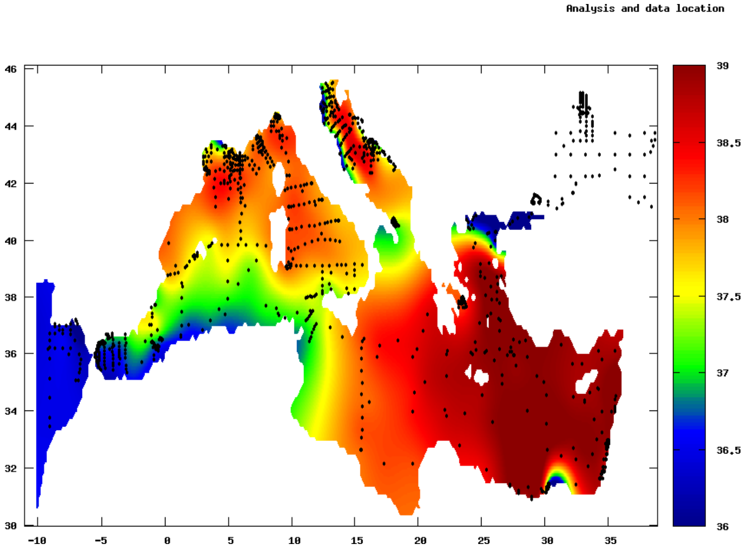


CHART SHOWING THE DISTRIBUTION OF SALTNESS IN THE OCEAN.
Arctic Circle
Tropic of Cancer
Equator
Tropic of Capricorn
Antarctic Circle

Solutions


Analysis and data location
46
44
42
40
38
36
34
32
30
-10
-5
0
5
10
15
20
25
30
35
39
38.5
38
37.5
37
36.5
36

Estimation

- Observer 1: 14°
- Observer 2: 16°

Your best guess ?

Estimation

- Observer 1: 14°
- Observer 2: 16°

Your best guess ?
15°

Estimation

- Observer 1: 14°
- Observer 2: 16°

Your best guess ?

15°

But what if observer 1 uses digital thermometer and observer 2 his finger ?

Estimation

- Observer 1: 14°
- Observer 2: 16°

Your best guess ?

15°

But what if observer 1 uses digital thermometer and observer 2 his finger ?

Best guess probably near 14°.

Estimation

- Observer 1: 14°
- Observer 2: 16°

Your best guess ?

15°

But what if observer 1 uses digital thermometer and observer 2 his finger ?

Best guess probably near 14°.

Exploit knowledge of errors !

Optimal estimate

$$T_1 = T^t + \epsilon_1, \quad \langle \epsilon_1 \rangle = 0, \quad T_2 = T^t + \epsilon_2, \quad \langle \epsilon_2 \rangle = 0 \tag{1}$$

statistical average, denoted by $\langle \quad \rangle$ with unbiased estimates $\langle \epsilon_* \rangle = 0$
Linear estimate

$$T = w_1 T_1 \ + \ w_2 T_2 \ = \ (w_1 + w_2) T^t \ + \ (w_1 \epsilon_1 + w_2 \epsilon_2) \tag{2}$$

$$\langle T \rangle = (w_1 + w_2) T^t, \tag{3}$$

we obtain an unbiased estimate of the true state if we take $w_1 + w_2 = 1$. This leaves one parameter free to chose: w_2

Exploit knowledge on errors to find optimal value of w_2

Choice of weighting ?

$$T^a = (1 - w_2)T_1 + w_2 T_2 \; = \; T_1 + w_2(T_2 - T_1) \tag{4}$$

while in reality there is an error

$$T^a - T^t = (1 - w_2)\epsilon_1 + w_2\epsilon_2, \tag{5}$$

This error is zero on average but its variance is not zero:

$$\left\langle (T^a - T^t)^2 \right\rangle = (1 - w_2)^2 \left\langle \epsilon_1^2 \right\rangle + w_2^2 \left\langle \epsilon_2^2 \right\rangle + 2(1 - w_2)w_2 \left\langle \epsilon_1 \epsilon_2 \right\rangle \tag{6}$$

The actual errors ϵ_1 and ϵ_2 are not known, but the error variance $\left\langle \epsilon_1^2 \right\rangle$ are. Often we can reasonably suppose that the errors ϵ_1 and ϵ_2 are uncorrelated $\left\langle \epsilon_1 \epsilon_2 \right\rangle = 0$. The error variance $\left\langle \epsilon^2 \right\rangle$ of the analysis is

$$\left\langle \epsilon^2 \right\rangle = (1 - w_2)^2 \left\langle \epsilon_1^2 \right\rangle + w_2^2 \left\langle \epsilon_2^2 \right\rangle. \tag{7}$$

So what ?

Minimisation

$$\left\langle \epsilon^2 \right\rangle = (1 - w_2)^2 \left\langle \epsilon_1^2 \right\rangle + w_2^2 \left\langle \epsilon_2^2 \right\rangle . \tag{8}$$

Naturally, the best estimate for T is the one with the lowest expected error variance and we will use w_2, which minimizes the right-hand side:

$$w_2 = \frac{\left\langle \epsilon_1^2 \right\rangle}{\left\langle \epsilon_1^2 \right\rangle + \left\langle \epsilon_2^2 \right\rangle} \tag{9}$$

$$T^a = \frac{\left\langle \epsilon_1^2 \right\rangle \left\langle \epsilon_2^2 \right\rangle}{\left\langle \epsilon_1^2 \right\rangle + \left\langle \epsilon_2^2 \right\rangle} \left(\frac{T_1}{\left\langle \epsilon_1^2 \right\rangle} + \frac{T_2}{\left\langle \epsilon_2^2 \right\rangle} \right) . \tag{10}$$

Best estimate

With (9) we obtain the minimal error variance

$$\left\langle \epsilon^2 \right\rangle = \frac{\left\langle \epsilon_1^2 \right\rangle \left\langle \epsilon_2^2 \right\rangle}{\left\langle \epsilon_1^2 \right\rangle + \left\langle \epsilon_2^2 \right\rangle} = \left(1 - \frac{\left\langle \epsilon_1^2 \right\rangle}{\left\langle \epsilon_1^2 \right\rangle + \left\langle \epsilon_2^2 \right\rangle} \right) \left\langle \epsilon_1^2 \right\rangle , \tag{11}$$

while the estimate of the temperature itself reads

$$T^a = T_1 + \left(\frac{\left\langle \epsilon_1^2 \right\rangle}{\left\langle \epsilon_1^2 \right\rangle + \left\langle \epsilon_2^2 \right\rangle} \right) (T_2 - T_1) . \tag{12}$$

Error variance on the combination of T_1 and T_2 is smaller than both $\left\langle \epsilon_1^2 \right\rangle$ and $\left\langle \epsilon_2^2 \right\rangle$.

Optimal Interpolation I

Same problem but data distributed in space and *a priori* information on background (with variance σ^2).
Weighting of background (zero value when working with anomalies and σ^2 local variance and covariances between points) information and data points (observed values and observational error variance)

- "Model forecast": Background field.
- Need for covariance of the background field between data points: each element i, j of $\mathbf{B}$ provides the covariance between points in location i and j. Covariance between a given point and all data points is stored in column vector $\mathbf{c}$ and the local variance at the analysis point is noted σ^2.
- Analysis ϕ of anomaly $\mathbf{y}$ with respect to background leads to spatial analysis at any desired location of covariance between any two points is known.

Optimal Interpolation II

$$\phi = \mathbf{c}^t (\mathbf{B} + \mathbf{R})^{-1} \mathbf{y} \tag{13}$$

with a local error variance of the analysis

$$\epsilon_a^2 = \sigma^2 - \mathbf{c}^t (\mathbf{B} + \mathbf{R})^{-1} \mathbf{c} \tag{14}$$

Note that inversion of matrix is needed (cost increases as the cube of number of data points).

Background covariance

Problem, how to specify background covariances (between all data points and between data points and the desired analysis location).

- c_i= covariance between location of the analysis and data location of point i = $C(x, x_i)$
- B_{ij}=covariance between location of data point i and location of point j= $C(x_i, x_j)$

Approaches

- Normally obtained via statistics on data. Seldom possible (noticable exception: satellite images).
- Standard OI: via functions $B_{ij} = f(r/L)$ where r is the distance between points i and j, but still function f needs to be determined. L is the so-called correlation length. Here statistics on all data couples as a function of distance. Example: $f = \sigma^2 \exp(-r^2/L^2)$.
- Via functionals (see Kernel of DIVA later)

Signal to noise ratio

$$\mathbf{B} = \sigma^2 \tilde{\mathbf{B}} \tag{15}$$

$$\mathbf{R} = \epsilon^2 \tilde{\mathbf{R}} \tag{16}$$

$$\mathbf{c} = \sigma^2 \tilde{\mathbf{c}} \tag{17}$$

with non-dimensional correlation matrixes $\tilde{\mathbf{B}}$, $\tilde{\mathbf{R}}$, $\tilde{\mathbf{c}}$

$$\phi = \tilde{\mathbf{c}}^t \left(\tilde{\mathbf{B}} + \frac{1}{\lambda} \tilde{\mathbf{R}} \right)^{-1} \mathbf{y} \tag{18}$$

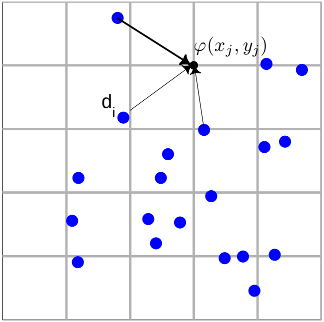
with signal-to noise ratio

$$\lambda = \frac{\sigma^2}{\epsilon^2} \tag{19}$$

Also the error field is only depending on the ratio.

DIVA: Data-Interpolating Variational Analysis

N_d data points d_i • $\rightarrow$ gridded field



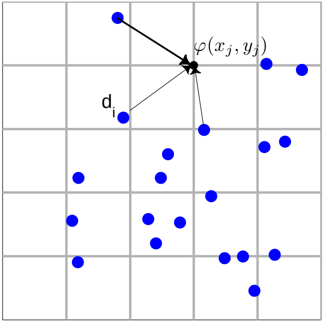
Formulation: minimize cost function $J[\varphi]$

$$\min J[\varphi] = \sum_{i=1}^{N} \mu_i \left[d_i - \varphi(x_i, y_i)\right]^2$$

$$+ \int_D \left(\nabla\nabla\varphi : \nabla\nabla\varphi + \alpha_1 \nabla\varphi \cdot \nabla\varphi + \alpha_0 \varphi^2\right) \mathrm{d}D$$

DIVA: Data-Interpolating Variational Analysis

N_d data points d_i • $\rightarrow$ gridded field



Formulation: minimize cost function $J[\varphi]$

$$\min J[\varphi] = \sum_{i=1}^{N} \mu_i \left[d_i - \varphi(x_i, y_i)\right]^2 \qquad \text{data–analysis misfit}$$

$$+ \int_D \left(\nabla\nabla\varphi : \nabla\nabla\varphi + \alpha_1 \nabla\varphi \cdot \nabla\varphi + \alpha_0 \varphi^2\right) \mathrm{d}D \qquad \text{field regularity}$$

Analysis parameters are related to data

Non-dimensional version:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{20}$$

$$\rightarrow \quad D = L^2\tilde{D} \tag{21}$$

Analysis parameters are related to data

Non-dimensional version:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{20}$$

$$\rightarrow \quad D = L^2\tilde{D} \tag{21}$$

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

Analysis parameters are related to data

Non-dimensional version:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{20}$$

$$\rightarrow \quad D = L^2 \tilde{D} \tag{21}$$

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

- $\alpha_0 \rightarrow L$ for which data-analysis misfit $\simeq$ regularity term: $\alpha_0 L^4 = 1$

Analysis parameters are related to data

Non-dimensional version:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{20}$$

$$\rightarrow \quad D = L^2 \tilde{D} \tag{21}$$

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

- $\alpha_0 \rightarrow L$ for which data-analysis misfit $\simeq$ regularity term: $\alpha_0 L^4 = 1$
- $\alpha_1 \rightarrow$ influence of gradients: $\alpha_1 L^2 = 2\xi, \quad \xi = 1$

Analysis parameters are related to data

Non-dimensional version:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{20}$$

$$\rightarrow \quad D = L^2 \tilde{D} \tag{21}$$

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

- $\alpha_0 \rightarrow L$ for which data-analysis misfit $\simeq$ regularity term: $\alpha_0 L^4 = 1$
- $\alpha_1 \rightarrow$ influence of gradients: $\alpha_1 L^2 = 2\xi, \quad \xi = 1$
- $\mu_i L^2 \rightarrow$ weight on data: $\mu_i L^2 = 4\pi \dfrac{\text{signal}}{\text{noise}_i}$

Analysis parameters are related to data

Non-dimensional version:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{20}$$

$$\rightarrow \quad D = L^2 \tilde{D} \tag{21}$$

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

Coefficients α_0, α_1 and μ_i related to

1. Correlation length L
2. Signal-to-noise λ
3. Observational noise standard deviation ϵ_i^2

Main analysis parameters

Correlation length L:

- Measure of the *influence* of data points
- Estimated by a least-square fit of the covariance function

Signal-to-noise ratio λ:

- Measure of the *confidence* in data
- Estimated with Generalized Cross Validation techniques

Correlation function

In variational analysis, the correlation function is not specified `a priori`, but its form is determined by the differential in equation (20). For an infinite domain, one can show that for $\alpha_0 = L^{-4}$ and $\alpha_1 = 2L^{-2}$ the correlation function is given by:

$$C(r) = \frac{r}{L} K_1 \left(\frac{r}{L} \right) \tag{22}$$

where r is the Euclidean distance, L is the correlation length and K_1 is the modified Bessel function.

Correlation function

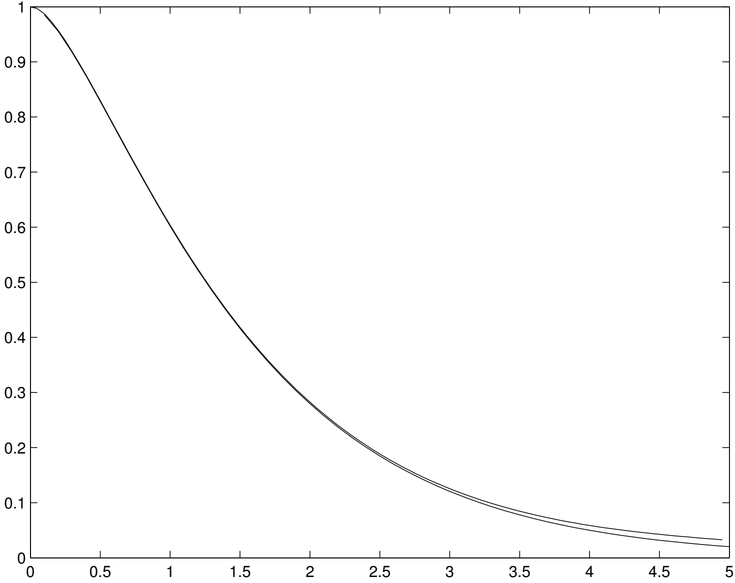
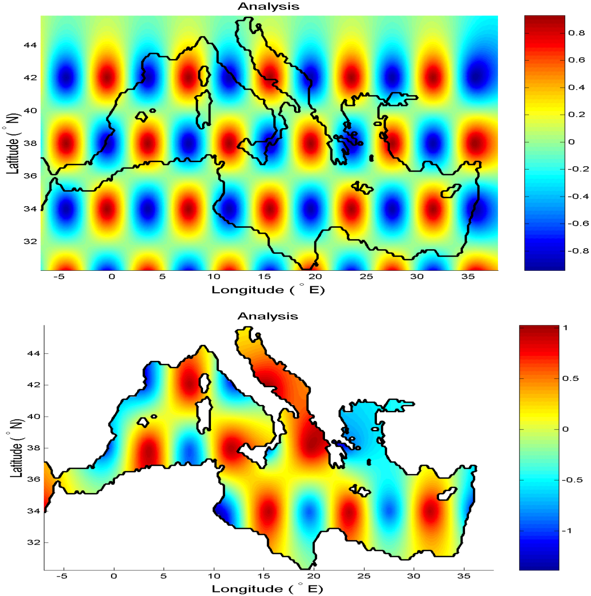
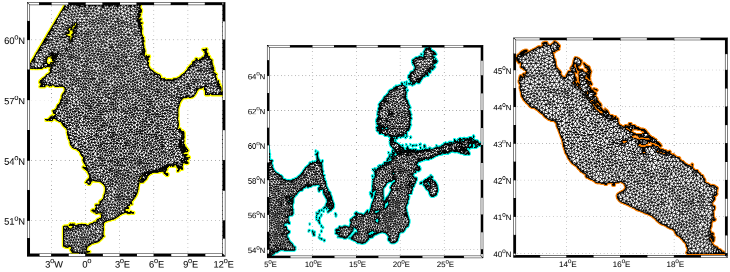


Illustration of covariance functions



Minimization with a finite-element method

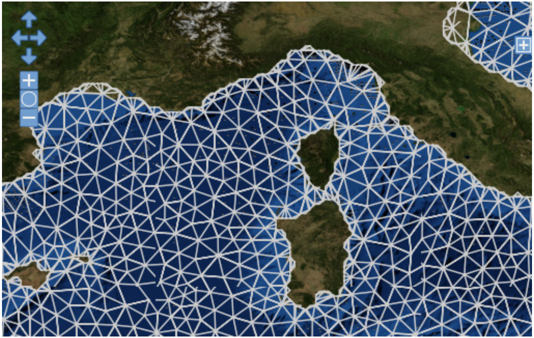
Field regularity $\rightarrow$ plate bending problem $\rightarrow$ finite-element solver



Advantages:

- boundaries taken into account
- numerical cost (almost independent on data number)
- no *a posteriori* masking (except if based on error level)

Minimization with a finite-element method



Solution by finite element method. Note decoupling of subbasins. (Each element is in fact composed by three sub-elements)

- The solution method relies on a finite-element resolution, the mesh being automatically generated using one of the tools provided within the software.
- The finite element mesh only covers actual sea so that land regions are excluded from the analysis and provide natural barriers for data-information propagation.

Minimization with a finite-element method

Triangular FE only covers sea: $$J[\varphi] = \sum_{e=1}^{N_e} J_e(\varphi_e) \tag{23}$$

In each element: $$\varphi_e(\mathbf{r_e}) = \mathbf{q_e}^T \mathbf{s}(\mathbf{r_e}) \quad \text{with} \begin{cases} \mathbf{s} & \rightarrow \text{shape functions} \\ \mathbf{q} & \rightarrow \text{connectors} \\ \mathbf{r_e} & \rightarrow \text{position} \end{cases} \tag{24}$$

(24) in (23) + variational principle

$$J_e(\mathbf{q_e}) = \mathbf{q_e}^T \mathbf{K_e} \mathbf{q_e} - 2\mathbf{q_e}^T \mathbf{g_e} + \sum_{i=1}^{N_{d_e}} \mu_i d_i \tag{25}$$

$$\text{where} \begin{cases} \mathbf{K_e} & \rightarrow \text{local stiffness matrix} \\ \mathbf{g} & \rightarrow \text{vector depending on local data} \end{cases}$$

Minimization with a finite-element method

On the whole domain:
$$J(\mathbf{q}) = \mathbf{q}^T\mathbf{K}\mathbf{q} - 2\mathbf{q}^T\mathbf{g} + \sum_{i=1}^{N_d} \mu_i d_i \tag{23}$$

Minimum:
$$\mathbf{q} = \mathbf{K}^{-1}\mathbf{g} \tag{24}$$

$$\boxed{\mathbf{q}} = \boxed{\mathbf{K}^{-1}}\,\boxed{\mathbf{g}} \tag{25}$$

- Stiffness matrix
- Connectors (new unknowns)
- Charge vector

Mapping of data on FEM $\rightarrow$ transfer operator $\mathbf{T_2}$ $\rightarrow$ $\mathbf{g} = \mathbf{T_2}(\mathbf{r})\mathbf{d}$

Solution at any location $\rightarrow$ transfer operator $\mathbf{T_1}$ $\rightarrow$ $\varphi(\mathbf{r}) = \mathbf{T_1}(\mathbf{r})\mathbf{q}$

Results obtained at any location $\rightarrow \varphi = \mathbf{T_1}(\mathbf{r})\mathbf{K}^{-1}\mathbf{T_2}(\mathbf{r})\mathbf{d}$

DIVA as OI

DIVA is identical to the well known Optimal Interpolation

- if so-called reproducing kernel of the norm = covariance function of OI,
- if the noise is random, spatially uncorrelated and the signal/noise ratio parameter is identical with OI.

In this case, the OI solution = DIVA solution.

- Advantages of DIVA: regularization, fast finite-element solution, boundary effects taken into account.
- Difficulties: generalizations to 3D and multivariate versions are "hybrid" (some expensive O.I. components).

Major direct advantage of DIVA: matrix to invert is related to the finite-element mesh, NOT the number of data. Useful for large data sets (Rixen *et al.*, 2000). Equivalence allows to calculate error fields with DIVA even if formulation does not rely on error minimisation.

Comparison

Method	$\min(\epsilon^2)$	3D	Multivar	Ops/image	$\epsilon(\boldsymbol{r})$	a priori	C.V.	anisotropy
Cressman		$\star$	$\star$	$N_d N_a$		$w(r/L)$	(L)	$(\star)$
O.I.	$\star$	$\star$	$\star$	$N_d^3 + N_d N_a$	$\star$	$c(r/L)$	$L, \sigma^2/\mu^2$	$(\star)$
DIVA	$\star$	$(\star)$	$(\star)$	$N_e^{5/2}$	$\star$	$K(r/L)$	$L, \sigma^2/\mu^2$	$\star$
DINEOF	$(\star)$	$\star$	$\star$	$N_a^{5/4}$	$(\star)$	stat.	N	$\star$

N_d : number of data points

N_a : number of grid points for analysis

N_e : number of finite elements

N : number of EOFs

L : correlation length

σ^2/ϵ^2 : signal to noise ratio

$\star$: available feature

$(\star)$: available with some adaptations

Diva Cocktail Recipe

Ingredients:

- 1 1/2 oz vodka
- 1/2 oz passion-fruit juice
- 1/2 oz lime juice
- 1 tbsp cherry juice
- fill with soda

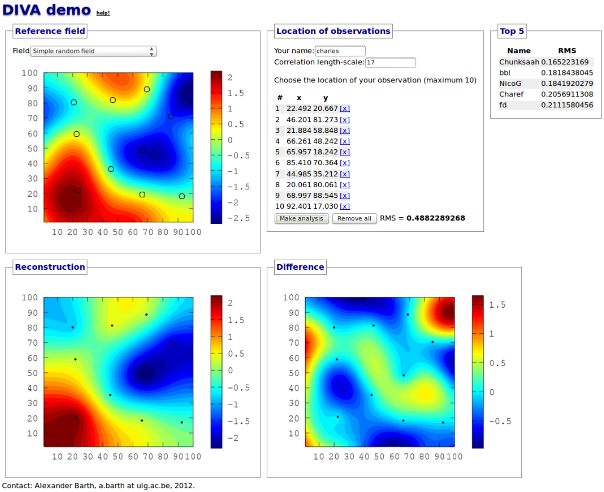
Diva Cocktail Recipe

Ingredients:

- Smoothness
- Observation constraint
- Behaviour constraint

Want to use Diva?

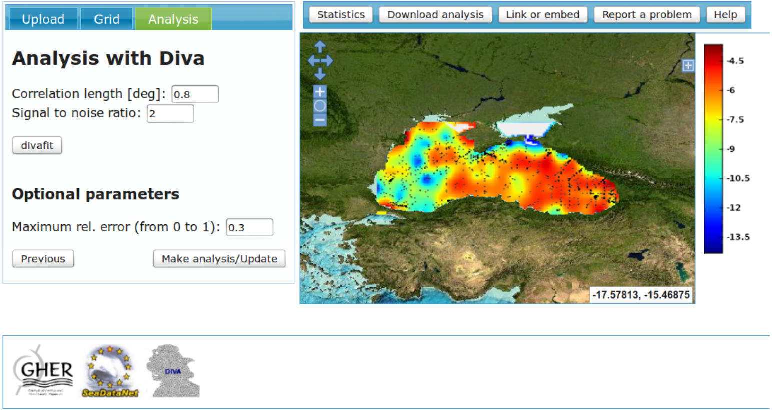
Playing…



http://data-assimilation.net/Tools/divand_demo/html/

Want to use Diva?

With your own data…

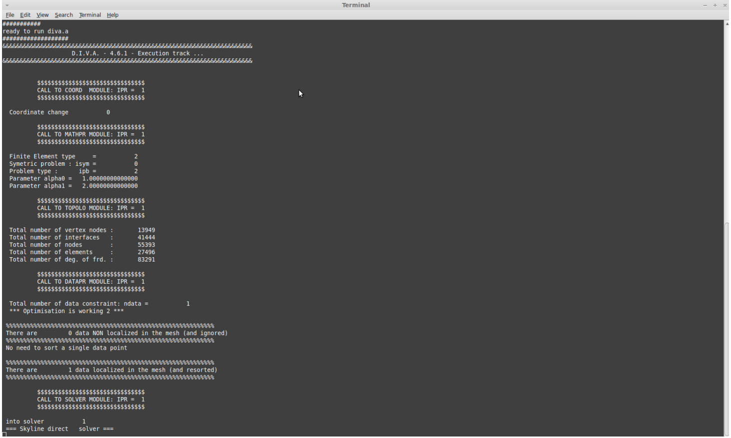


http://gher-diva.phys.ulg.ac.be/web-vis/diva.html or ODV or matlab wrapper

Want to use Diva?

For serious work:

- 2D version (for production), open source, GPL
- nD version (for research), open source, GPL



```
###########
ready to run diva.a
###################
&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&
                    D.I.V.A. - 4.6.1 - Execution track ...
&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&


          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
          CALL TO COORD  MODULE: IPR =  1
          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

  Coordinate change          0

          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
          CALL TO MATHPR MODULE: IPR =  1
          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

  Finite Element type     =            2
  Symetric problem : isym =            0
  Problem type :      ipb =            2
  Parameter alpha0 =   1.00000000000000
  Parameter alpha1 =   2.00000000000000

          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
          CALL TO TOPOLO MODULE: IPR =  1
          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

  Total number of vertex nodes :       13949
  Total number of interfaces   :       41444
  Total number of nodes        :       55393
  Total number of elements     :       27496
  Total number of deg. of frd. :       83291

          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
          CALL TO DATAPR MODULE: IPR =  1
          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

  Total number of data constraint: ndata =           1
  *** Optimisation is working 2 ***

 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 There are          0 data NON localized in the mesh (and ignored)
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 No need to sort a single data point

 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 There are          1 data localized in the mesh (and resorted)
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
          CALL TO SOLVER MODULE: IPR =  1
          $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

 into solver           1
 === Skyline direct   solver ===
```

http://modb.oce.ulg.ac.be/mediawiki/index.php/DIVA

Running Diva in 2D: input files

1. `data.dat`: contains the observations — x|y|value

```
36.5500 45.163 17.7138
33.7500 44.167 18.135
32.7500 44.167 18.51
36.2500 43.833 18.5892
33.2500 45.083 18.2326
32.7833 43.917 18.477
32.7500 43.500 18.59
37.2433 44.833 18.1555
36.5000 44.000 18.19
35.8333 43.750 18.62
34.2500 43.832 18.29
35.6500 44.000 18.75
38.0000 44.000 18.155
37.8200 44.368 17.1916
39.0000 42.500 18.23
33.1333 44.433 18.001
33.0500 44.433 18.09
33.2500 44.167 18.231
32.5333 44.833 18.014
38.0167 44.447 18.0568
```

Running Diva in 2D: input files

1. `data.dat`: contains the observations — x|y|value
2. `coast.cont`: delimits land and sea — (coastline or isobaths)

```
         7
       552
  27.4375000       40.3499985
  27.4500008       40.3375015
  27.4666672       40.3375015
  27.4833336       40.3375015
  27.5000000       40.3375015
  27.5166664       40.3375015
  27.5333328       40.3375015
  27.5499992       40.3375015
  27.5666676       40.3375015
  27.5791664       40.3499985
  27.5833340       40.3541679
  27.6000004       40.3541679
  27.6124992       40.3666649
  27.6166668       40.3708344
  27.6291676       40.3833351
  27.6291676       40.4000015
  27.6291676       40.4166679
  27.6291676       40.4333344
  27.6166668       40.4458351
  27.6124992       40.4500008
```

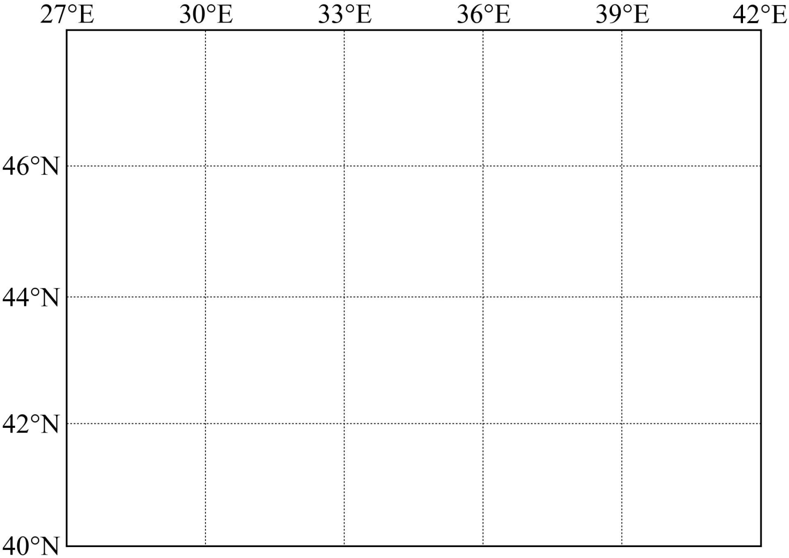
Running Diva in 2D: input files

1. `data.dat`: contains the observations — x|y|value
2. `coast.cont`: delimits land and sea — (coastline or isobaths)
3. `param.par`: analysis parameters — L, λ, resolution, . . .

```
# Lc: correlation length (in units coherent with your data)#
1.5
# icoordchange (=0 if no change of coordinates is to be performed; =1 if positions are in degree
2
# ispec: output files required#
0
# ireg: mode selected for background field: 0=null guess; 1=mean of data; 2=regression plan if a
2
# xori: x-coordinate of the first grid point of the output#
27
# xori: y-coordinate of the first grid point of the output#
40
# dx: step of output grid#
0.1
# dy: step of output grid#
0.1
# nx: number of grid points in the x-direction#
151
# ny: number of grid points in the y-direction#
76
# valex: exclusion value#
-99
# snr: signal to noise ratio of the whole dataset#
0.5
# varbak: variance of the background field. If zero, no error fields are produced. If one, relat
1.0
```

Workflow in 2D

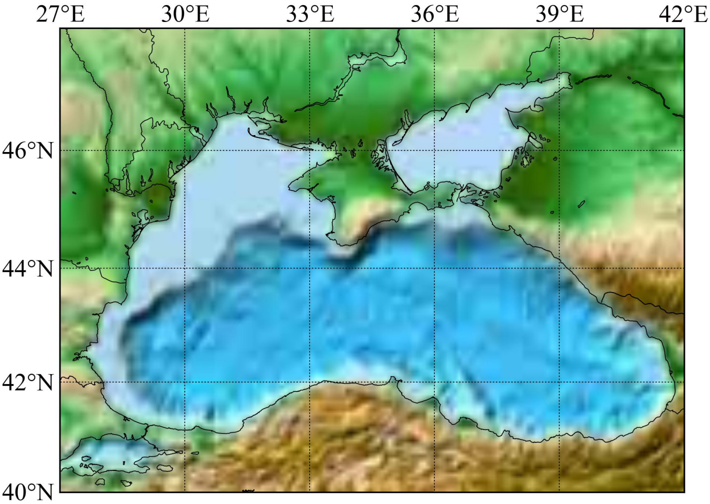
Select region of study



Workflow in 2D

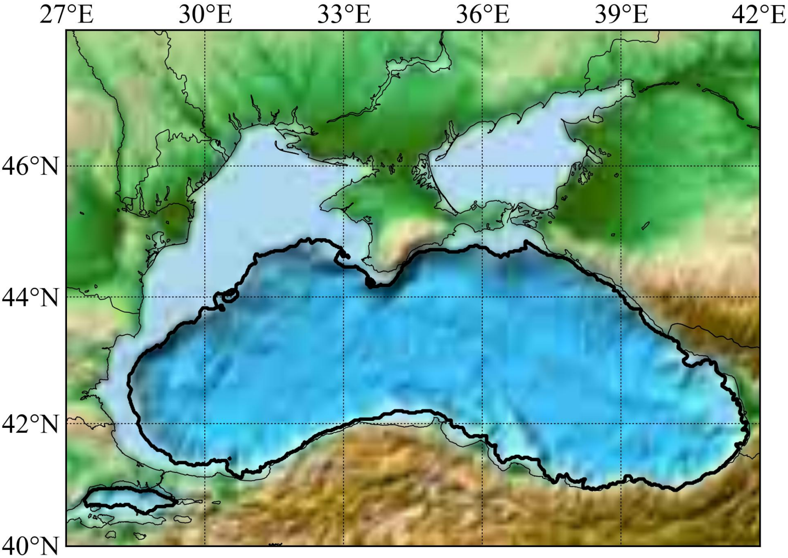
Extract topography, for example via

http://gher-diva.phys.ulg.ac.be/web-vis/diva.html



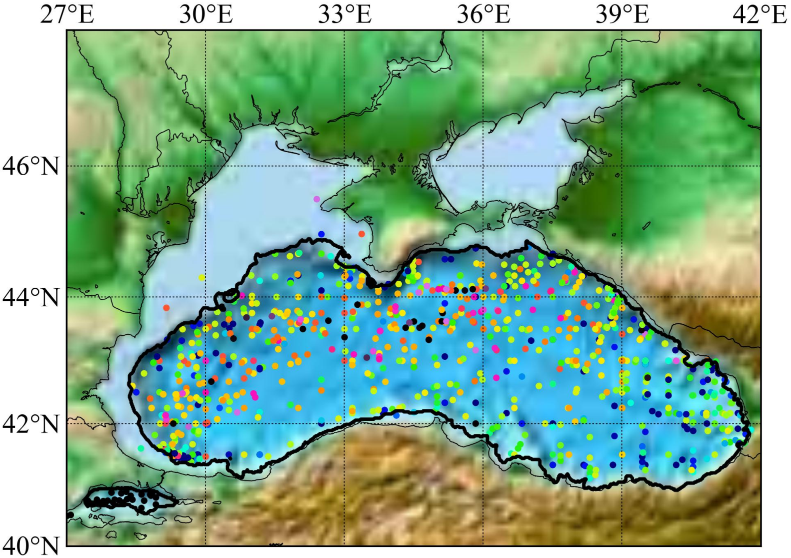
Workflow in 2D

Generate contour



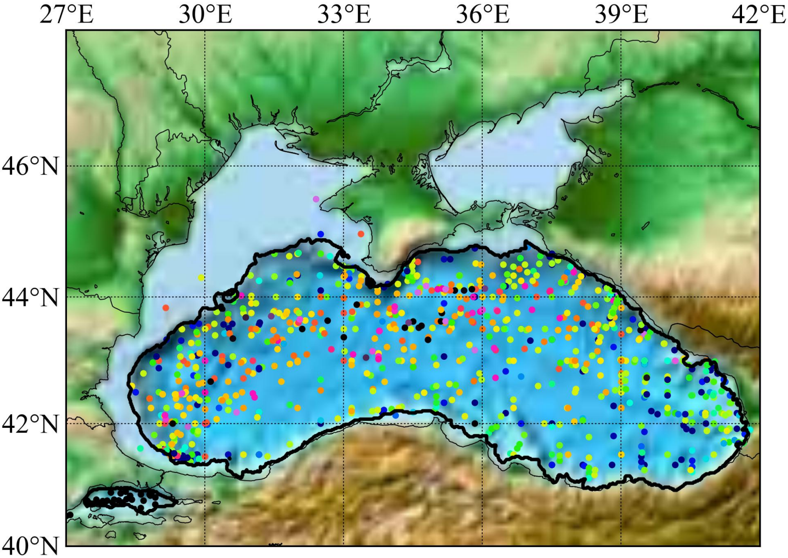
Workflow in 2D

Extract data



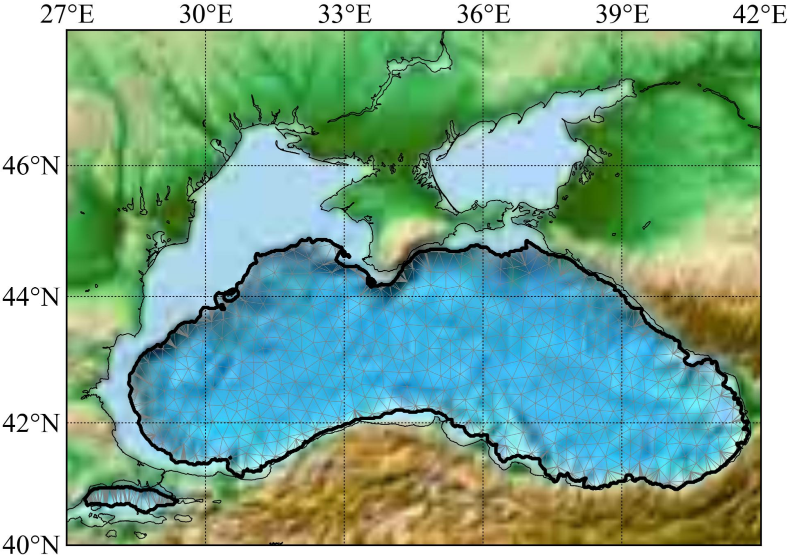
Workflow in 2D

Evaluate analysis parameters



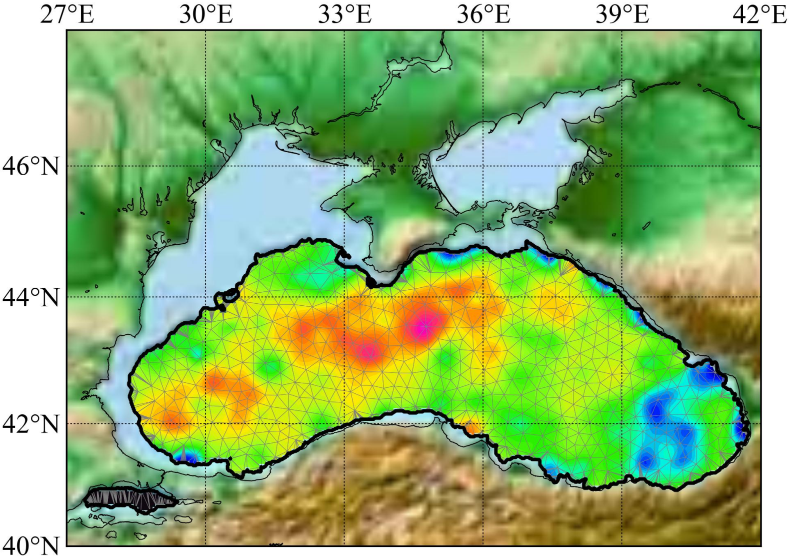
Workflow in 2D

Create finite-elementmesh



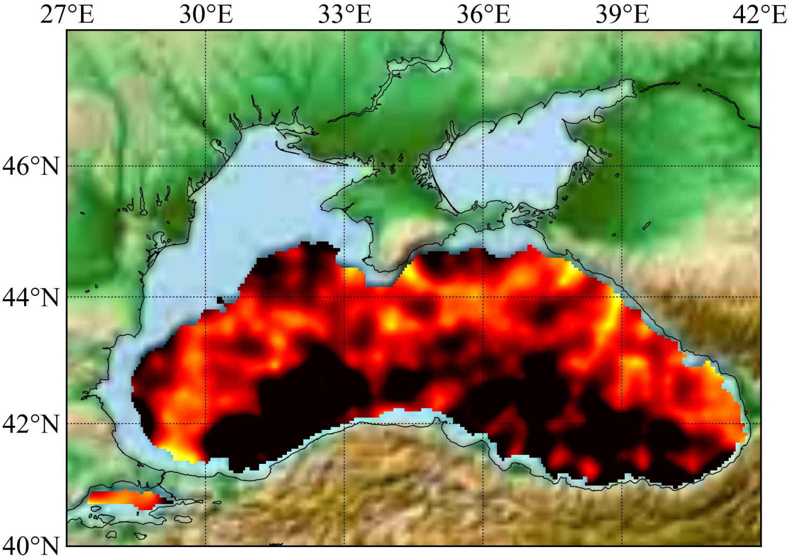
Workflow in 2D

Generate analysis



Workflow in 2D

Generate error field



When to use 2D version

- occasional use
- 2D fields like benthic properties
- for implementation of special features by your own (eg multiplicative bias correction, special background field creation based on habitats
- ...

otherwise: use 3D or 4D version directly

Next. . .

Diva in 4 dimensions