Using Diva on large datasets: applications and tips

C. Troupin, M. Ouberdous, F. Lenartz, D. Sirjacobs, J.-M. Beckers
& the GHER team

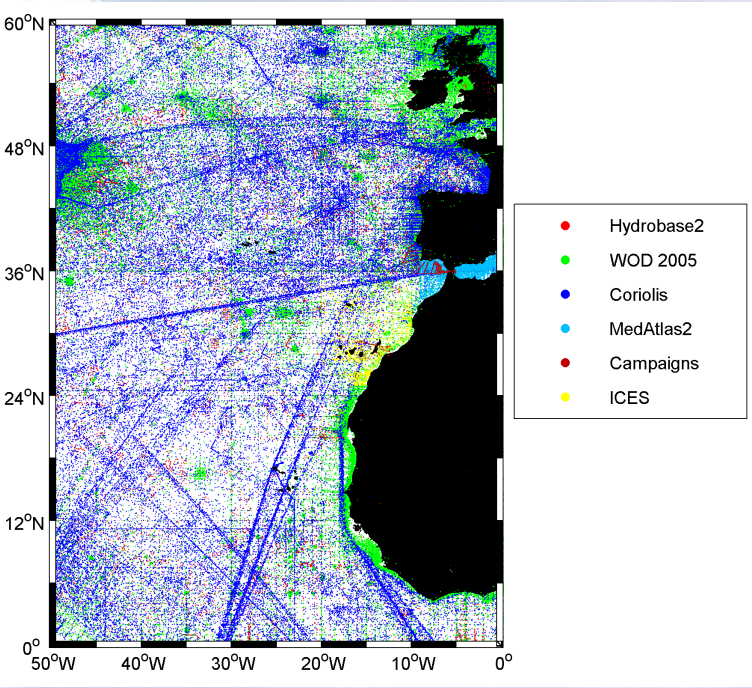
October 16, 2008

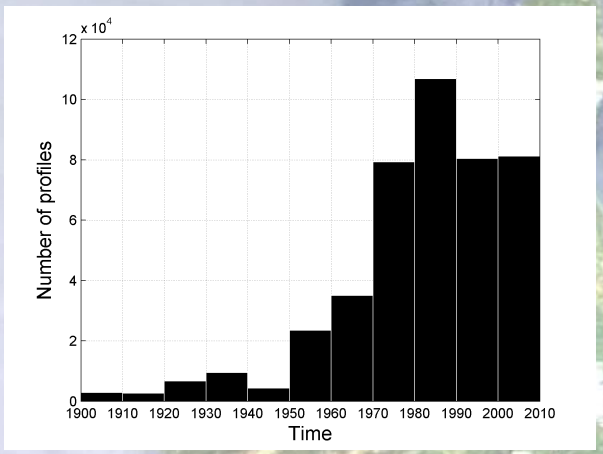
Outline

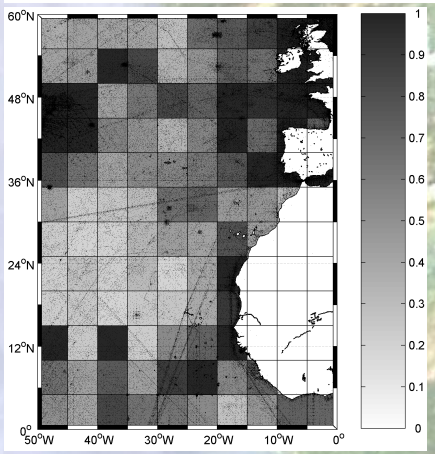
1. some applications of the tools presented yesterday
2. some tips from my tries to get a climatology in the Atlantic
3. application of Diva + satellite imagery + numerical model
4. lots of pictures

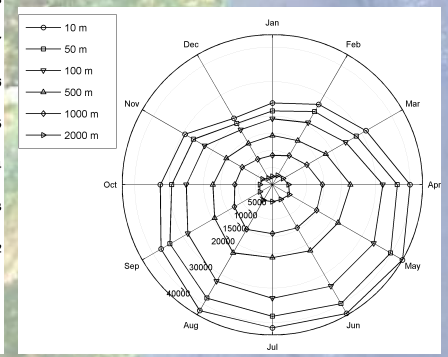
Data

- $0 - 50^\circ W, 0 - 60^\circ N$
- WOA + Coriolis + Hydrobase + MedAtlas + campaigns + misc
- from 1890 to August 2008



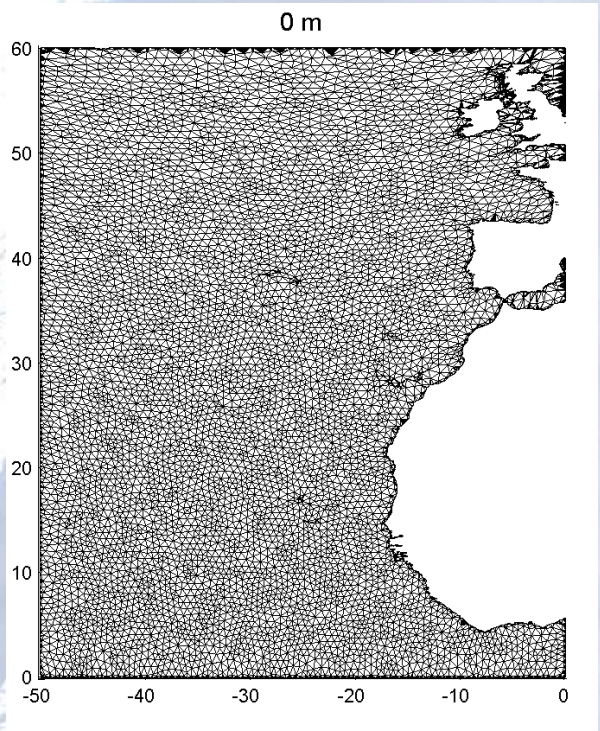






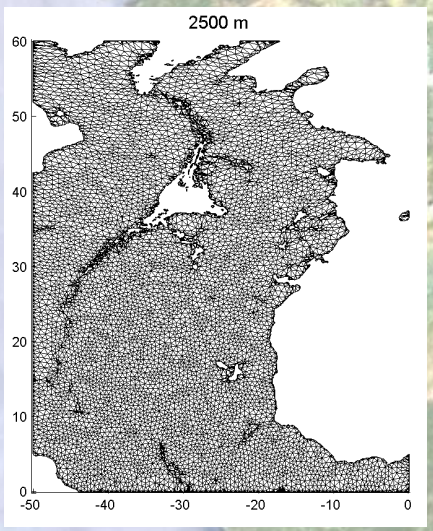
Contour and mesh

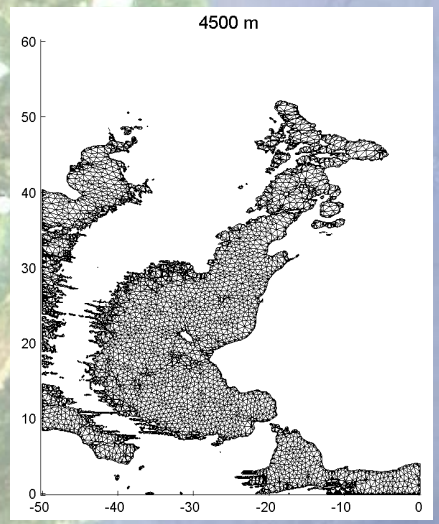
- contours from DBDBV (GEBCO too fine)
- $L = 3.0$
- same L for each level



Tips:

- generate the mesh only once for each level
  $\longrightarrow$ implementation in `3Dinfo` ?
- choose L not too small (memory issue)
- choose L not too large (resolution issue)





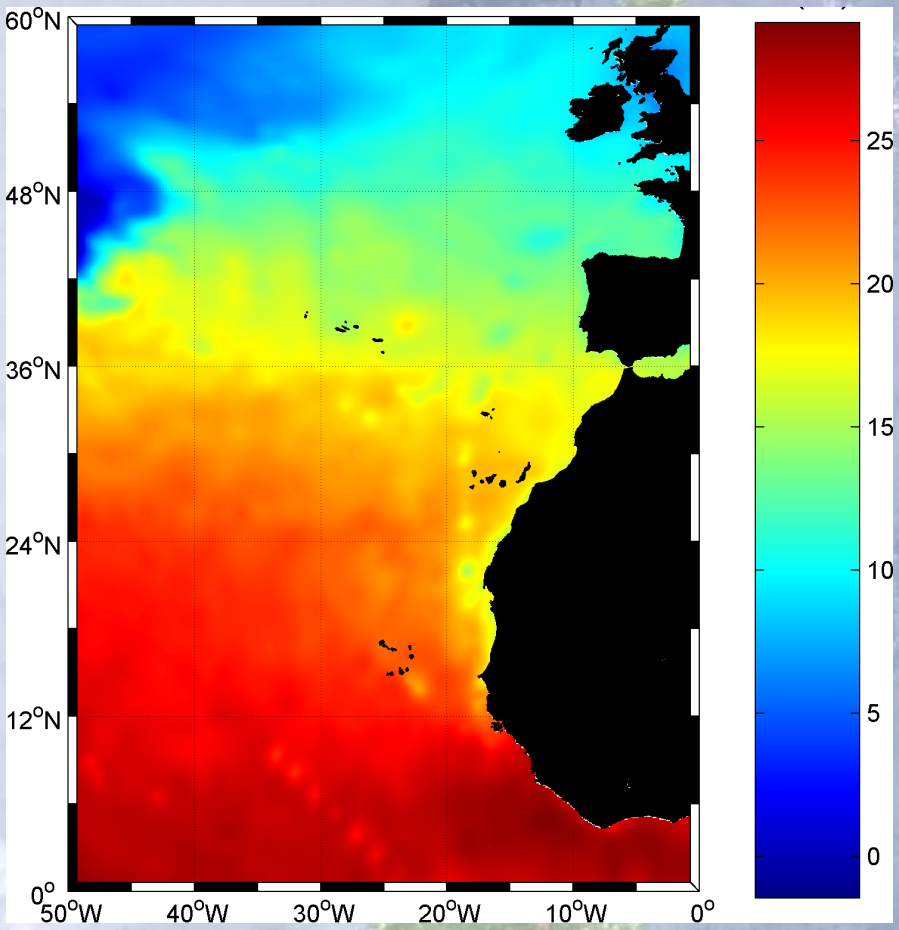
Analysis

Configuration:

- $L = 2.73$ from `divafit`
- $\lambda = 4.0$ from `divafit`
- nothing magic

Conclusions:

- `divafit` too slow with lots of data
  $\rightarrow$ use `divafit nsamples`
- visible outliers...



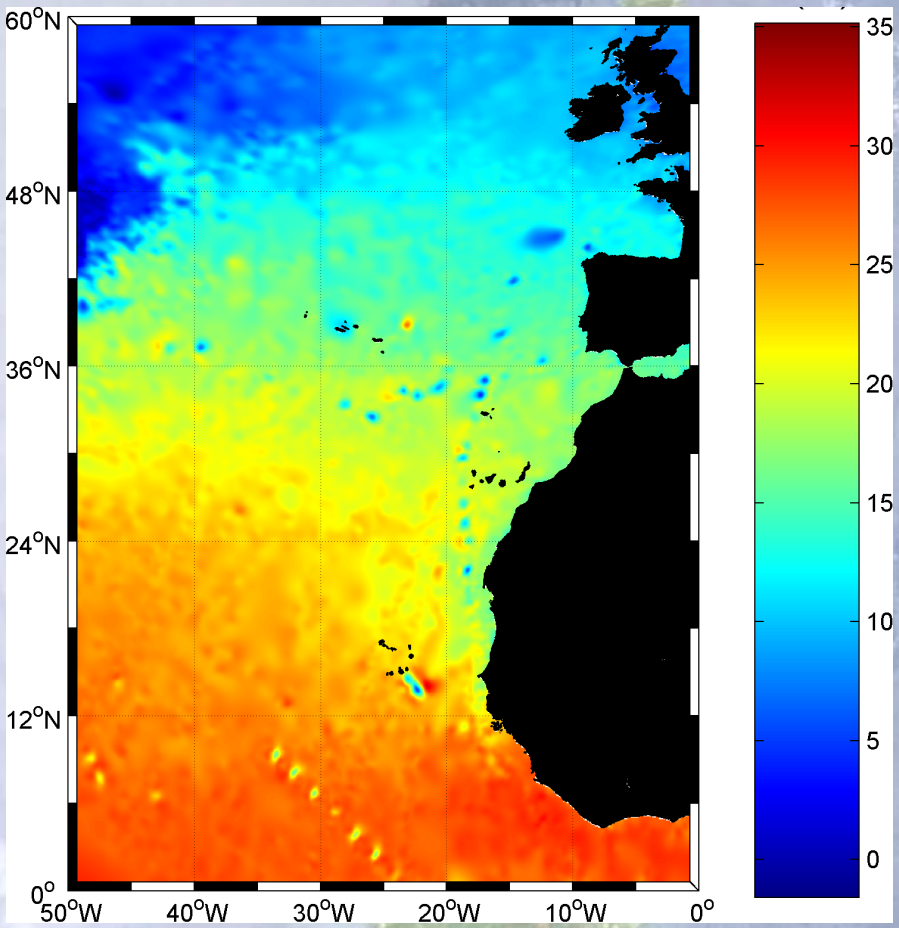
Analysis

Configuration:

- L from `divafit`
- $\lambda = 300.0$ from `divagcv`

Conclusions:

- L too small, λ too large
- outliers more visible



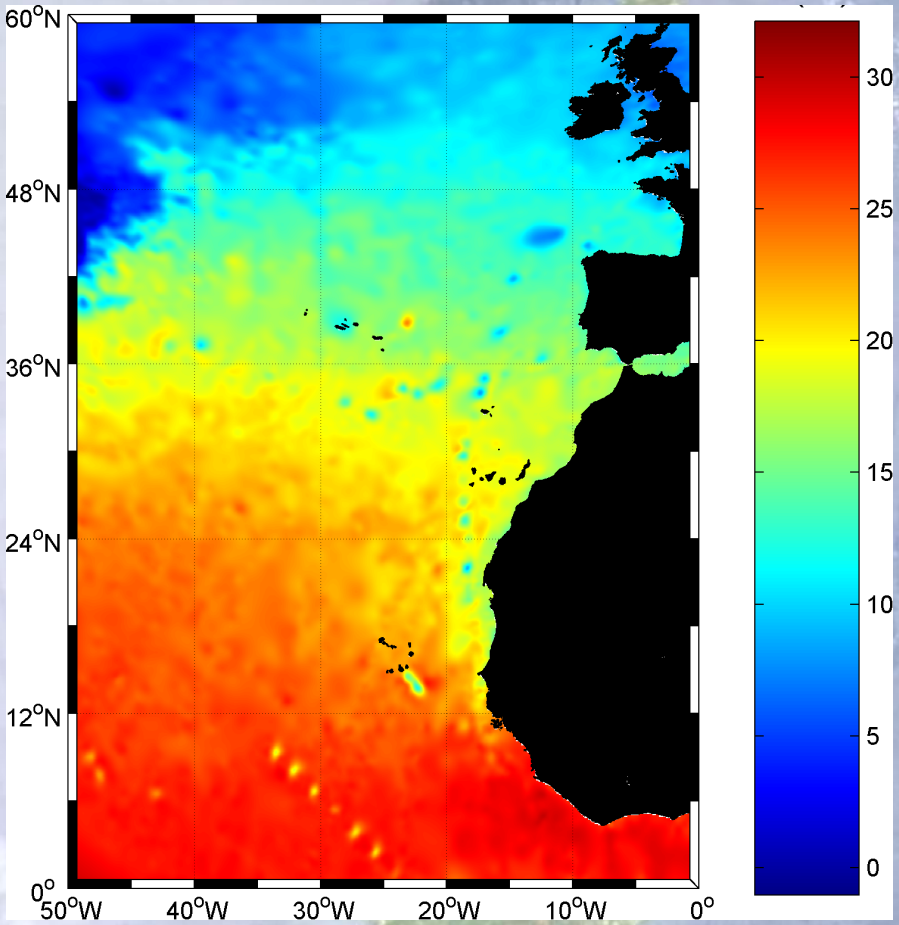
Analysis

Configuration:

- $L = 5$ (experience, physics, ...)
- $\lambda = 300.0$ from `divagcv`

Conclusions:

- λ still too large
- outliers did not disapear (no magic)



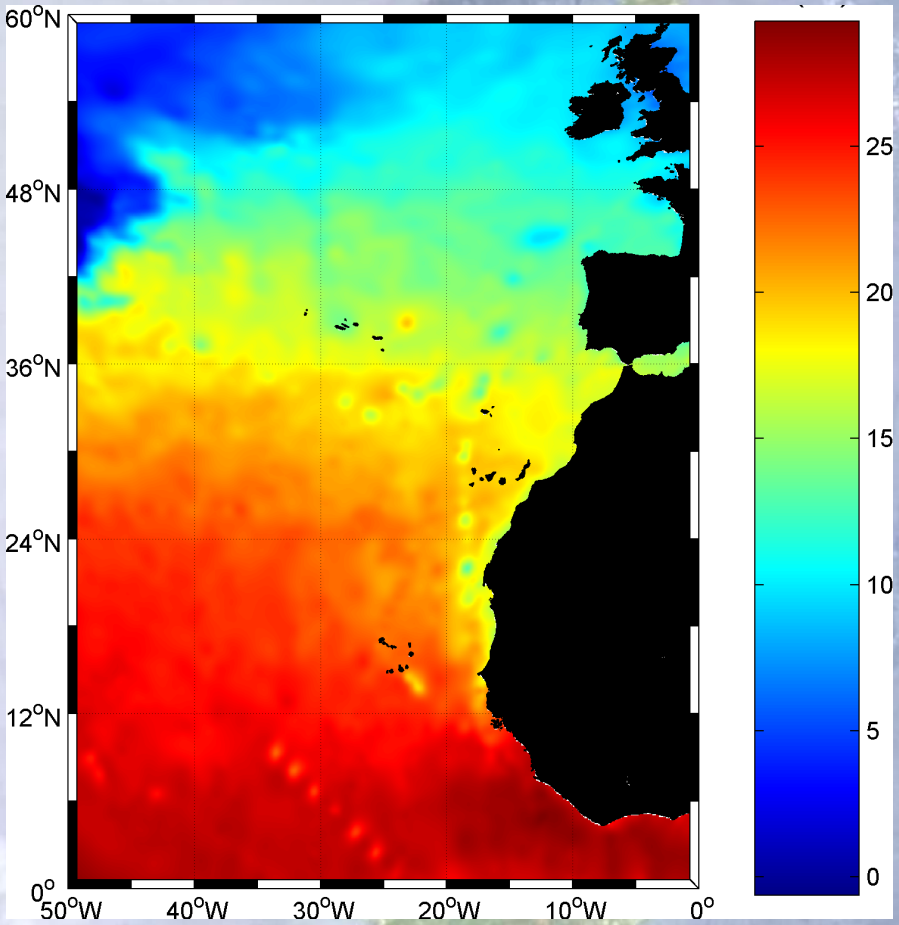
Analysis

Configuration:

- L from `divafit`
- $\lambda = 14.4$ from `divacvrand 2000 2 -r`

Conclusions:

- more realistic estimate of λ using `divacvrand`



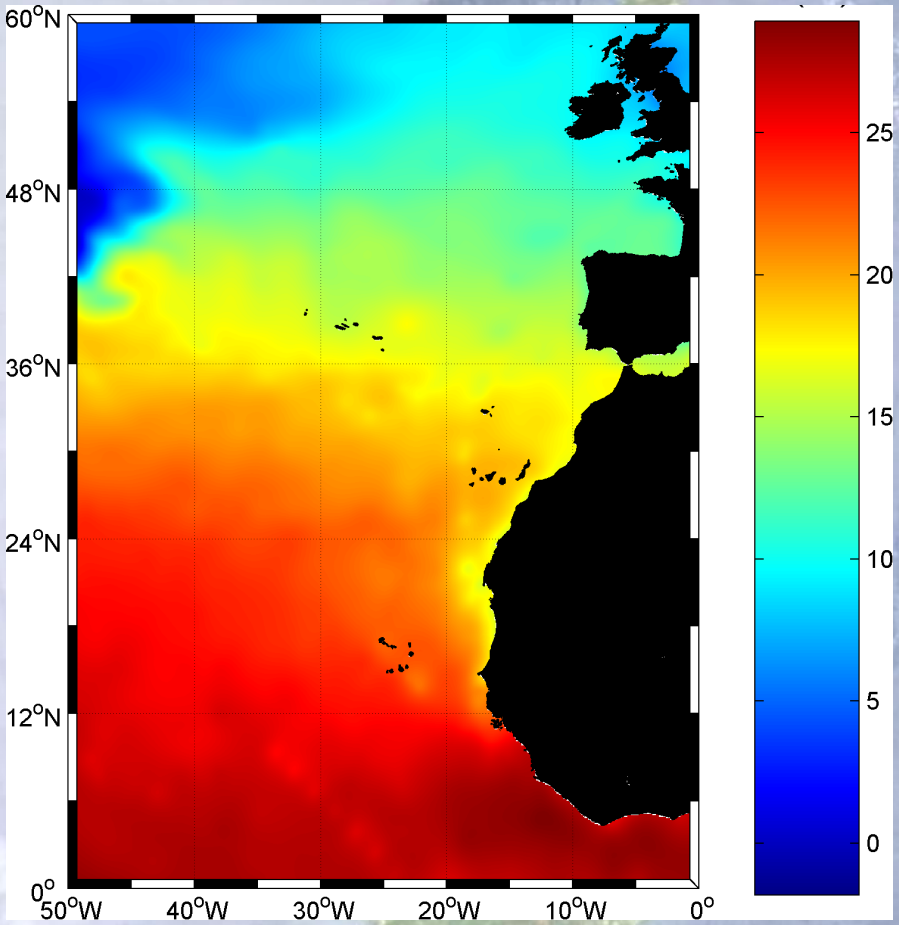
Analysis

Configuration:

- L from `divafit`
- $\lambda = 1$

Conclusions:

- very smooth field
- outlier still visible



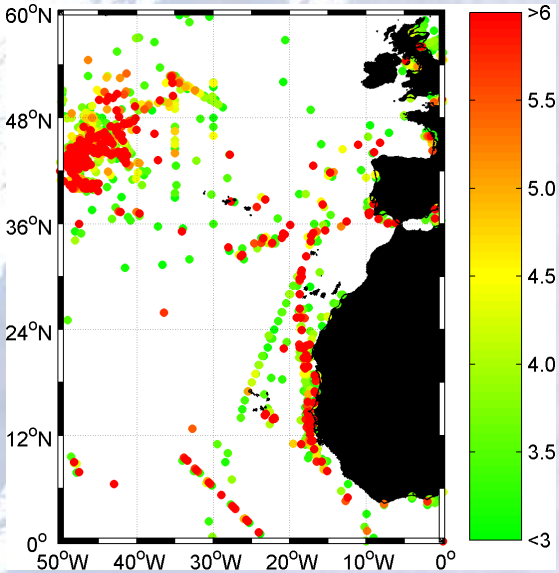
Analysis

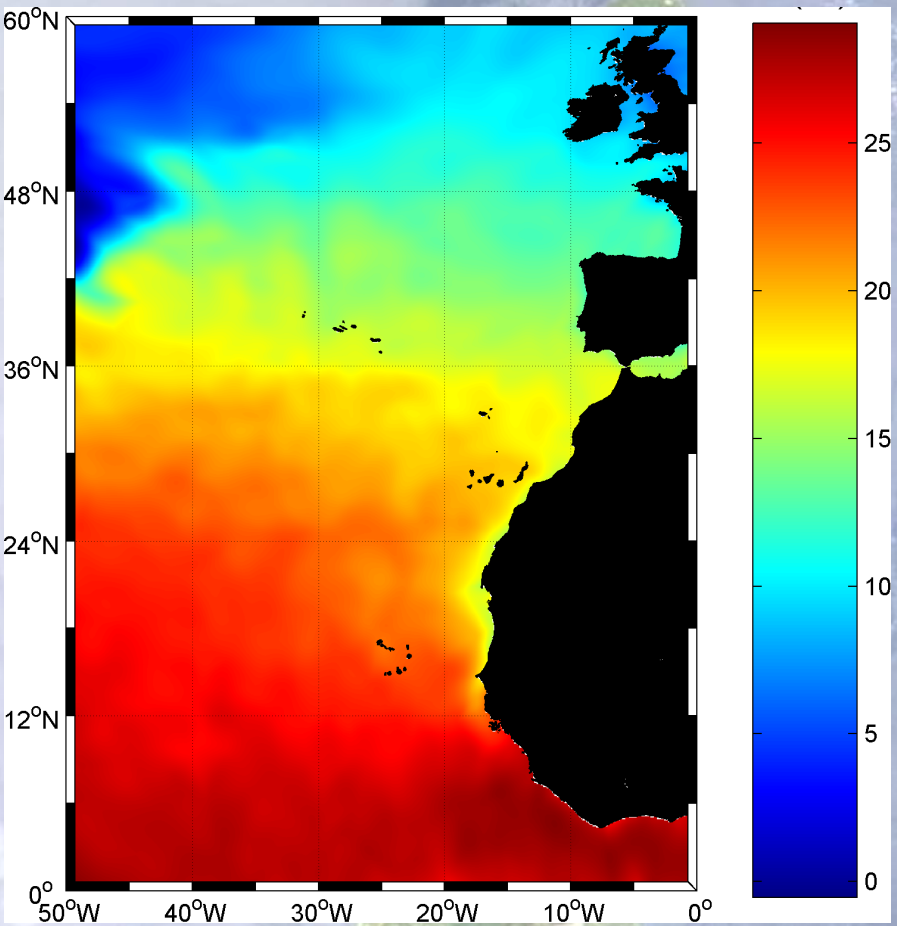
Configuration:

- L from `divafit`
- λ from `divacvrand`
- outlier removal with `divaqcbis`

Conclusions:

- 696 outliers detected
- improved results





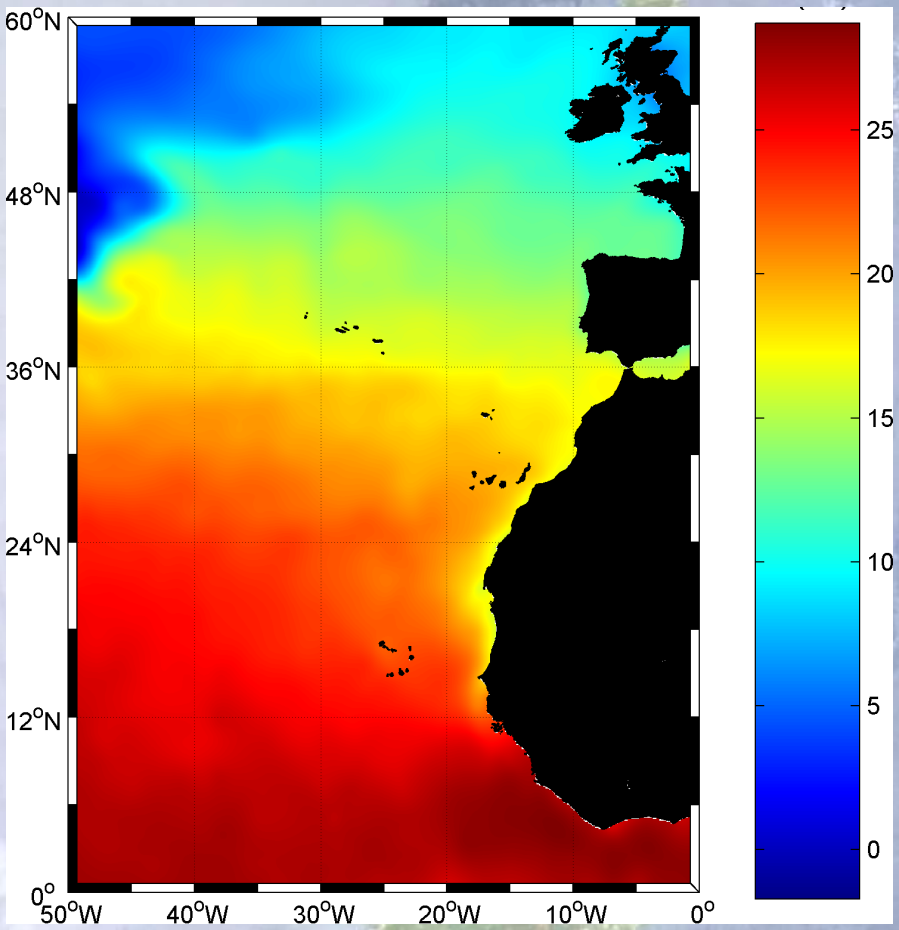
Analysis

Configuration:

- outlier removal
- L from `divafit`
- $\lambda = 1$ (with `divagcv` 63.5)
- semi-normed analysis

Conclusions:

- lots of steps before satisfying field
- need good parameters before using `divaqc`
- semi-normed analysis helps where few data are available



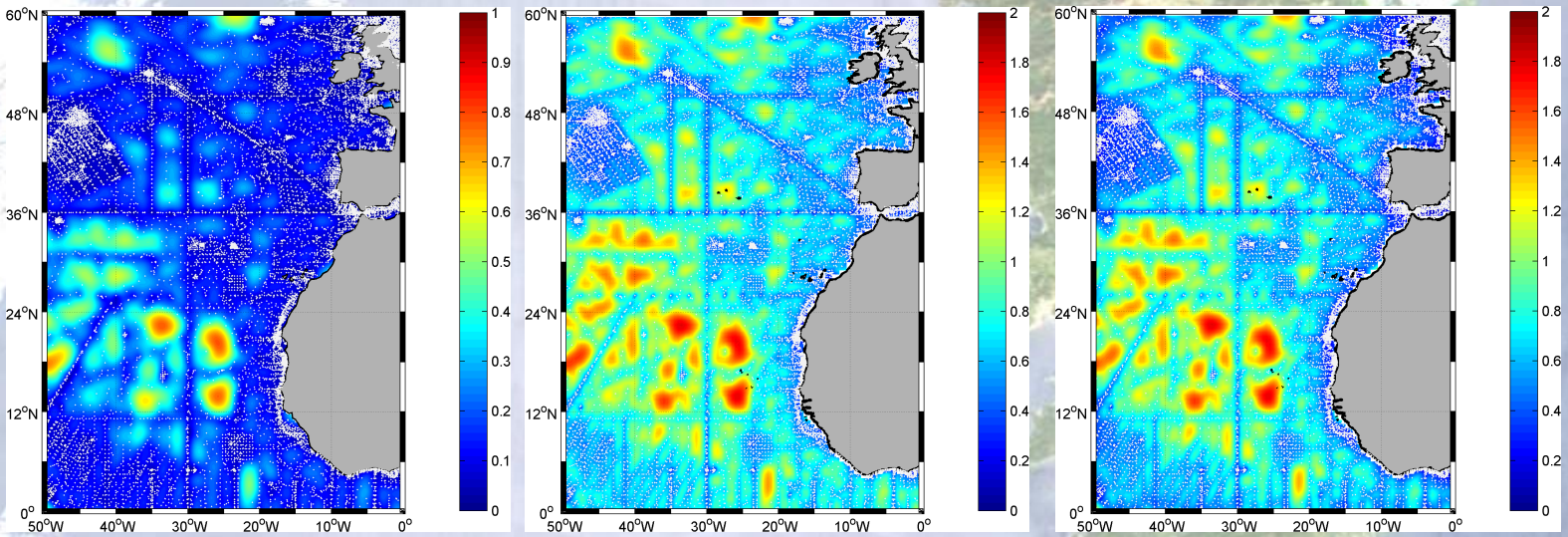
Error fields

Methods:

1. *poor man's* error estimate (quick but underestimated)
2. hybrid approach, analogy with O.I
3. real covariance function → don't use your laptop!

Conclusions:

1. highly depends on data coverage
2. same distribution of errors
3. different orders of magnitude



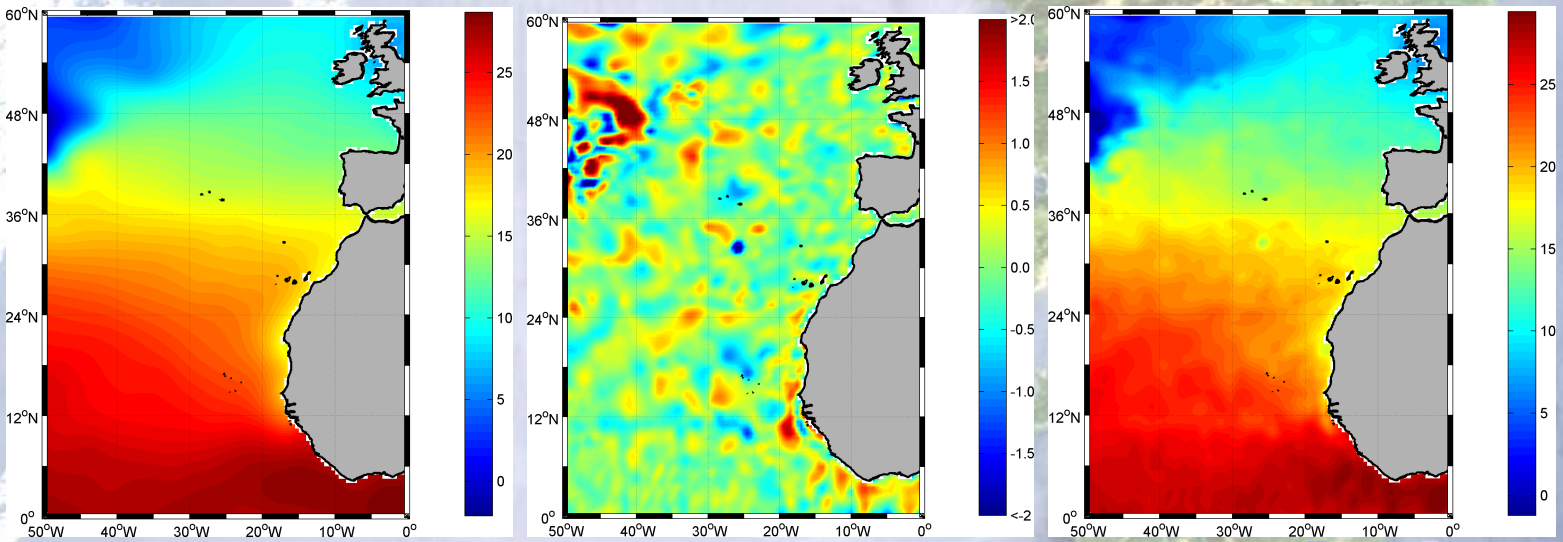
Semi-normed analysis

Methods:

1. create a reference field with large L and small λ
2. compute anomalies with respect to reference
3. analyse the anomalies
4. sum up the 2 fields

Utility:

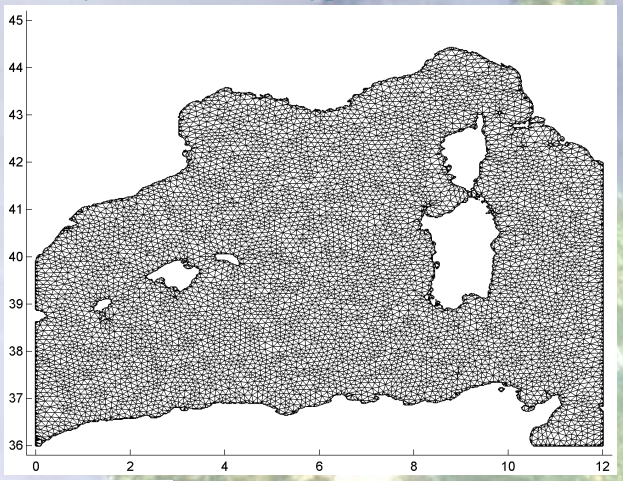
1. better field where no data are available

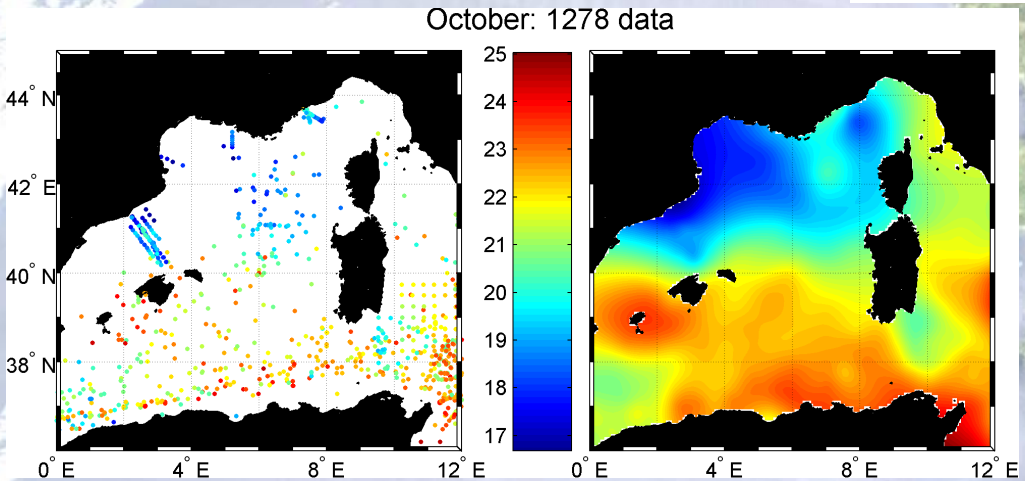


Diva

Data-Interpolating Variational Analysis
[Brasseur et al. (1996), Brankart and Brasseur (1996, 1998)]

- = advanced data-gridding method
- + finite-element resolution
- + coastlines + advection influence
- + error maps
- + many more

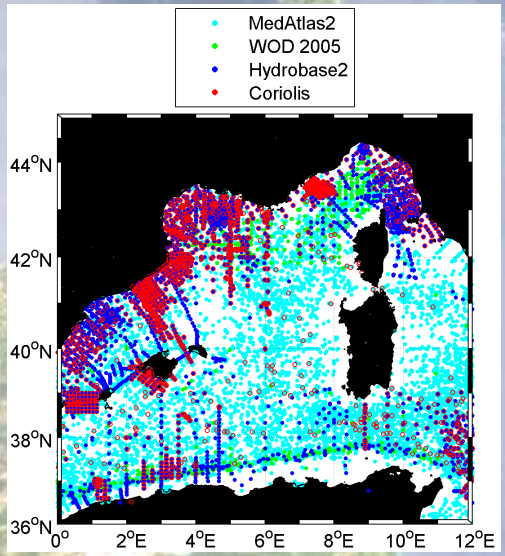


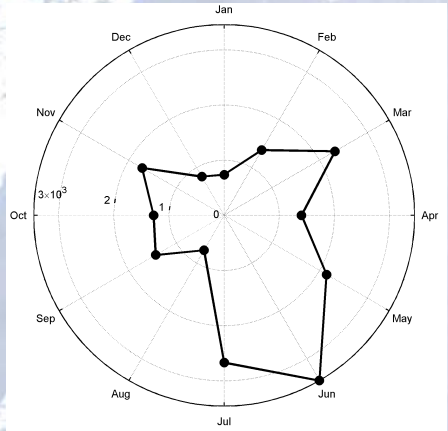


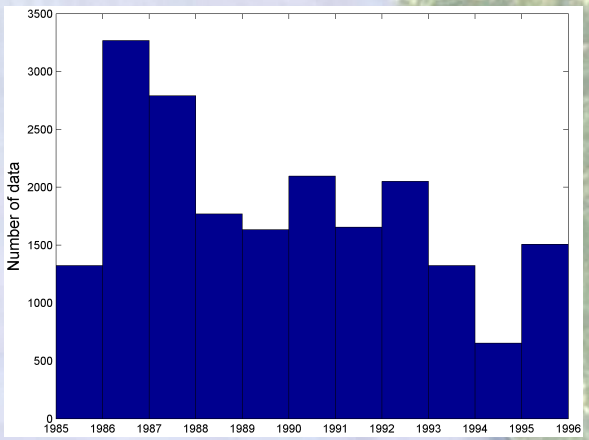
Data used

Data set = aggregation between:
WOD05 + MedAtlas2 + Hydrobase2 + Coriolis

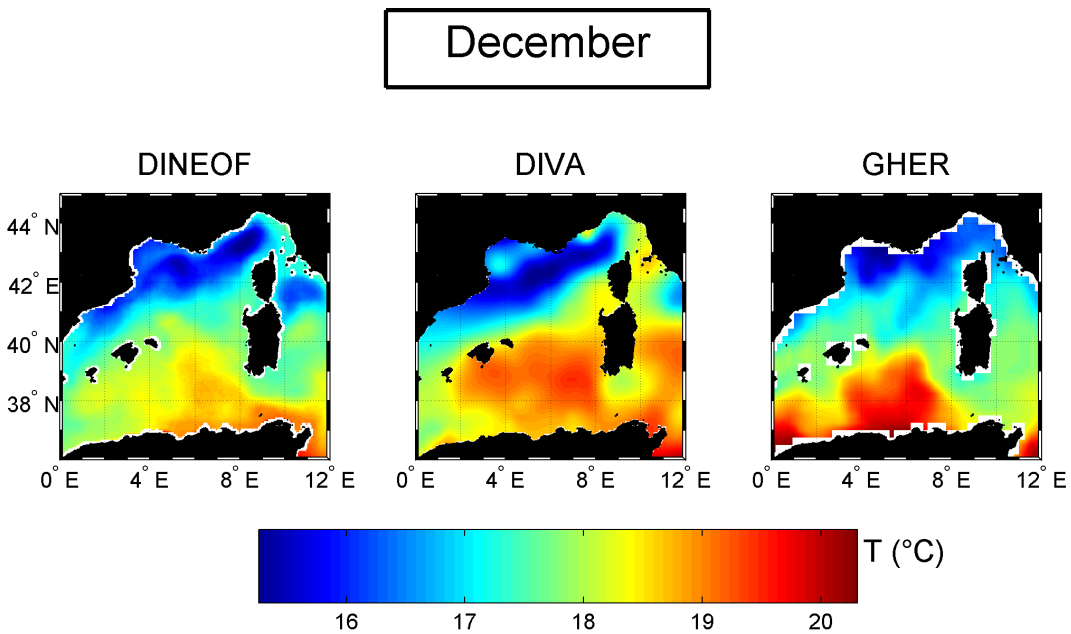
- 24293 unique profiles
- observations with depth $< 5\ m$ (82.6% of the profiles)
- 1986: 3267 measurements, 1994: 653
- $> 40\%$ of data between May and July
- mean: $18.4769^\circ C$, standard deviation: $3.8773^\circ C$
- 99.3% of data have $12 < T < 27^\circ C$







Monthly temperature fields



Thanks for your attention!