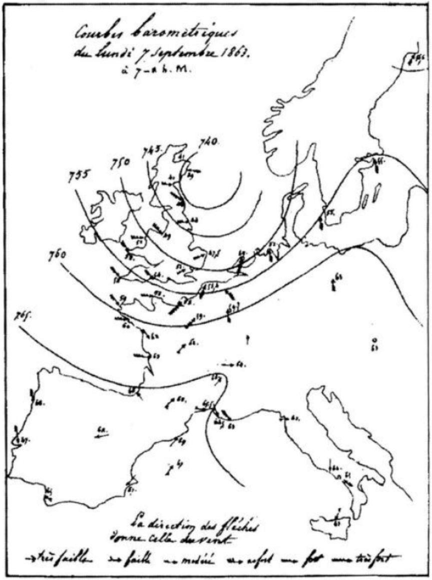
Diva workshop 2015

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

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

On-going:

- Modernisation of the code structure
- n-dimensional generalisation
- Spatially correlated observations errors
- optimized and approximate error calculations (clever poor man)
- Analysis at a specific distance from bottom
- ... ☕

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

On-going:

General: $\boxed{\text{user-driven developments}}$

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

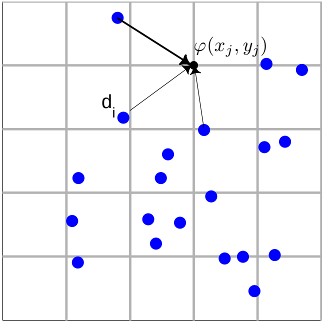
OceanBrowser: visualisation tool of 4D NetCDF files

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

divaformatlab: wrapper to use in matlab

DIVA: Data-Interpolating Variational Analysis

N_d data points d_i • $\rightarrow$ gridded field



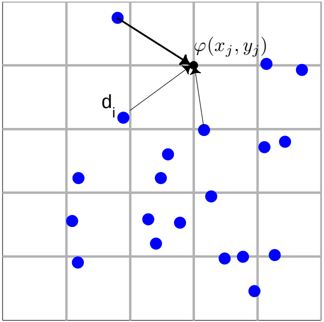
Formulation: minimize cost function $J[\varphi]$

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

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

DIVA: Data-Interpolating Variational Analysis

N_d data points d_i • $\rightarrow$ gridded field



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

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

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

Analysis parameters are related to data

Non-dimensional version:

$$
\begin{aligned}
L = \text{length scale} \quad &\rightarrow \quad \tilde{\boldsymbol{\nabla}} = L \boldsymbol{\nabla} && (1) \\
&\rightarrow \quad D = L^2 \tilde{D} && (2)
\end{aligned}
$$

Analysis parameters are related to data

Non-dimensional version:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{1}$$

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

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

Analysis parameters are related to data

Non-dimensional version:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{1}$$

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

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

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

Analysis parameters are related to data

Non-dimensional version:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{1}$$

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

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

- $\alpha_0 \rightarrow L$ 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:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{1}$$

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

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

- $\alpha_0 \rightarrow L$ 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:

$$L = \text{length scale} \quad \rightarrow \quad \tilde{\boldsymbol{\nabla}} = L\boldsymbol{\nabla} \tag{1}$$

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

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

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:

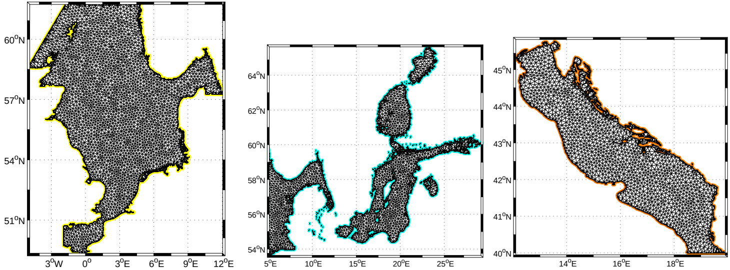
- Measure of the *influence* of data points
- Estimated by a least-square fit of the covariance function

Signal-to-noise ratio λ:

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

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

$$\text{Triangular FE only covers sea:} \qquad J[\varphi] = \sum_{e=1}^{N_e} J_e(\varphi_e) \tag{3}$$

$$\text{In each element: } \varphi_e(\mathbf{r_e}) = \mathbf{q_e}^T \mathbf{s}(\mathbf{r_e}) \quad \text{with} \begin{cases} \mathbf{s} & \rightarrow \text{shape functions} \\ \mathbf{q} & \rightarrow \text{connectors} \\ \mathbf{r_e} & \rightarrow \text{position} \end{cases} \tag{4}$$

(4) in (3) + variational principle

$$J_e(\mathbf{q_e}) = \mathbf{q_e}^T \mathbf{K_e} \mathbf{q_e} - 2\mathbf{q_e}^T \mathbf{g_e} + \sum_{i=1}^{N_{d_e}} \mu_i d_i \tag{5}$$

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

Minimization with a finite-element method

On the whole domain: $$J(\mathbf{q}) = \mathbf{q}^T\mathbf{K}\mathbf{q} - 2\mathbf{q}^T\mathbf{g} + \sum_{i=1}^{N_d} \mu_i d_i \tag{3}$$

Minimum: $$\mathbf{q} = \mathbf{K}^{-1}\mathbf{g} \tag{4}$$

$$\mathbf{q} = \mathbf{K^{-1}}\,\mathbf{g} \tag{5}$$

- 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 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

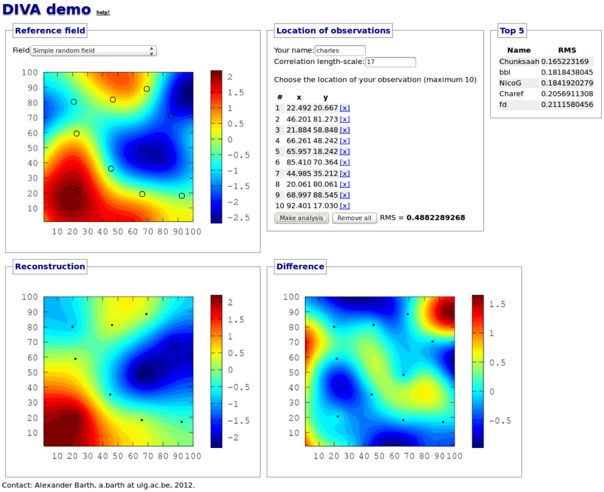
Diva Cocktail Recipe

Ingredients:

- Smoothness
- Observation constraint
- Behaviour constraint

Want to use Diva?

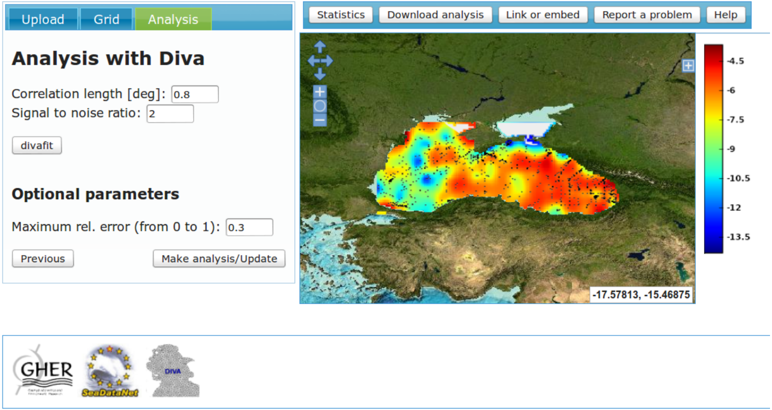
Playing. . .



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

Want to use Diva?

With your own data. . .



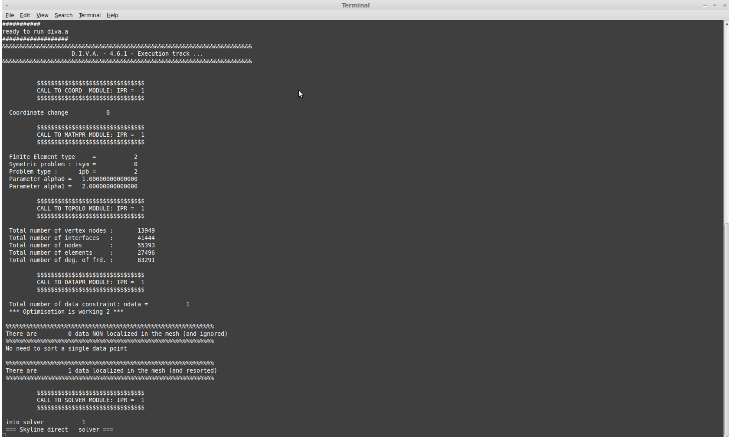
http://gher-diva.phys.ulg.ac.be/web-vis/diva.html or ODV or matlab wrapper

Want to use Diva?

For serious work:

2D version (for production), open source, GPL

nD version (for research), open source, GPL



```
Terminal
File  Edit  View  Search  Terminal  Help
###########
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
```

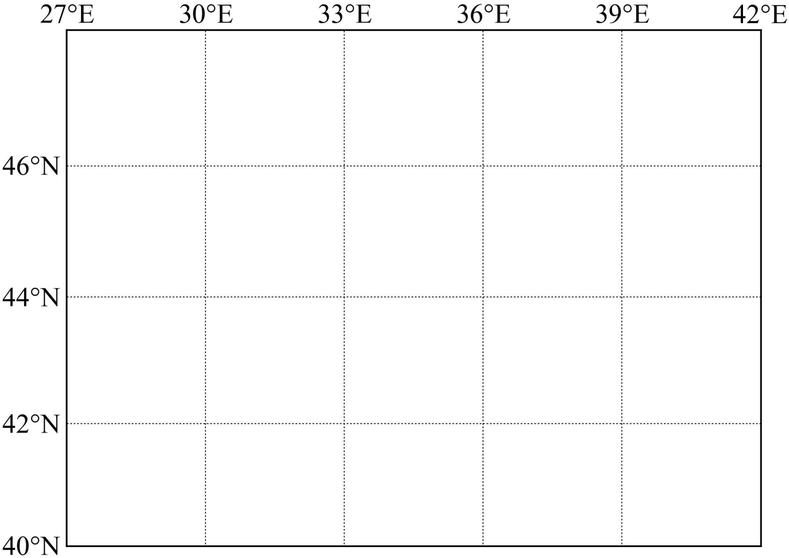
Running Diva in 2D: input files

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

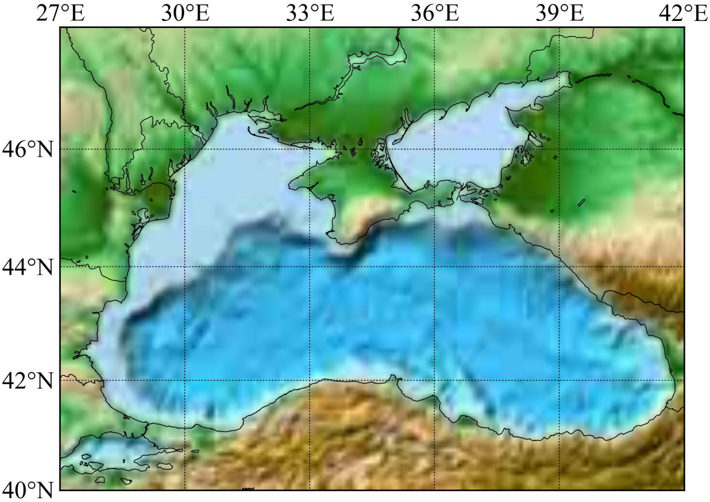
Select region of study



Workflow in 2D

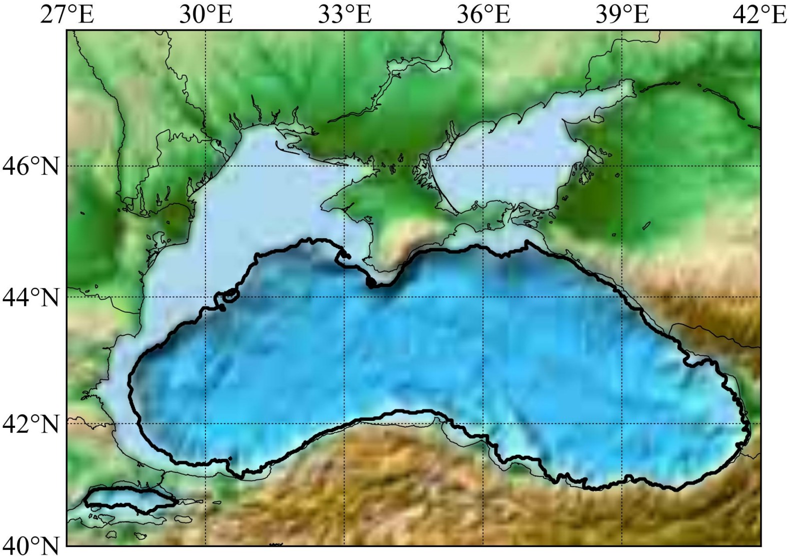
Extract topography, for example via

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



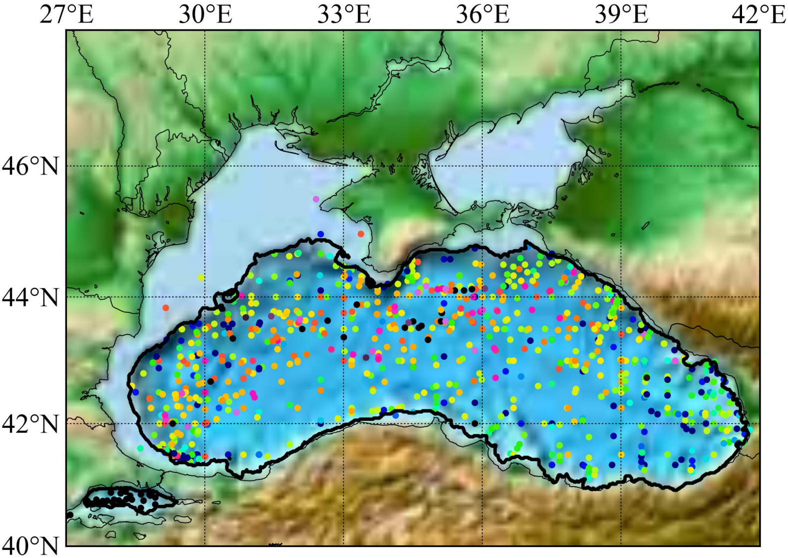
Workflow in 2D

Generate contour



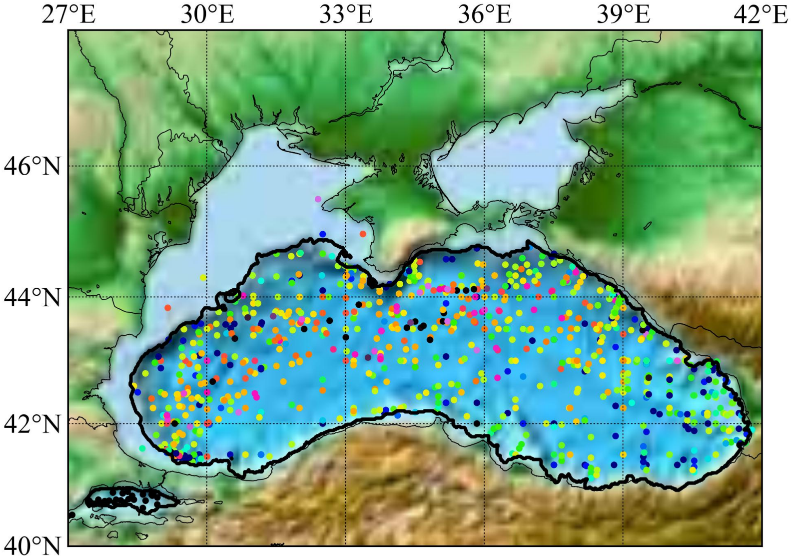
Workflow in 2D

Extract data



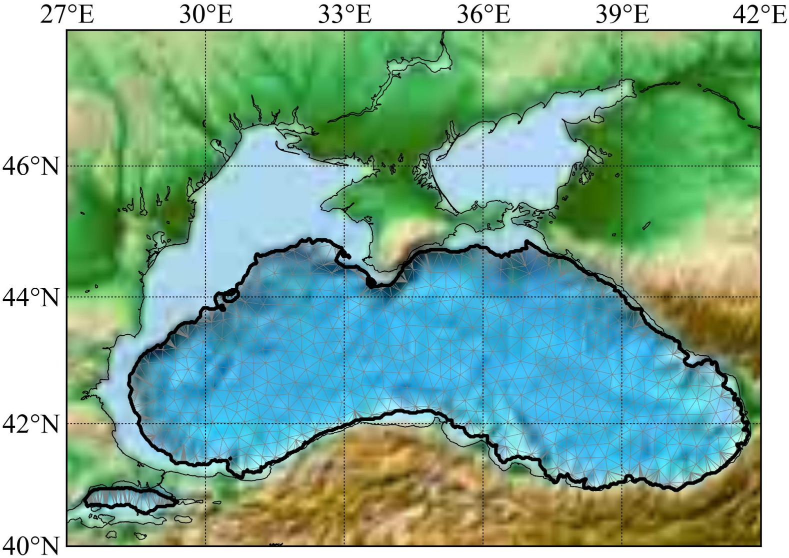
Workflow in 2D

Evaluate analysis parameters



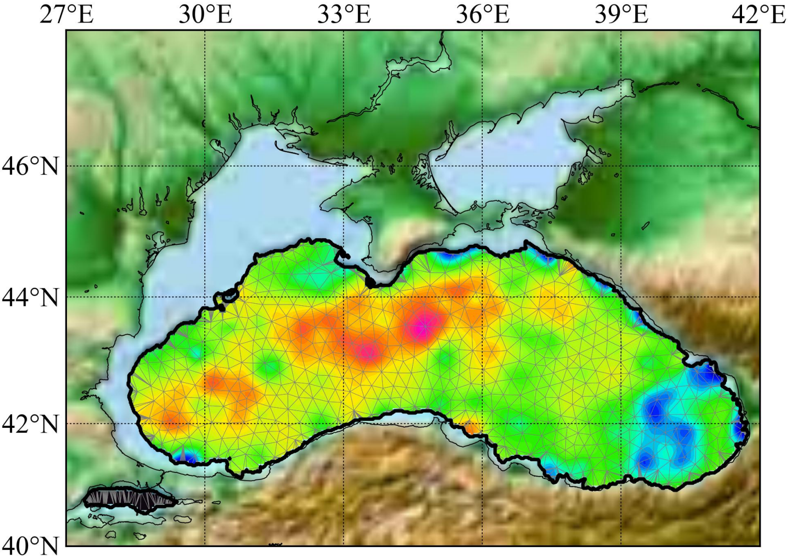
Workflow in 2D

Create finite-elementmesh



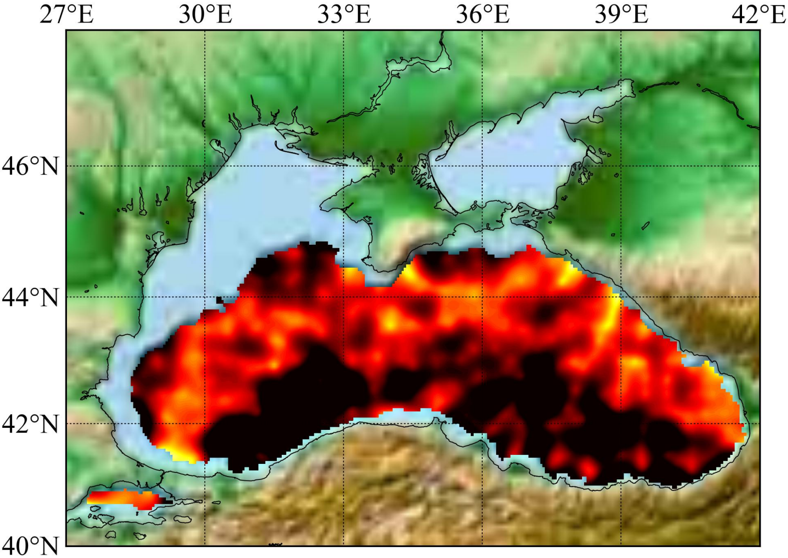
Workflow in 2D

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