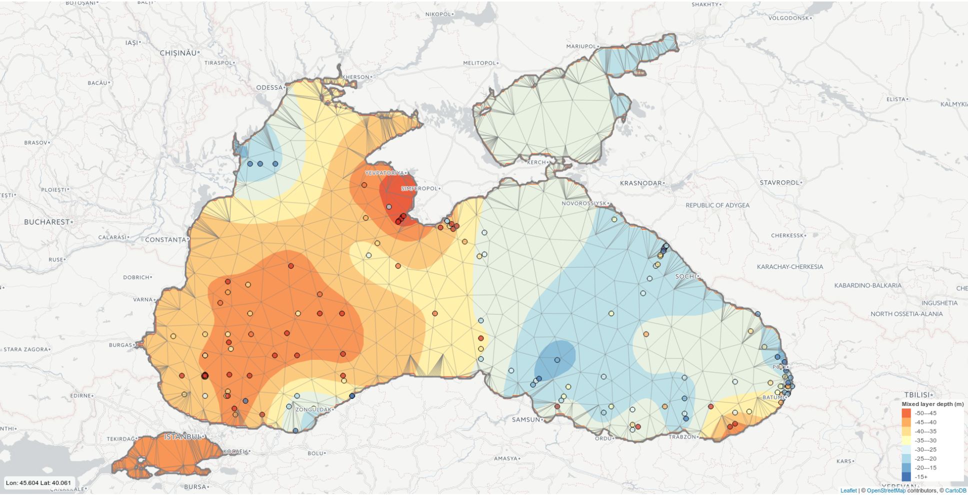
A. Barth, C. Troupin, S. Watelet &
J.-M. Beckers

GHER-University of Liège (ULiège)

New Diva capabilities for climatologies

SeaDataCloud – Product Meeting

Diva: from in situ data to gridded fields



https://github.com/gher-ulg/DIVA

DivaND: generalised, n-dimensional interpolation

https://www.geosci-model-dev.net/7/225/2014/gmd-7-225-2014.pdf

divand-1.0: *n*-dimensional variational data analysis for ocean observations

A. Barth[1,*], J.-M. Beckers[1], C. Troupin[2], A. Alvera-Azcárate[1], and L. Vandenbulcke[3,4]

[1]GHER, University of Liège, Liège, Belgium
[2]IMEDEA, Esporles, Illes Balears, Spain
[3]seamod.ro/Jailoo srl, Sat Valeni, Com. Salatrucu, Jud. Arges, Romania
[4]CIIMAR, University of Porto, Porto, Portugal

[*] *Invited contribution by A. Barth, recipient of the EGU Arne Richter Award for Outstanding Young Scientists 2010.*

Correspondence to: A. Barth (a.barth@ulg.ac.be)

Received: 7 June 2013 – Published in Geosci. Model Dev. Discuss.: 23 July 2013
Revised: 18 October 2013 – Accepted: 12 December 2013 – Published: 29 January 2014

https://github.com/gher-ulg/divand.jl

2013: Octave/MATLAB

2016: Julia

faster, better, stronger

DivaND: generalised, n-dimensional interpolation

- Variational inverse method
- Smoothness and other constraints

Differences with Diva (2D)

- *n*-dimensional, $n \geq 2$
- Different formulations, kernels & solvers
- Programming languages Fortran vs. Julia

User interfaces:

Jupyter notebooks and WPS

Notebooks: interactive computational environments

Notebooks combine:

1. code fragments that can be executed,
2. text for the description of the application and
3. figures illustrating the data or the results.

In [2]:
```
import numpy as np
import matplotlib.pyplot as plt
```

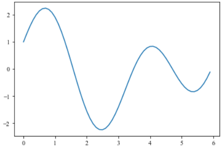
Data

Let's create a simple function.

In [6]:
```
x = np.arange(0, 6, .1)
y = np.cos(x) + 1.5 * np.sin(2 * x)
```

Make a simple plot

In [7]:
```
plt.plot(x, y)
plt.show()
```



Notebooks: interactive computational environments

Notebooks combine:

1. code fragments that can be executed,
2. text for the description of the application and
3. figures illustrating the data or the results.

"Digital Playground"

"Data Story Telling"

"Computational Narratives"

Notebooks: interactive computational environments

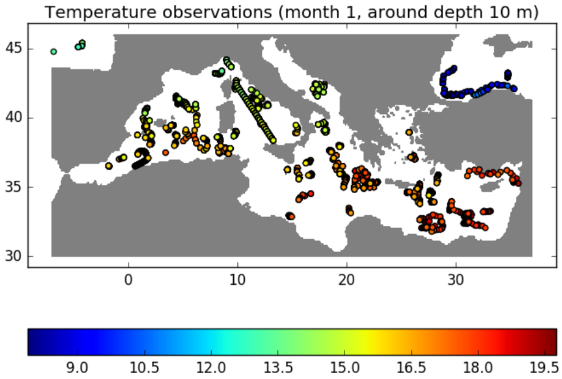
Notebooks combine:

1. code fragments that can be executed,
2. text for the description of the application and
3. figures illustrating the data or the results.

"Interactive notebooks: Sharing the code", Nature (2014)
http://www.nature.com/news/interactive-notebooks-sharing-the-code-1.16261

divand in a notebook

In [36]:
```
# sets the correct aspect ratio
gca()[:set_aspect](1/cos(mean(latr) * pi/180))
```



Setup the domain using the bathymetry from the file `bathname`.

Example online

New climatologies and products

Specifications for the new products

Depth levels:

- Common to all the products — allows combined product
- Follow IODE levels — 33 levels from 0 to 5500 m
- Consider World Ocean Atlas — more than 100 levels!

Specifications for the new products

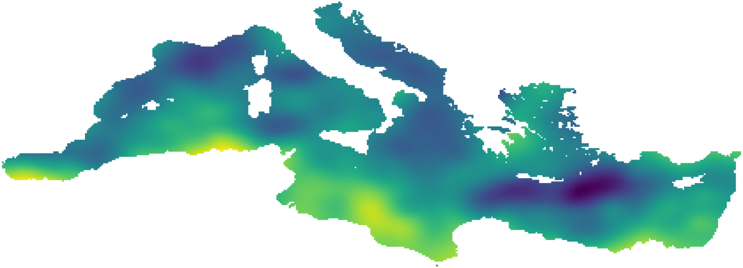
Spatial resolution:

Grid resolution $\neq$ real resolution!!

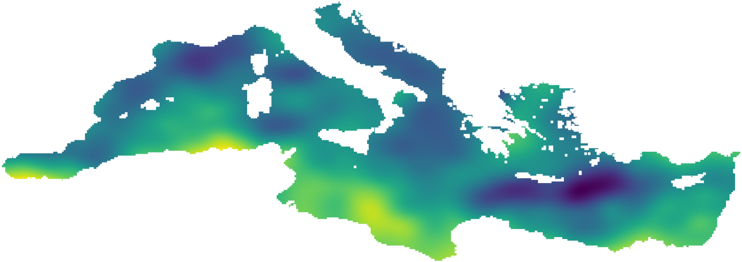
1. divand: $1/8° \times 1/8°$ based on data availability
2. post-processing: $1/24° \times 1/24°$ match model resolution

Find the differences between Product 1 & Product 2

Product 1



Product 2



Find the differences between Product 1 & Product 2

1. Field 1 is 161×426
   Field 2 is 641×1701

Find the differences between Product 1 & Product 2

1. Field 1 is 161×426
   Field 2 is 641×1701
2. File 2 is 30 times larger

Find the differences between Product 1 & Product 2

1. Field 1 is 161×426
   Field 2 is 641×1701
2. File 2 is 30 times larger
3. Product 2 would take wayyyyyyyyyyyyyyyyy longer to be created with divand

Re-gridding/re-interpolation: use nco?

nco	=	netCDF Operators
	=	set of standalone programs to manipulate netCDF files
	→	renaming, averaging, **re-gridding**, binary operations...

http://nco.sourceforge.net/

Re-gridding based on Earth System Modeling Framework (ESMF, https://www.earthsystemcog.org/projects/esmf/)

Re-gridding/re-interpolation: use `nco`?

Usage:

```
ncremap -i data.nc -d grid.nc -o output.nc
```

where:

data.nc = original netCDF containing field

grid.nc = file containing the new (finer) grid

output.nc = new netCDF with field interpolated onto the new grid

Further ideas on model-climatology comparison

1. Decrease the model resolution ($1/24°$)
   to match that of the climatology ($1/8°$)

Further ideas on model-climatology comparison

1. Decrease the model resolution (1/24°)
   to match that of the climatology (1/8°)
2. Detect where there is a high density of data
   and perform an analysis over that area with a higher resolution

Further ideas on model-climatology comparison

1. Decrease the model resolution (1/24°)
   to match that of the climatology (1/8°)
2. Detect where there is a high density of data
   and perform an analysis over that area with a higher resolution
3. Increase resolution locally
   where the geometry forces it (strait, islands, ...)

Further ideas on model-climatology comparison

1. Decrease the model resolution ($1/24°$)
   to match that of the climatology ($1/8°$)
2. Detect where there is a high density of data
   and perform an analysis over that area with a higher resolution
3. Increase resolution locally
   where the geometry forces it (strait, islands, ...)
4. Re-interpolate the climatologies at $1/24°$ and use it as a background
   Perform analysis with divand by window (larger than correlation length and shorter than domain size)

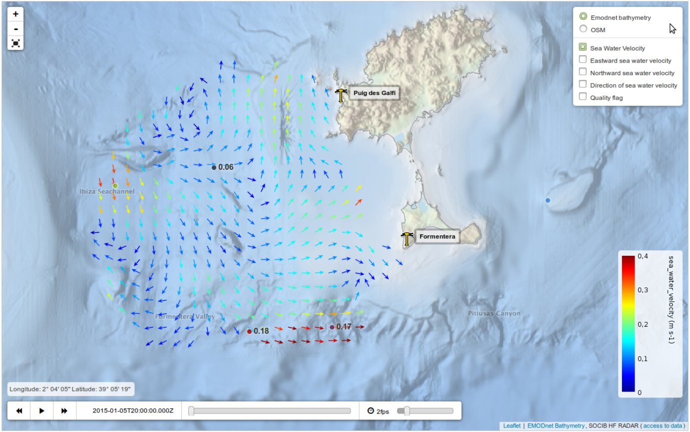
Further ideas on model-climatology comparison

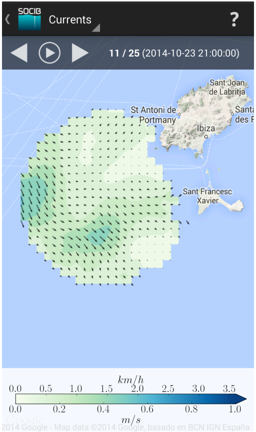
1. Decrease the model resolution ($1/24°$)
   to match that of the climatology ($1/8°$)
2. Detect where there is a high density of data
   and perform an analysis over that area with a higher resolution
3. Increase resolution locally
   where the geometry forces it (strait, islands, ...)
4. Re-interpolate the climatologies at $1/24°$ and use it as a background
   Perform analysis with divand by window (larger than correlation length and shorter than domain size)
5. Assign lower weight for high-resolution data

(Very) new product

Velocity field from HF radar

H2020 SeaDataCloud call: emphasis on coastal data



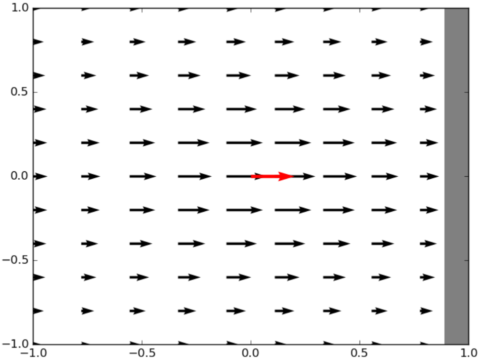


Data: SOCIB HF radar in the Ibiza Channel

New product: currents

$\rightarrow$ hypothetical measurement

$\rightarrow$ analyzed field



- Analysis of radial currents to derive total currents
- Observation operator links the radial currents of the different radar sites

Formulation: couple velocity components

$$\text{Norm :} \quad |\varphi|^2 = \int_\Omega (\alpha_2 \nabla\nabla\varphi : \nabla\nabla\varphi + \alpha_1 \nabla\varphi \cdot \nabla\varphi + \alpha_0 \varphi^2) \; d\Omega$$

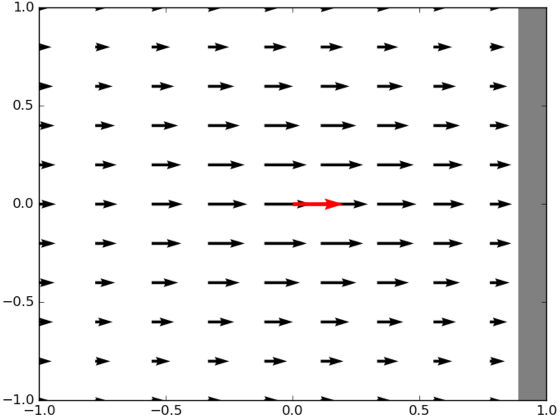
$$\text{Cost function:} \quad J(\vec{u}) = |u|^2 + |v|^2 + \sum_{\imath=1}^{N} \frac{(\vec{u}_\imath \cdot \vec{p}_\imath - u_{r\imath})^2}{\epsilon_\imath^2}$$

$\vec{u} = (u, v)$
$\vec{p}_\imath =$ normalized vector pointing toward the correspond HF radar site of the $\imath$-th radial observation $u_{r\imath}$

Coastline as a boundary condition ($\vec{u} \cdot \vec{n} = 0$)

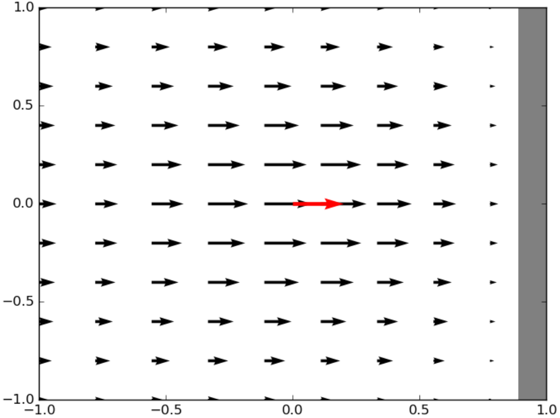
Cost function (OFF)

$$J_{bc}(\vec{u}) = \frac{1}{\epsilon_{bc}^2} \int_{\partial\Omega} (\vec{u} \cdot \vec{n})^2 ds$$



Cost function (ON)

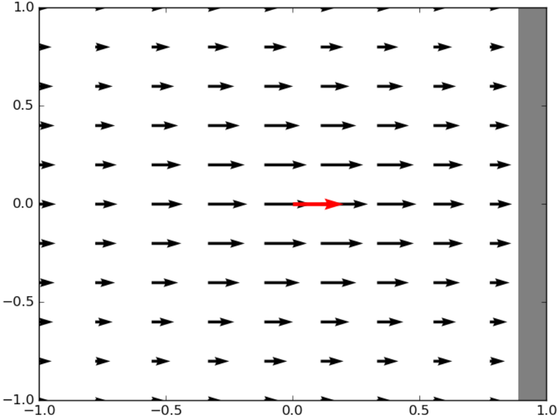
$$J_{bc}(\vec{u}) = \frac{1}{\epsilon_{bc}^2} \int_{\partial\Omega} (\vec{u} \cdot \vec{n})^2 ds$$



Low horizontal divergence of currents ($\nabla \cdot \vec{n} = 0$)

Cost function (OFF)

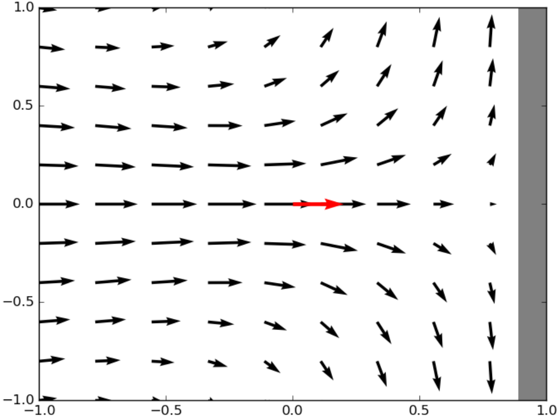
$$J_{\mathrm{div}}(\vec{u}) = \frac{1}{\epsilon_{\mathrm{div}}^2} \int_\Omega (\vec{\nabla} \cdot \vec{u})^2 dx$$



Low horizontal divergence of currents ($\nabla \cdot \vec{n} = 0$)

Cost function (ON)

$$J_{\mathrm{div}}(\vec{u}) = \frac{1}{\epsilon_{\mathrm{div}}^2} \int_\Omega (\vec{\nabla} \cdot \vec{u})^2 dx$$



3D analysis: longitude, latitude and time

- Include the data the hour before and after
- Temporal correlation length
- Coriolis force

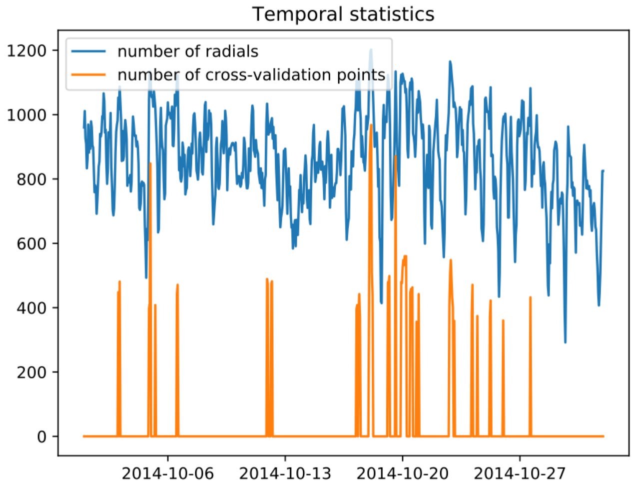
Coriolis force and geostrophically balanced mean flow

$$\frac{\partial u}{\partial t} = fv - g\frac{\partial \eta}{\partial x}$$
$$\frac{\partial v}{\partial t} = -fu - g\frac{\partial \eta}{\partial y}$$

f = Coriolis frequency
η = sea surface elevation

Cross validation

In 30 current maps with the best coverage,
some data points are marked as missing (for both sites)



Test cases: more constrains (physics) in the interpolat

Case	Description
2D	classical 2D-analysis (longitude, latitude)
2D_bc	as 2D, but with boundary conditions
2D_iv	as 2D, but imposing small horizontal divergence
3D	3D-analysis (longitude, latitude, time)
3D_Coriolis	3D-analysis with the Coriolis force
3D_Coriolis_geo	3D-analysis with the Coriolis force and the surface pressure gradient

Skill score

$$\text{S(Case)} = 1 - \frac{\text{MSE(Case)}}{\text{MSE}(2D)}$$

- The 2D case is the base-line
  for computing the relative improvement
- $\text{MSE}(C)$ is the mean square error
  (relative to the cross-validation dataset)
- If $\text{S} = 0$: reconstruction as "good/bad" as the base-line
- If $\text{S} = 1$: reconstruction matches perfectly the validation dataset.

Comparison: increased skill with more constrains

Case	RMS	Skill score	Optimal parameter(s)
2D	0.0652	0.000	ϵ^2=0.0001161
2D_bc	0.0652	0.000	ϵ^2=0.0001, ϵ^2_{bc}=10
2D_div	0.0650	0.005	ϵ^2=9.799e-05, ϵ^2_{div}=2.778e+08
3D	0.0606	0.134	ϵ^2=0.1219, lent=6904
3D_Coriolis	0.0547	0.295	ϵ^2=5.673e-05, ϵ^2_{Cor}=9.207e-05
3D_Coriolis_geo	0.0485	0.447	ϵ^2=5.37e-05, ϵ^2_{Cor}=5.65e-05, ratio=26.46

Conclusions

1. DIVA framework was extended to handle surface currents and able to handle observations when only one component of the velocity vector is measured.

Conclusions

1. DIVA framework was extended to handle surface currents and able to handle observations when only one component of the velocity vector is measured.
2. 2D analyses were used as a base-line for different test cases.

Conclusions

1. DIVA framework was extended to handle surface currents and able to handle observations when only one component of the velocity vector is measured.
2. 2D analyses were used as a base-line for different test cases.
3. Including boundary conditions and the constrain on small divergence did not improve the accuracy of the constructions.

Conclusions

1. DIVA framework was extended to handle surface currents and able to handle observations when only one component of the velocity vector is measured.
2. 2D analyses were used as a base-line for different test cases.
3. Including boundary conditions and the constrain on small divergence did not improve the accuracy of the constructions.
4. Taking for every time instance the previous and the following radial maps into account (i.e. a 3D analysis) improves the skill score .

Conclusions

1. DIVA framework was extended to handle surface currents and able to handle observations when only one component of the velocity vector is measured.
2. 2D analyses were used as a base-line for different test cases.
3. Including boundary conditions and the constrain on small divergence did not improve the accuracy of the constructions.
4. Taking for every time instance the previous and the following radial maps into account (i.e. a 3D analysis) improves the skill score .
5. Additional dynamical information improves the skill score.

Conclusions

1. DIVA framework was extended to handle surface currents and able to handle observations when only one component of the velocity vector is measured.
2. 2D analyses were used as a base-line for different test cases.
3. Including boundary conditions and the constrain on small divergence did not improve the accuracy of the constructions.
4. Taking for every time instance the previous and the following radial maps into account (i.e. a 3D analysis) improves the skill score .
5. Additional dynamical information improves the skill score.
6. Dynamical information appears to be highly beneficial when analyzing surface currents.