Data transformation in DIVA

J.-M. Beckers, A. Barth, M. Belounis, M. Ouberdous, C. Troupin,
with the help of M.-E. Toussaint, A. Alvera-Azcárate,
A. Capet, L. Geron, F. Lenartz

MARE, AGO Department
University of Liège

3-7 November, 2010

Why transform the data?

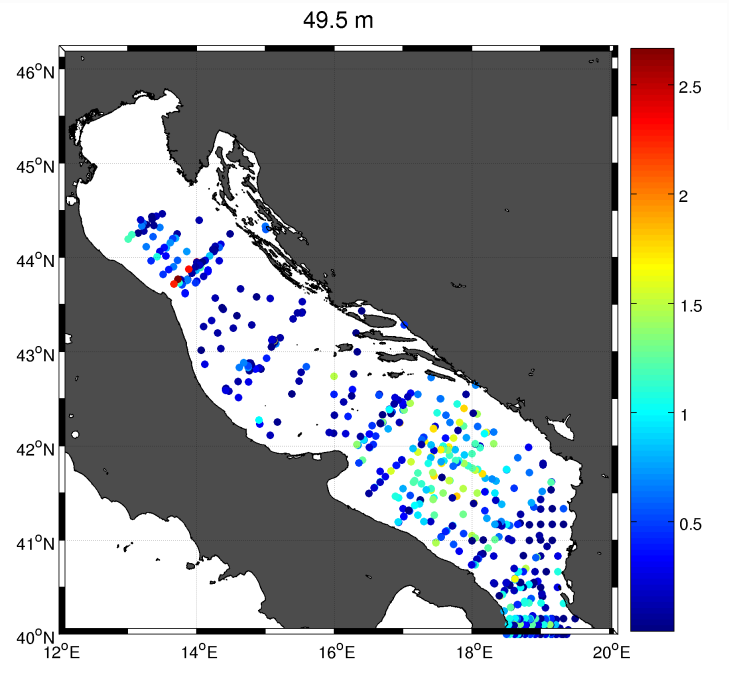
Problem:

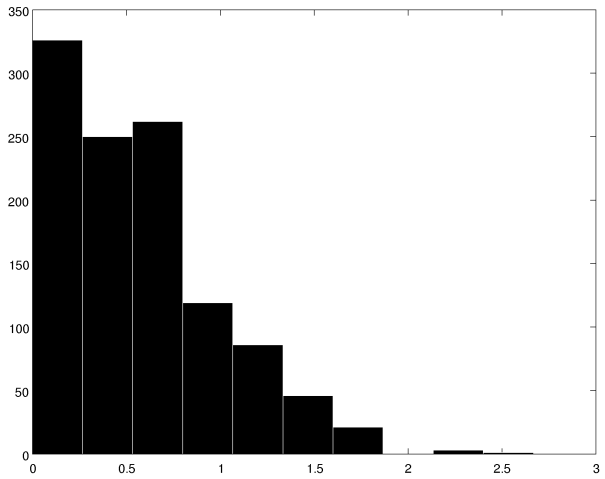
- Non-Gaussian variable $\longrightarrow$ unrealistic values in analysis
- Example: negative concentration, negative salinity, ...

Solution:

1. Apply a transformation on the data, prior to analysis
2. Perform the analysis
3. Apply inverse transformation on analyzed field

Example: nitrate concentration in the Adriatic

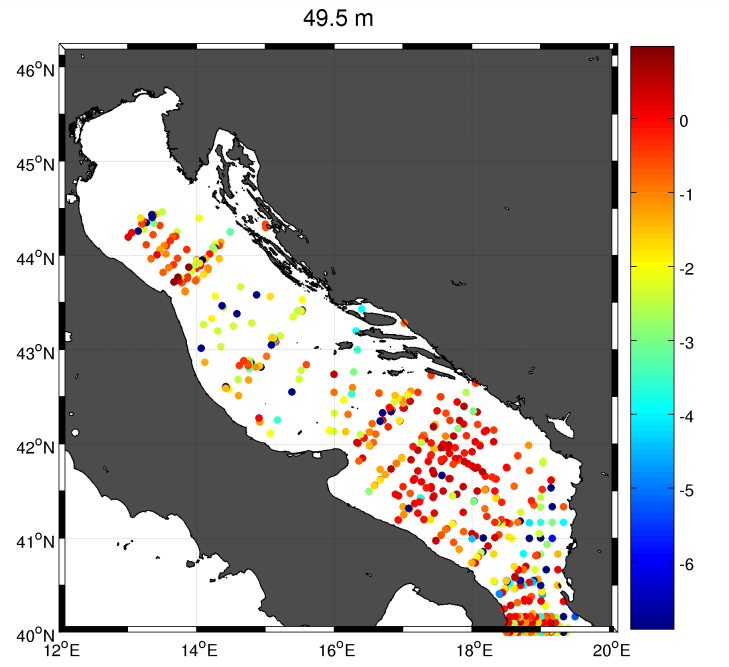


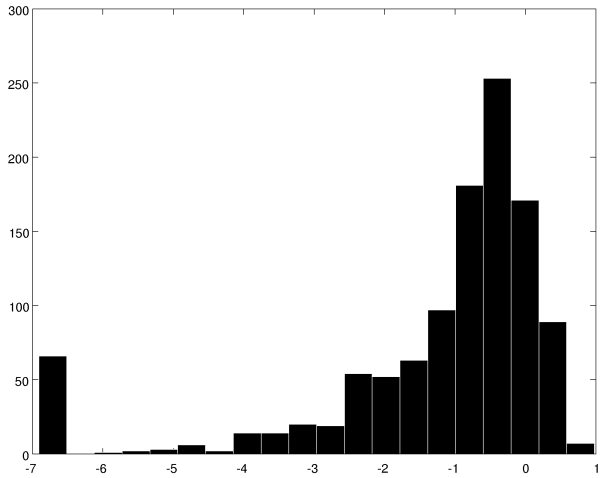


Transformation 1

Logarithm:

- \+ transformation not dependent on the data
- \+ inverse transformation is known
- – needs strictly positive values
  $\longrightarrow$ problem with null concentrations
- – inverse transformation can create extrema

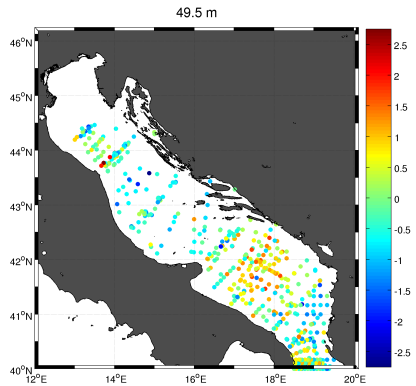


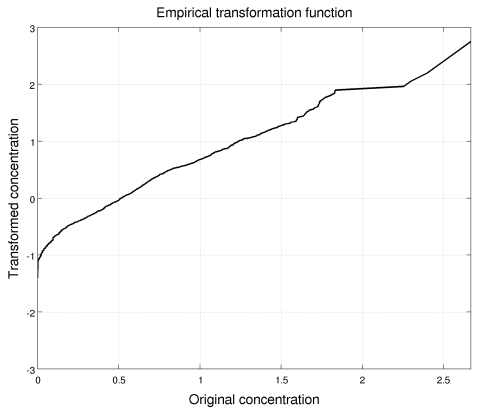


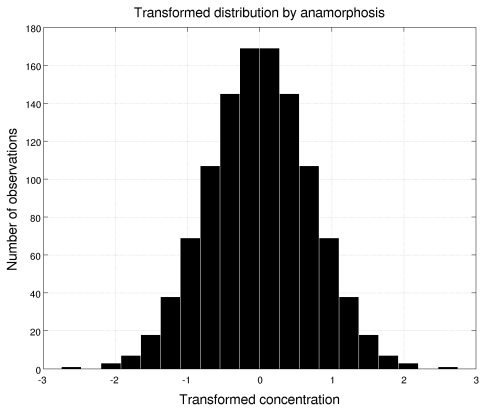
Transformation 2

Anamorphosis:

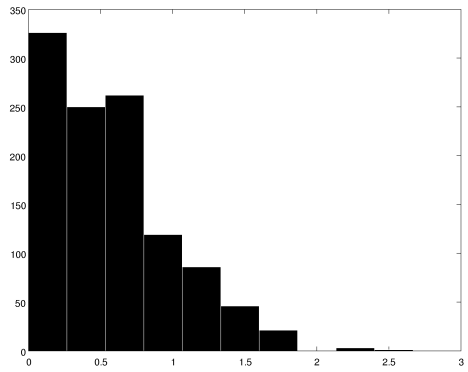
- \+ transformation computed from the data themselves
- \+ final distribution is close to Gaussian
- – transformation function may require extrapolation

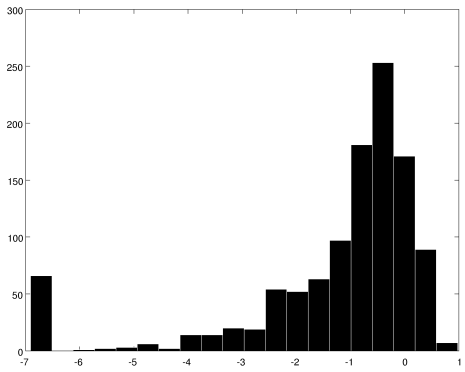


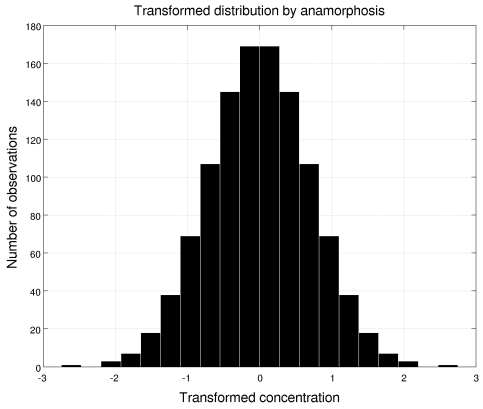




Comparison: histograms







Implementation in GODIVA

Analysis option in the `driver`

- 10 : exponential applied on data
- 11 : logarithm of data (small value added to avoid $\log(0)$)
- 12 : data^2
- 13 : anamorphosis (optimization phase)
- 14 : user-chosen transformation (Fortran file to edit)

Remark: during extraction

use file `var.bounds` to specify `min` and `max`