Empirical correction on XBT data and global field reconstruction using EOFs

Hamon M., Le Traon P.Y., Reverdin G.

The First XBT Science Workshop
Melbourne 07/2011

- Introduction

- 1 The XBT problem

  1.1 Database and collocation method

  1.2 Test of the W08 correction

  1.3 A new correction

- 2 EOF reconstruction

  2.1 Method

  2.2 Test with ORCA025 G70

- 3. Results

- Conclusion

Introduction

- Identifying and quantifying the effects of global warming is one of the most important research areas for the international oceanographic community.
- It is difficult to accurately estimate temperature (and salinity) trends and variations because of sparse sampling of data and instrumental biases.

  → Estimations depend on the quality of data and the method for filling gaps.
- Expendable BathyThermograph (XBT) system does not measure directly the depth of the probe, it uses a fall rate to estimate it.
- Gouretski and Koltermann (2007) used a CTD climatology to identify a positive temperature bias of XBT.
- Wijffels et al (2008) (W08) proposed a yearly correction of immersion which is a linear function of depth.
- Levitus et al (2009) used a simpler temperature correction to estimate the ocean heat content.
- Ishii and Kimoto (2009) proposed a correction by type of XBT to avoid regionally biases.
- Gouretski and Reseghetti (2010) (GR10) proposed a new depth correction dependent on ocean temperature added to a thermal offset.

The XBT problem

Data and collocation method

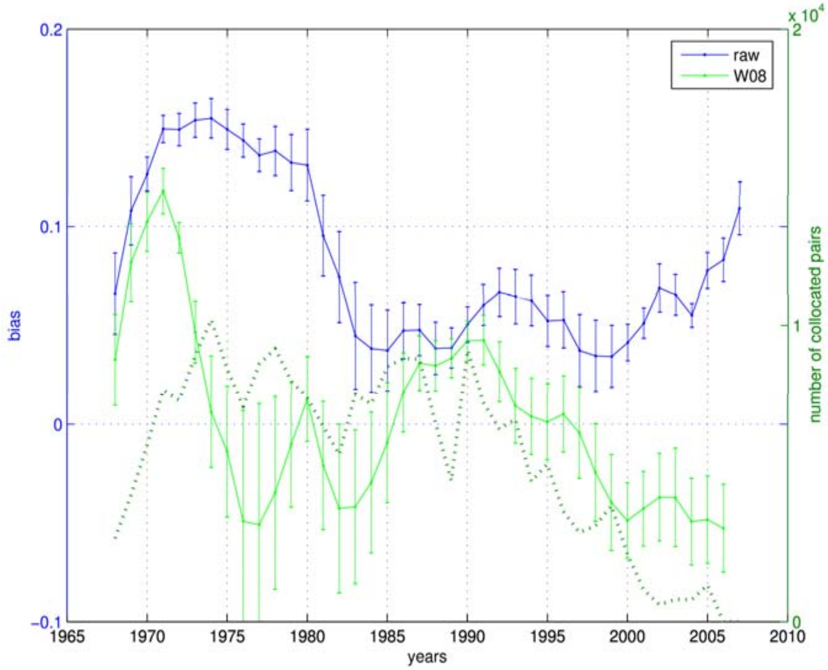
- We use WOD05 profiles, interpolated to standard levels.
- CTD and OSD are our reference profiles
- XBT have been processed when identification was possible with the Hanawa correction (Hanawa et al 1995).
- Rather than to use climatologies as W08, we use a collocation method (1°*2°*15 days).
- For each individual XBT profile, we calculate the median of all CTD/OSD selected in the collocation area, to obtain a single reference profile.
- Using the median is preferred for this kind of data distribution, it reduces influence of outliers.
- We ensure that the difference of bathymetry is less than 500m (comparisons between collocated profiles can yield unrealistic biases in continental slope regions).
- This method allows us to capture about 10^4 profiles per year between 1968 and 2007.

The XBT problem

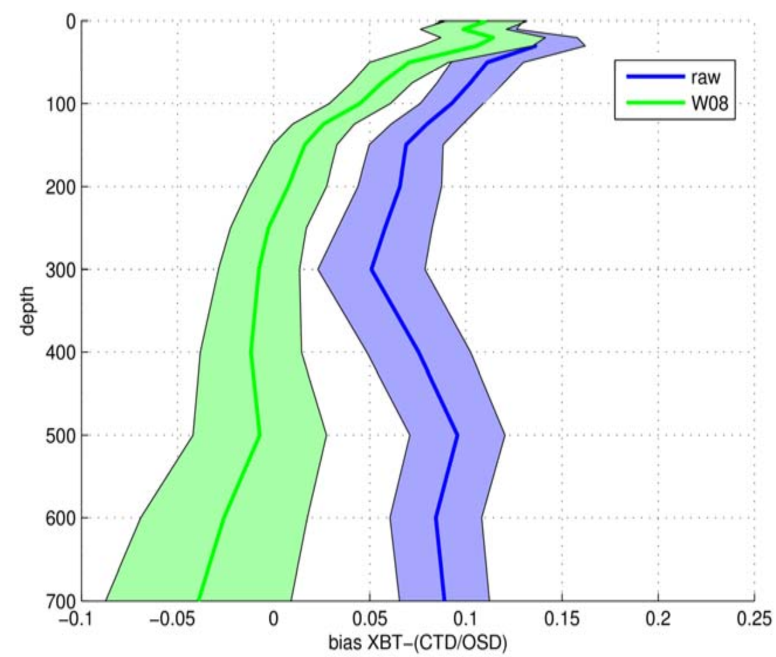
Test of the W08 correction

- The W08 is a linear annual correction on depth. It distinguishes XBTS (shallow) and XBTD (deep):

$$Z_{true} = Z(1-r)$$



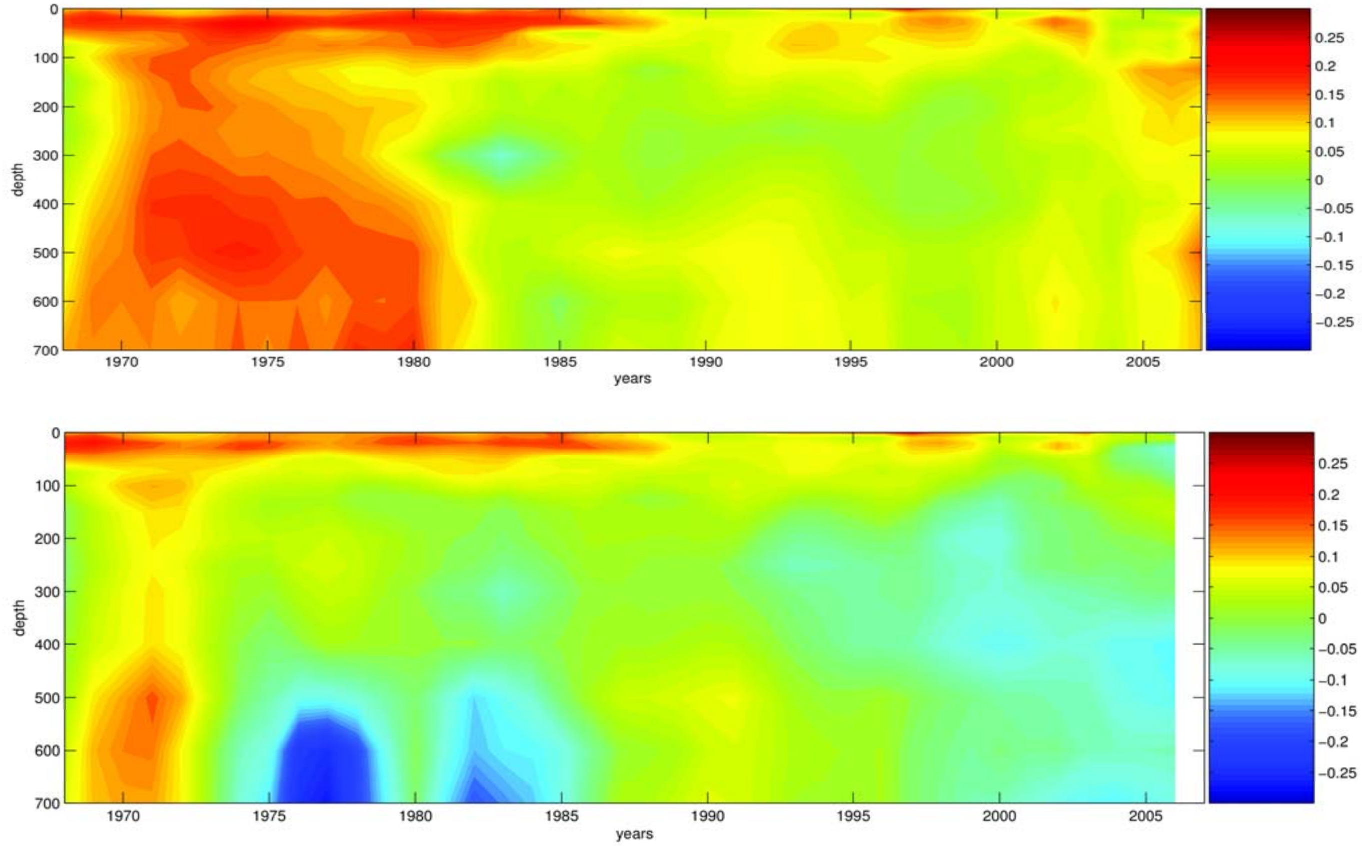
Evolution of XBT-CTD median raw bias (blue) and corrected by W08 (green) integrated between 0m and 700m.



Median raw bias (blue) and corrected by W08 (green) function of depth on average on the study period.

The XBT problem

Test of the W08 correction

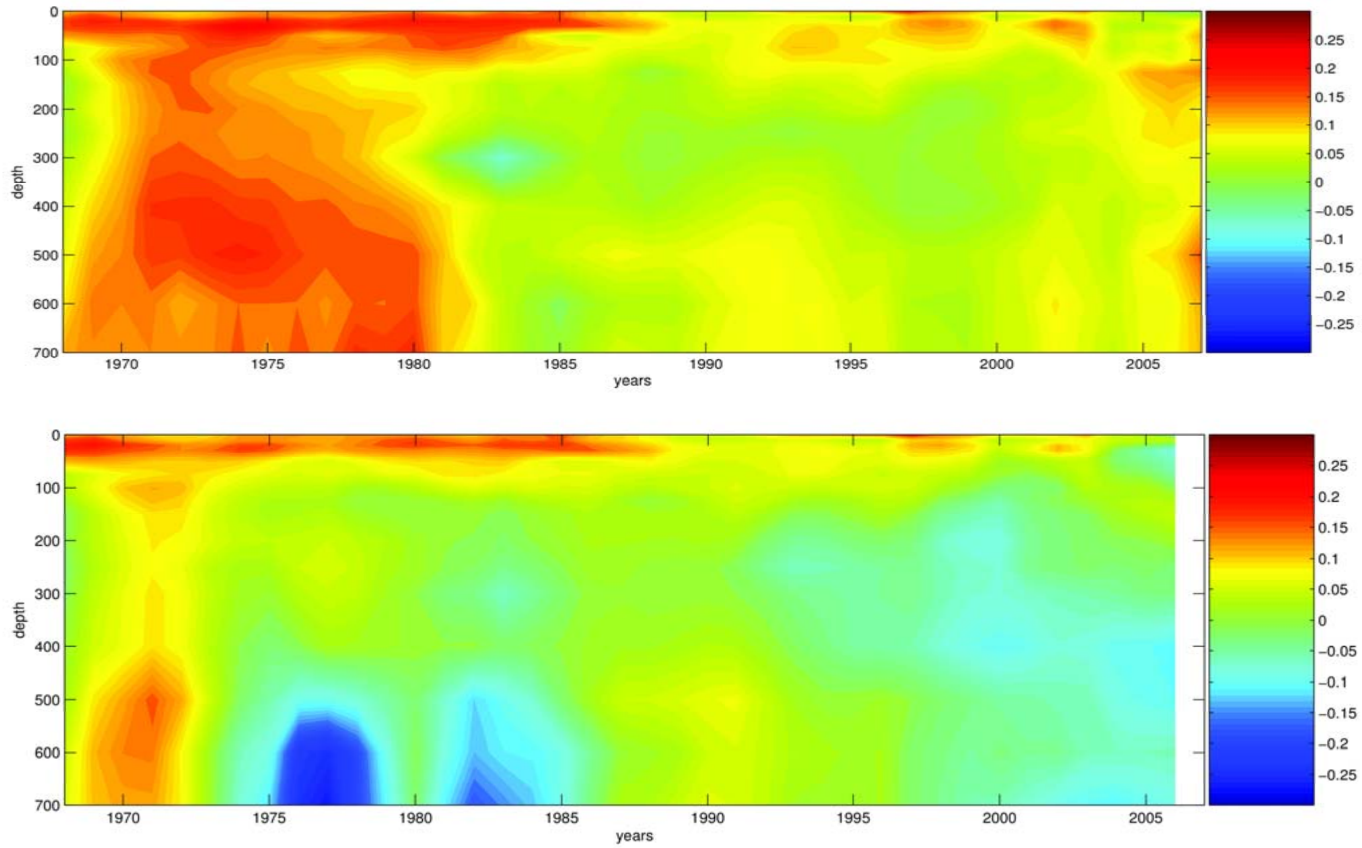


Evolution of the median raw bias (top) and corrected by W08 function of depth and years (°C)

- The linear correction is not always performing well (with our collocation method and dataset) especially between 1975 and 1985. It provides too strong correction below 500m depth and a too small correction for surface layers.

The XBT problem

Test of the W08 correction



Evolution of the median raw bias (top) and corrected by W08 function of depth and years (°C)

- The linear correction is not always performing well (with our collocation method and dataset) especially between 1975 and 1985. It provides too strong correction below 500m depth and a too small correction for surface layers.

The XBT problem

A new correction - Definition

- Correction in 2 steps :

  1. *Correction of thermal offset*

  $$\boxed{T = T_{XBT} - T_{off}}$$

  2. *Correction of the depth bias*

  $$\boxed{Z = Z_{obs}.(1 - A - B.Z_{obs}) - Z_{off}}$$

- Correction based on previous studies (W08, GR10).

- We separate XBTS (predominantly T7/Deep Blue) and XBTD (predominantly T4/T6) as in W08.

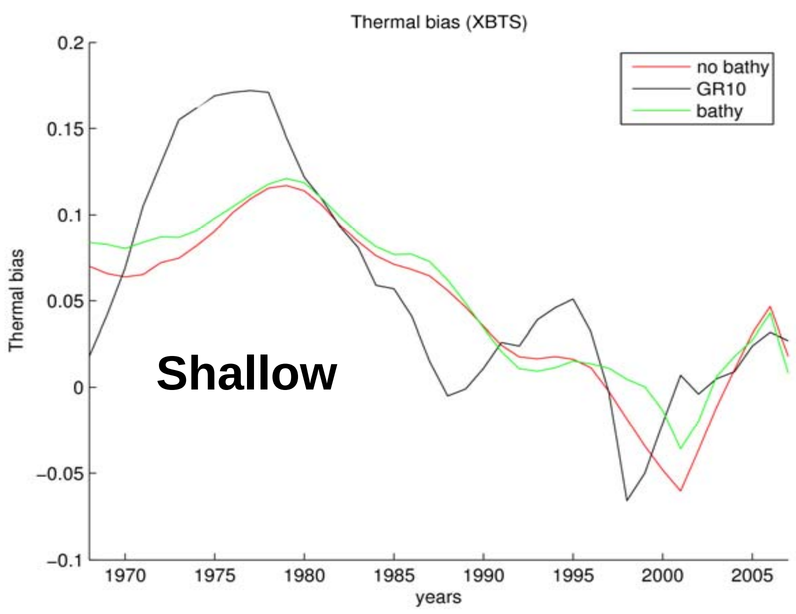
- Time-dependant correction on temperature (as in GR10).

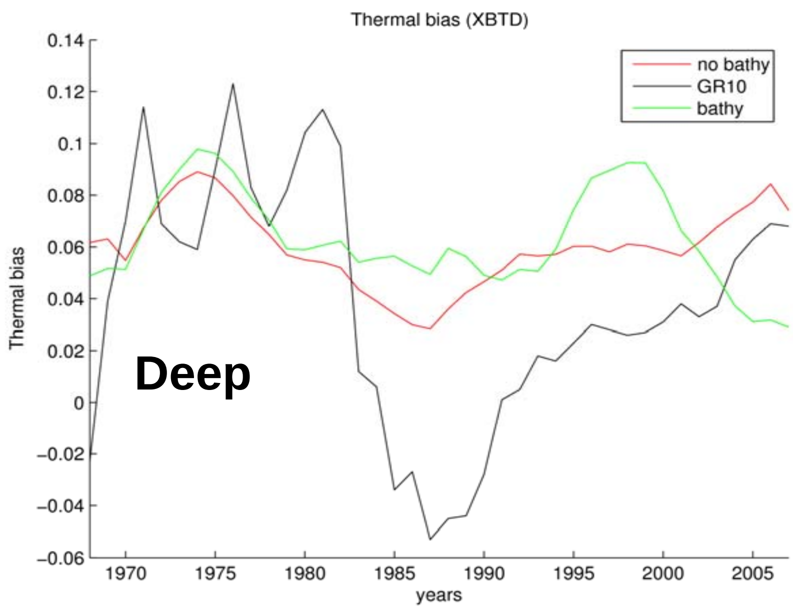
- Time-dependant correction on the depth (as in W08).

The XBT problem

A new correction - Thermal Offset

- A thermal offset is necessary and is computed using profiles with a low temperature gradient (< 0.0025°C/m) in the upper 30m.
- Unrealistic biases due to differences of bathymetry between collocated XBTs and CTDs (red and green line).
- Similarities with the GR10 correction's evolution, but the values are quite different.





Thermal Offset (°C) calculated for XBTS (left) and XBTD (right) function of years.The GR10 thermal offset is indicated with the black line

Ifremer LOCEAN Coriolis OPERATIONAL OCEANOGRAPHY

The XBT problem

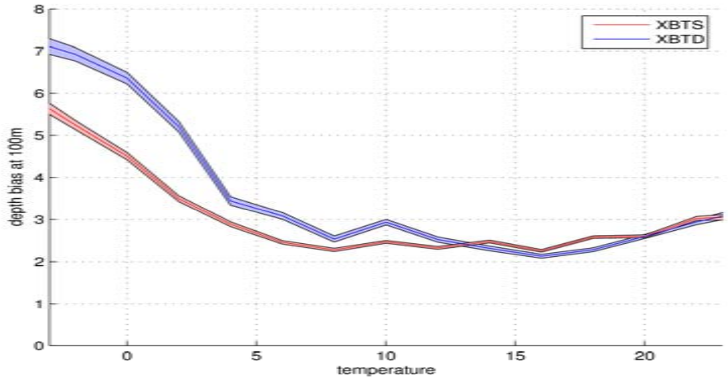
A new correction - Second order correction

- We compute the annual median depth bias using, $$dZ = (T_{CTD} - T_{XBT}) \frac{\delta Z}{\delta T_{CTD}}$$

- The difference between collocated profiles does not seem to indicate a depth linear function correction, but rather a second order function and an offset.

- Between the surface and 30m, the bias doesn't follow a parabolic behavior because of high variability noise and low gradient in the surface mixed layer.

- Relation between depth bias and the temperature where the probe has been deployed.



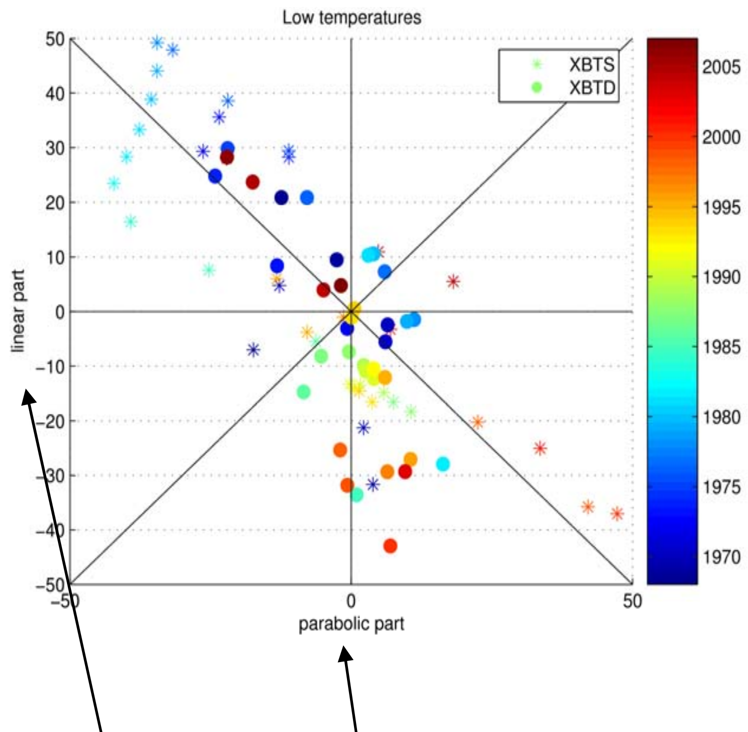
→ We can't distinguish XBTS from XBTD at a given depth.

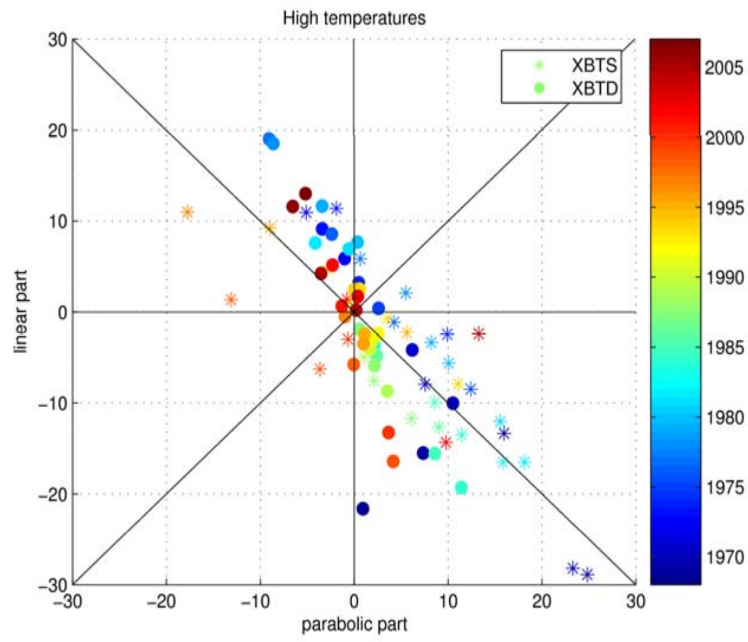
Median XBT-CTD depth bias (m) at 100m function of absolute latitude for XBTS (red) and XBTD (blue)

The XBT problem

A new correction - Second order correction

Linear part function of parabolic part and years in meters, at 400m for XBTS (stars) and XBTD (filled circles) at low temperatures (left) and high temperatures (right).





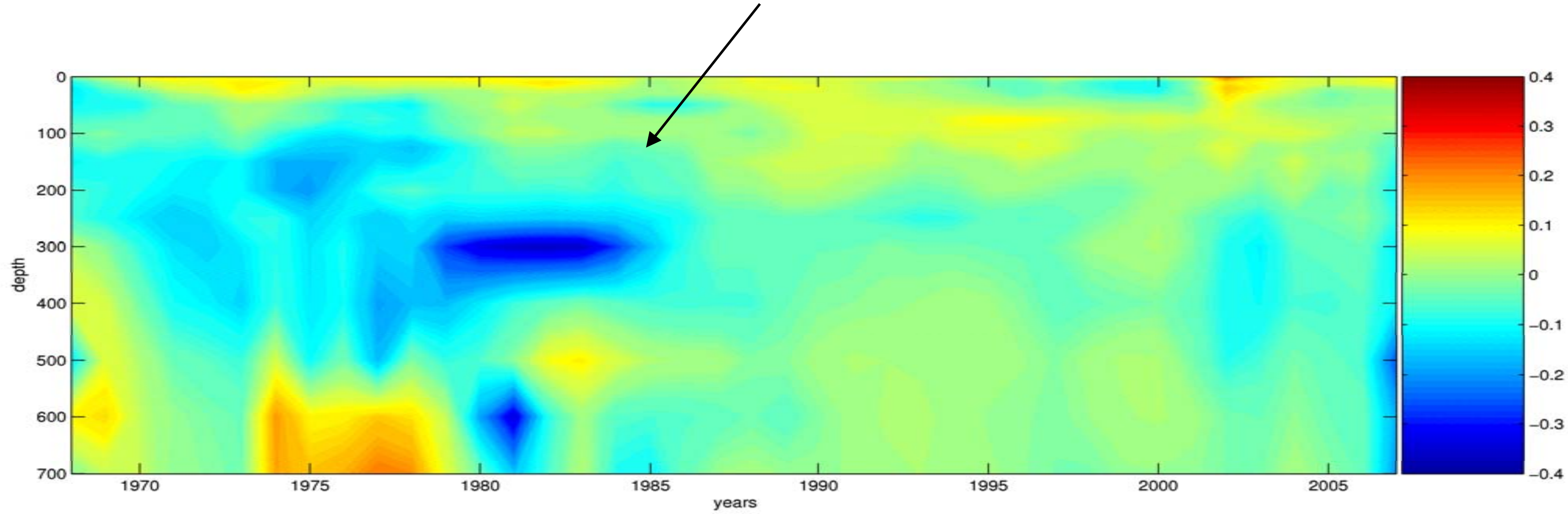
$$Corr = A(t, xbt, lat)z + B(t, xbt, lat)z^2 + OFFSET(t, xbt, lat)$$

- → Separation of XBT into 4 classes:
  - XBTS and XBTD
  - Low and high temperatures (10°C)

The XBT problem

A new correction - Specific case

- A strong negative temperature bias is found in the western pacific basin after the global correction.
- It is predominantly located at 300m between 1970 and 1985.



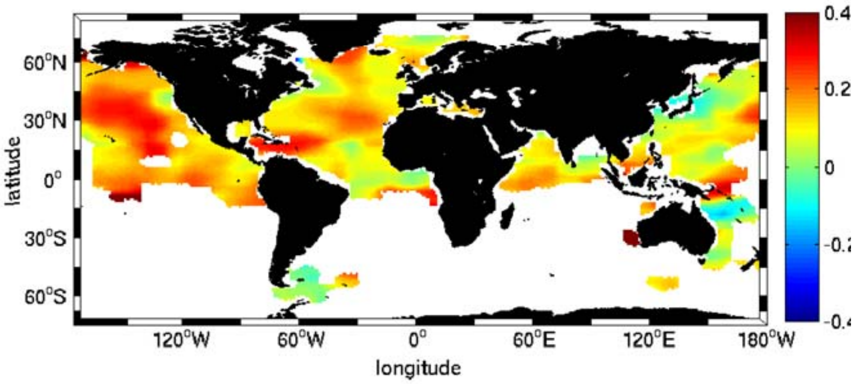
Evolution of XBT-CTD median globally corrected bias for XBT deployed in western pacific bassin, function of depth and years (°C).

→ need to apply a regional correction.

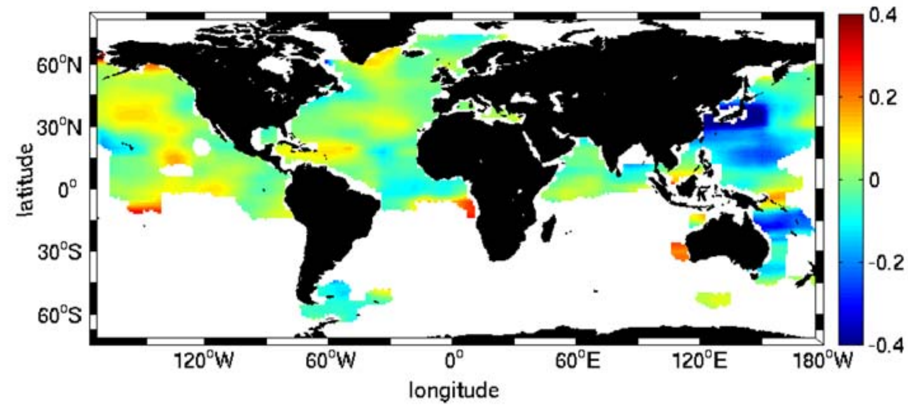
The XBT problem

A new correction - Specific case

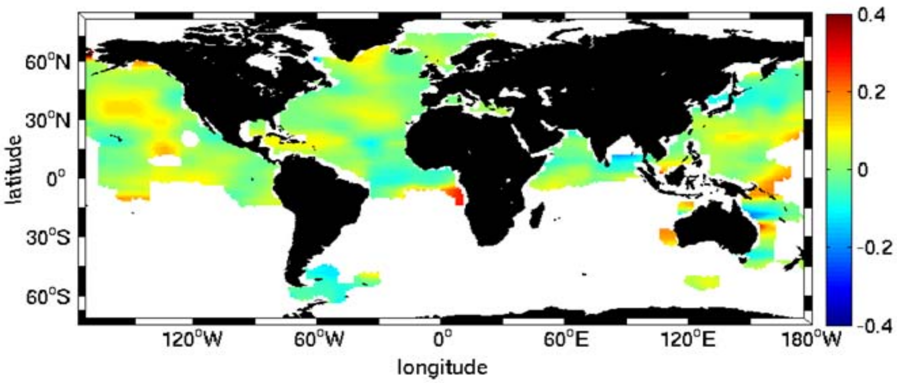
A



B



C



Residual average bias in XBT-CTD temperature at 300m between 1968 and 1985 (°C) for raw XBTs (A), globally corrected XBTs (B) and regionally corrected XBTs.

- We detect negative XBT biases in the West Pacific but we are uncertain about the origin of the problem.

→ Particular probes ?

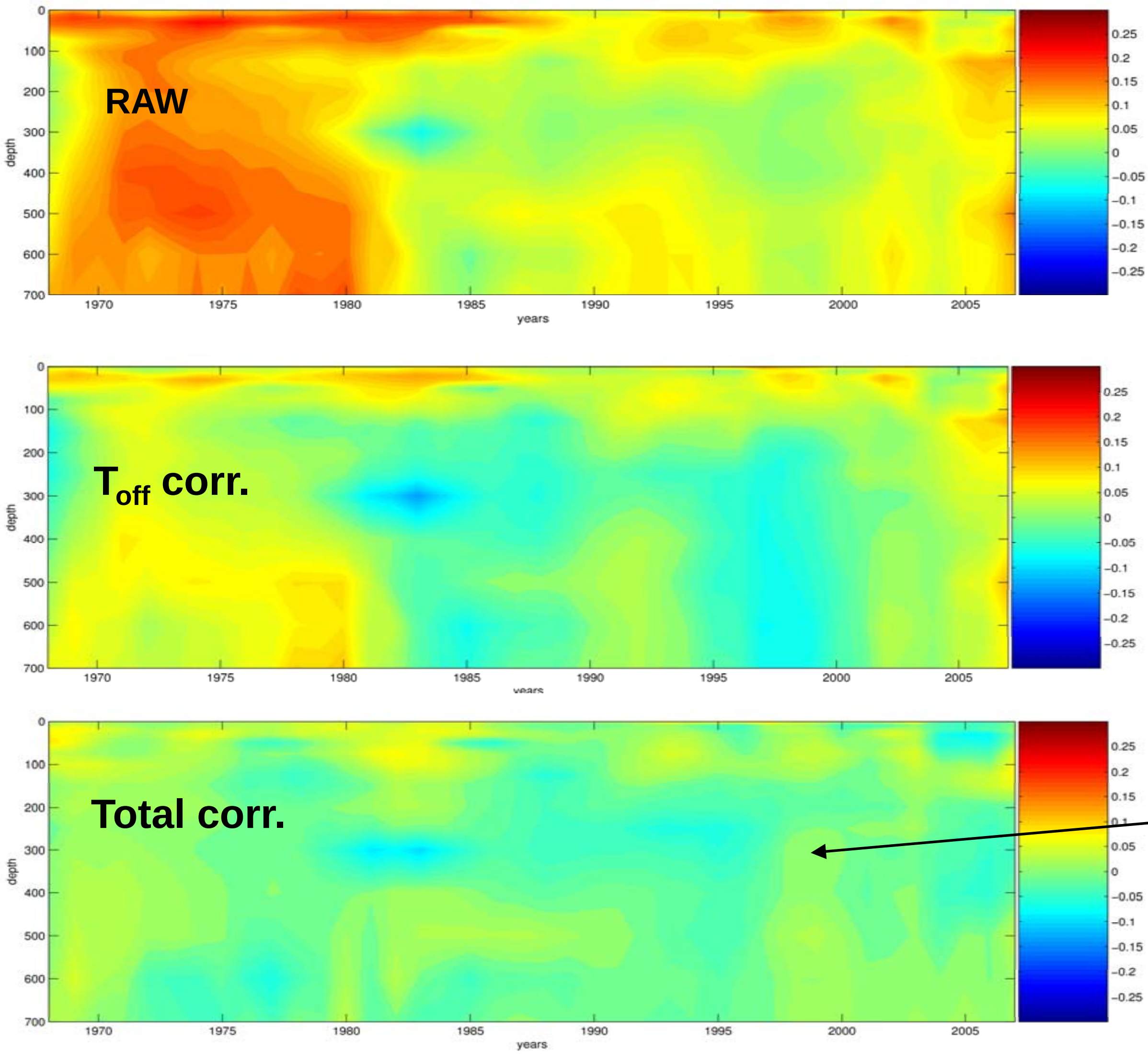
→ Collocation method ?

→ Interpolation at standard depth in WOD05?

Ifremer LOCEAN Coriolis OPERATIONAL OCEANOGRAPHY

The XBT problem

A new correction - Synthesis



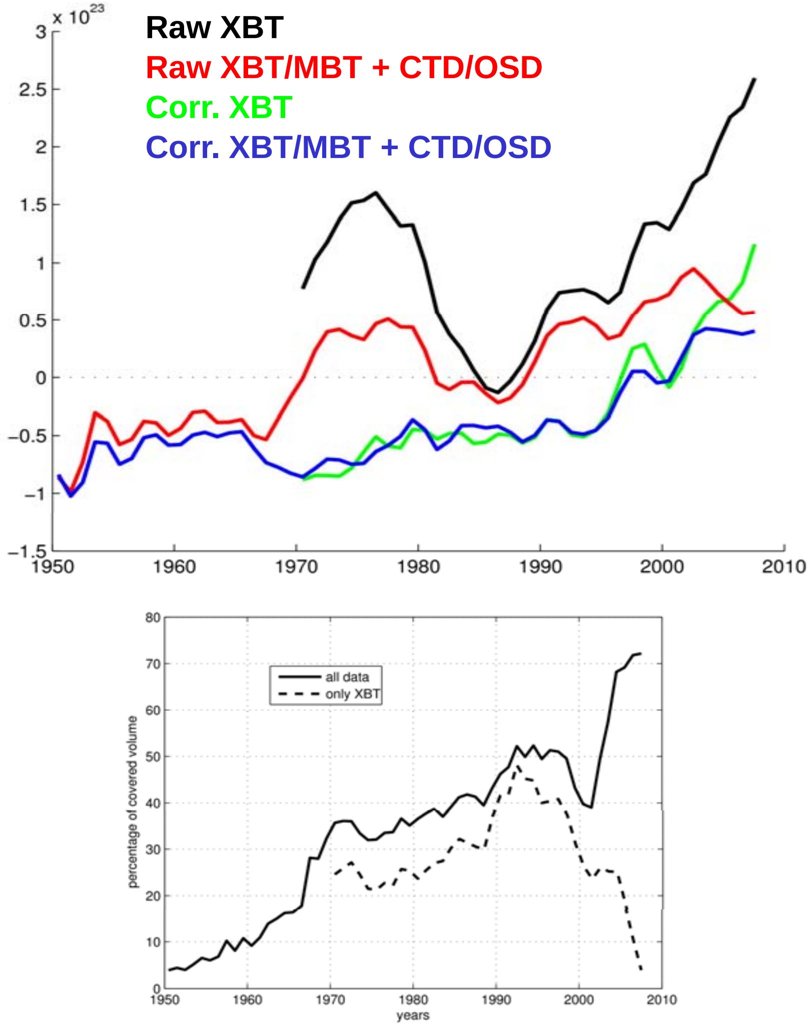
- The correction reduces the median temperature bias.
- Contribution of the offset is significant.

Optimal reduction of the global median bias.

Evolution of the median raw bias (top), corrected by the thermal offset (middle) and by a parabolic function added an offset (below) funtion of depth and years (°C).

The XBT problem

A new correction - Validation



Evolution of the 0-700m integrated OHC calculated from only WOD05 XBT (black), corrected XBT (green),all data from WOD05 (red) and all corrected data (blue) function of the time.

- We used an integrated value to validate our correction over the entire database.
- The calculation of the OHC confirms that on average XBT temperature data are now closer to CTD temperature data.
- Using the same methodology, we corrected MBT (second order correction and an offset, temperature classes).
- The maximum of heat content during the 70's can be explained by the XBT bias (Domingues et al, 2008, Ishii and Kimoto, 2009; Levitus et al, 2009; Wijffels et al, 2008; Gouretski and Reseghetti 2010).

EOF reconstruction

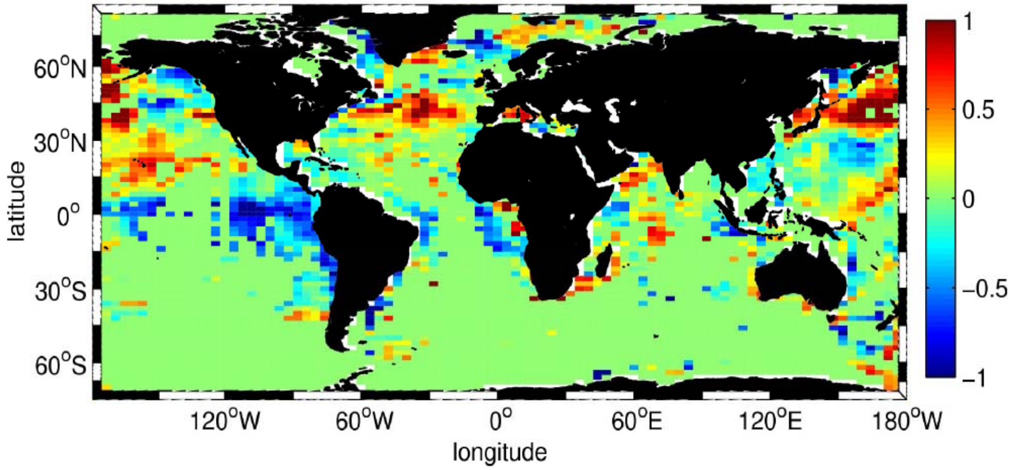
Method

- Method based on DINEOF (Beckers et al, 2003), interpolation process which had been developed to fill gappy satellite maps.
- Reconstruction period:1955-2009.
- The observations (WOD05+ARGO Coriolis) are mapped on a latitude and longitude grid (2°*4°) on standard levels (until 700m).
- Annual mean anomalies are obtained subtracting the WOA05 climatology .
- Iterative process to compute EOFs and fill the gaps.
  - 1. gappy boxes are filled with zeros.
  - 2. Calculation of the first EOF. We fill the gaps with the values of the field reconstructed with the first EOF.
  - 3. Repeat step 2 until we reach convergence of the first EOF.
  - […]
  - n. Calculation of the Nth EOF. We fill the gaps with the values of the field reconstructed with the Nth EOF.

EOF reconstruction

Method

SST



Example of field reconstruction using EOFs

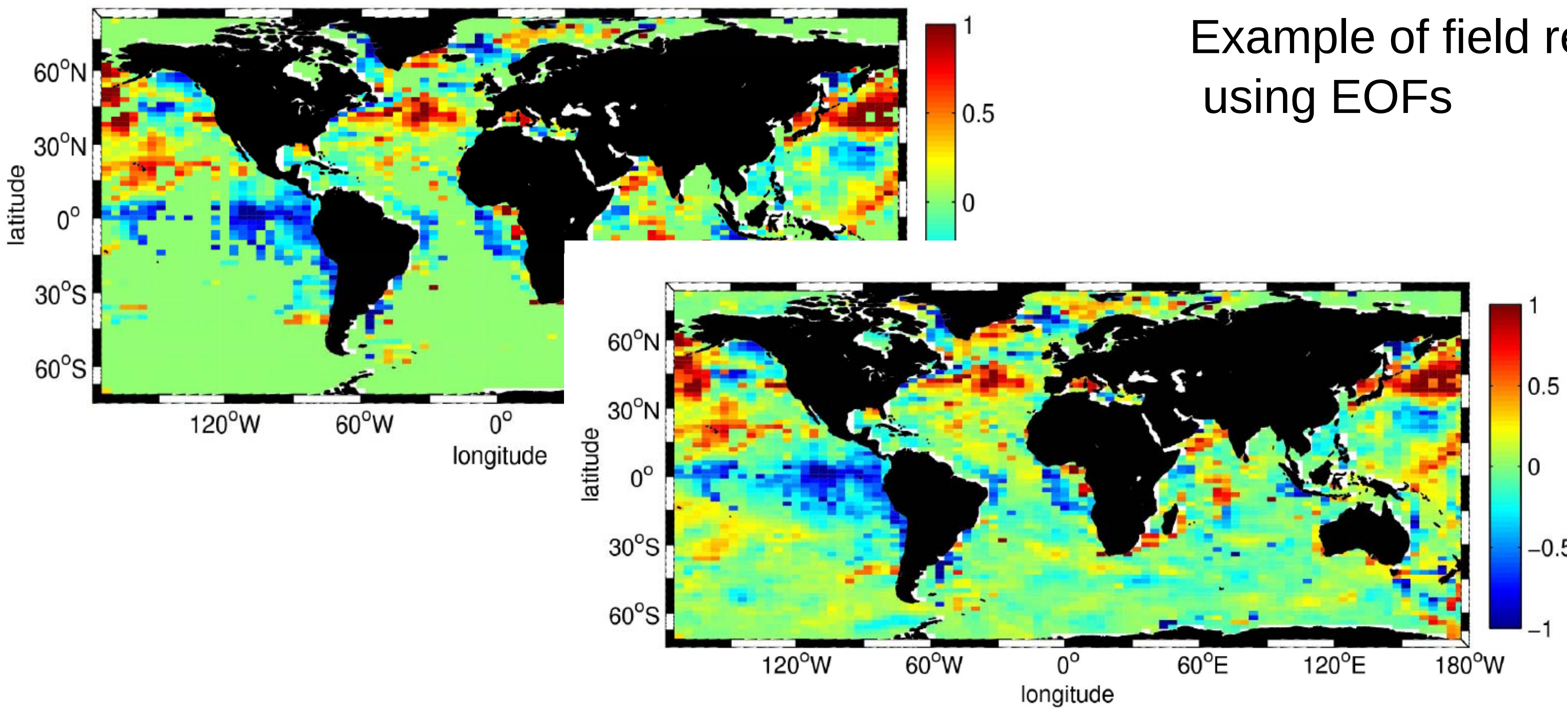
Steps of the SST reconstruction process. On the upper panel, the initial field (gaps filled with 0), in the middle One, gaps are filled with the result of 4 EOFs and on the lower panel , the final reconstruction with 20 EOFs.

Ifremer LOCEAN Coriolis OPERATIONAL OCEANOGRAPHY

EOF reconstruction

SST *Method*

Example of field reconstruction using EOFs

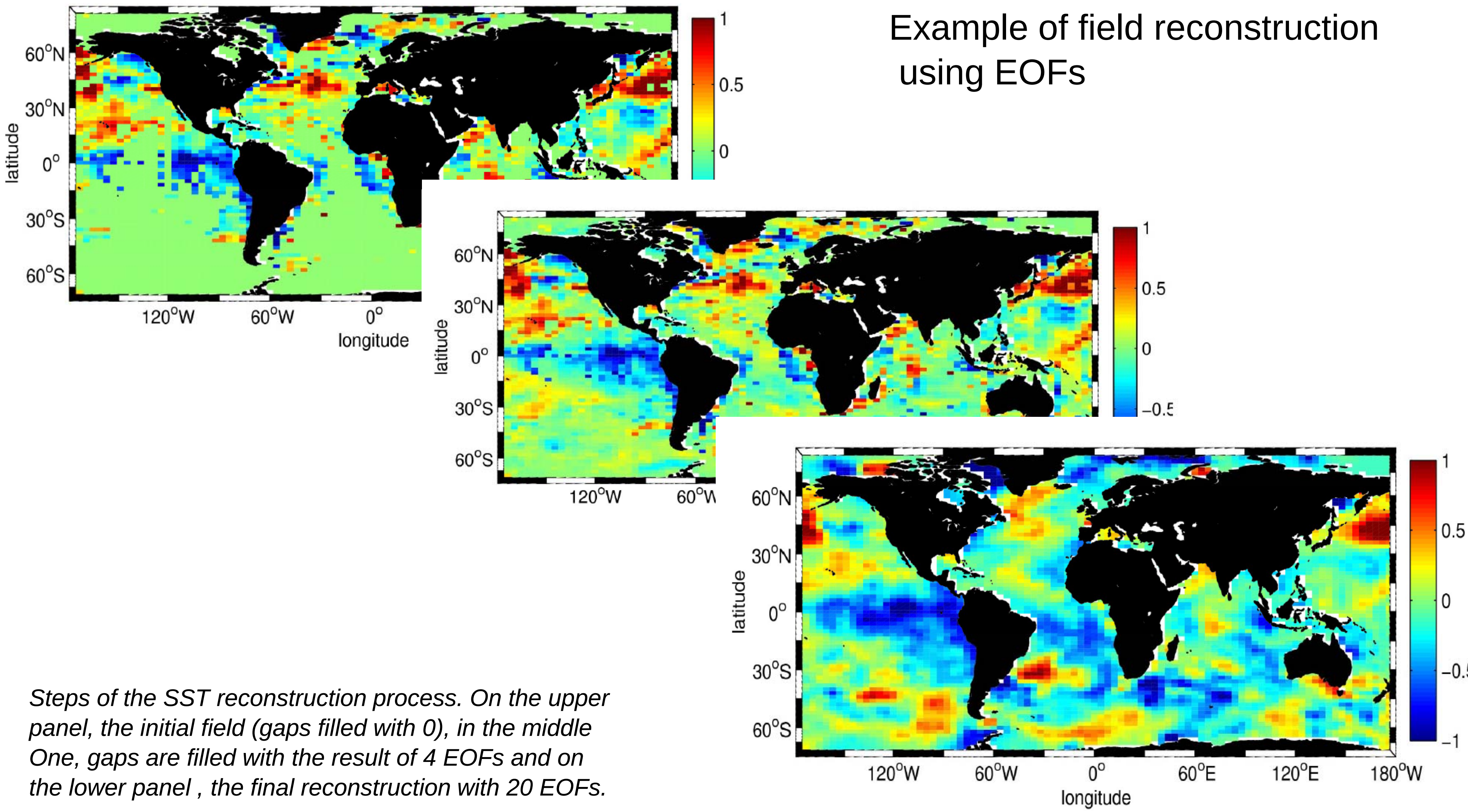


Steps of the SST reconstruction process. On the upper panel, the initial field (gaps filled with 0), in the middle One, gaps are filled with the result of 4 EOFs and on the lower panel , the final reconstruction with 20 EOFs.

EOF reconstruction

Method

SST



Example of field reconstruction using EOFs

Steps of the SST reconstruction process. On the upper panel, the initial field (gaps filled with 0), in the middle One, gaps are filled with the result of 4 EOFs and on the lower panel , the final reconstruction with 20 EOFs.

Ifremer LOCEAN Coriolis OPERATIONAL OCEANOGRAPHY

EOF reconstruction

Method

Application to the in situ dataset

- *How many EOFs to retain ?*

  We removed data randomly to the initial gappy fields. Then we compare with the interpoled values (cross validation).

  In DINEOF, the optimal number of EOFs is the one that minimizes the RMS error between the interpolated point and the « true » point.

- For in situ data, this criterion is not robust due to its noise. Applying this method, the number of EOFs is too small and the resulting field underestimates the variance.

- We prefer to maximize:

$$C(N_{eof}) = \frac{R_{Var}(N_{eof})}{\sigma^2(N_{eof})}$$

with

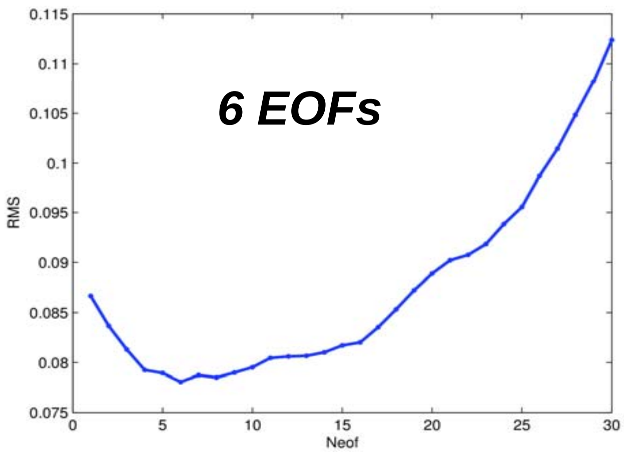
$$R_{Var}(N_{eof}) = \frac{Var_t(N_{eof})}{Var_f(N_{eof})}$$

EOF reconstruction

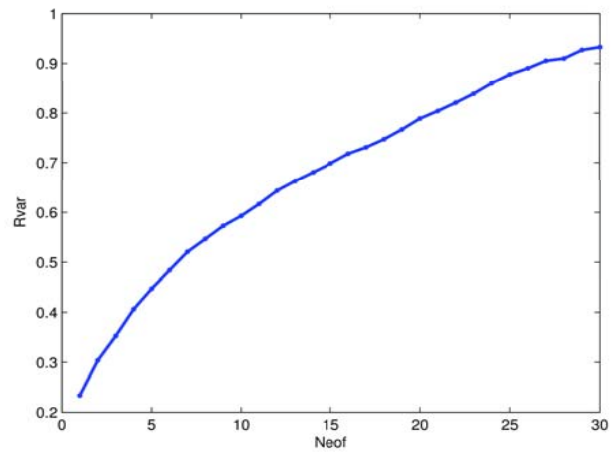
Method

<u>Application to the in situ dataset</u>

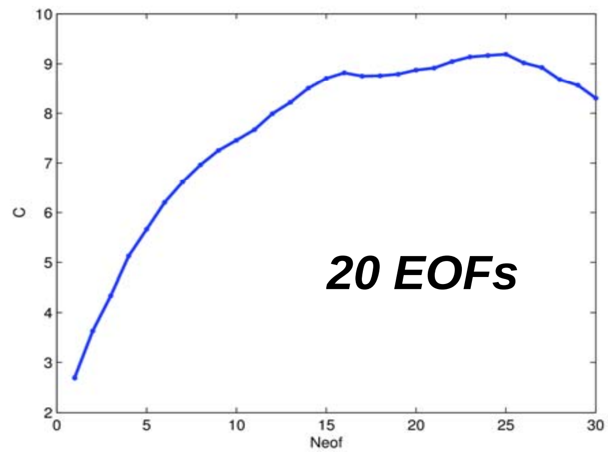
$$C(N_{eof}) = \frac{R_{Var}(N_{eof})}{\sigma^2(N_{eof})}$$



RMS(Neof)



Rvar(Neof)



C(Neof)

RMS error, variance ratio between true initial values and interpolated ones of cross validation boxes and parameter C function of the number of EOFs retained for filling gaps on the surface layer.

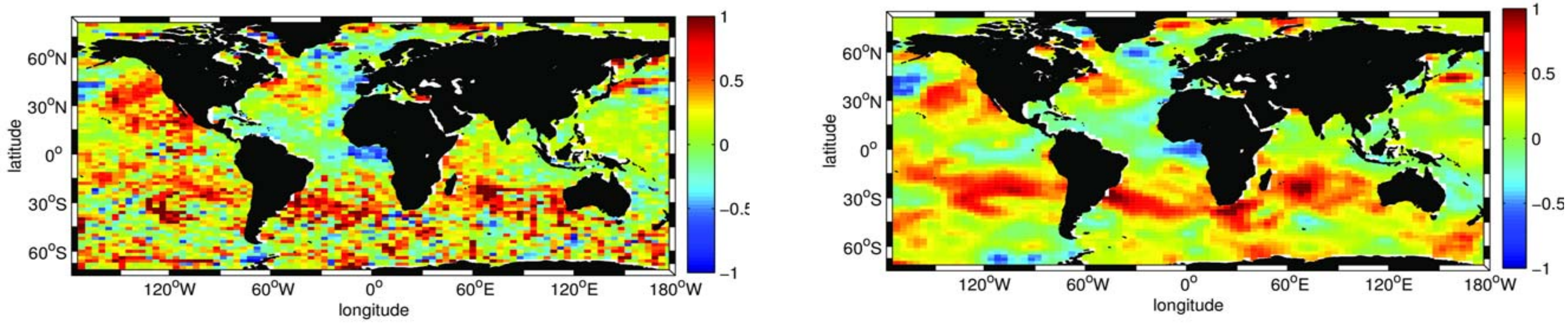
EOF reconstruction
Method

<u>Other improvements of the method...</u>

⅄ Because of the iterative character of the method, it is impossible to exactly compute the error field of the reconstruction.

⅄ Based on Beckers et al 2006, we perform a new method to estimate the error field applied to in situ data.

⅄ At each end of iteration, a smoothing is important.
→ spatial modes are more coherent (smoothing ~ add small-scale covariance) .



Example of field reconstruction without smoothing (left) and with smoothing (right) at each end of iteration for the year 1977 (ORCA).

EOF reconstruction

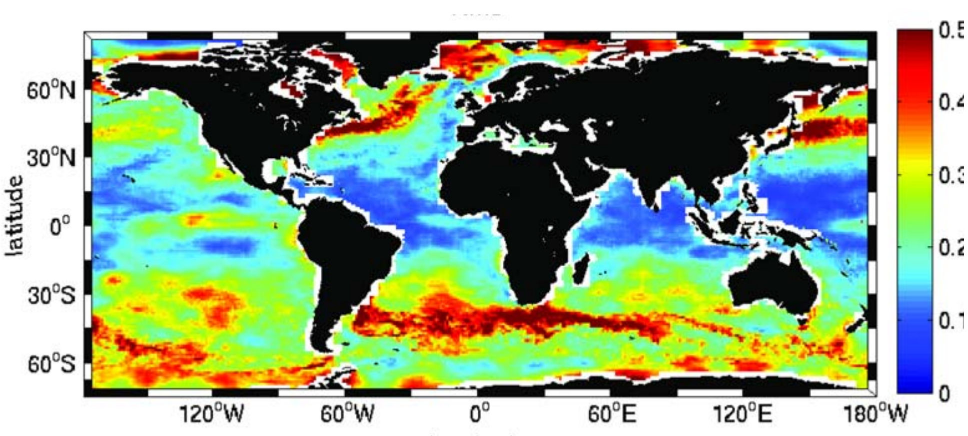
validation using a global model ORCA025 G70

- Test of the iterative method using subsampling of ORCA025 G70 fields (diagnostic model simulation of variability in 1958-2006 at intermediate resolution performed by the DRAKKAR project; LEGI-DRA-2-11-2006i).
- Model field values collocated with the EN3 dataset (1960-2006).
- Test of the impact of the quality/number of boxes on the interpolation robustness.
- In the southern ocean, using poorly sampled boxes result in increased signal variance but also higher RMS error .

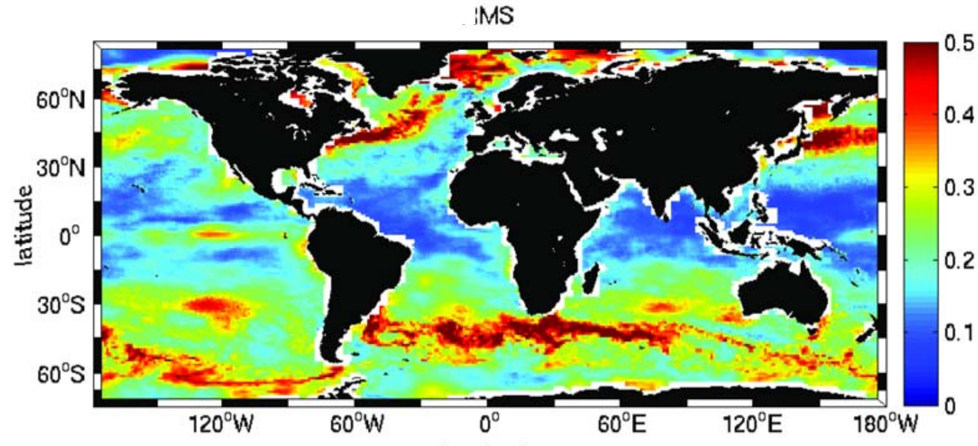
EOF reconstruction

validation using a global model ORCA025 G70

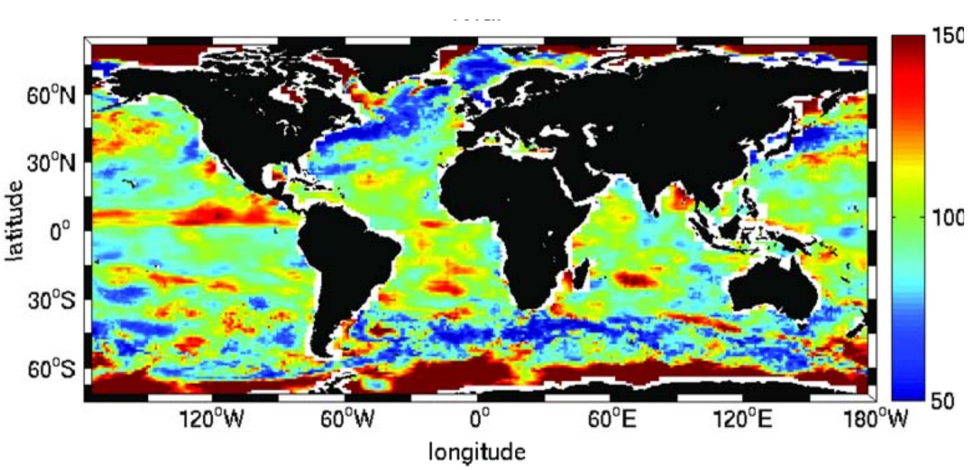
RMS



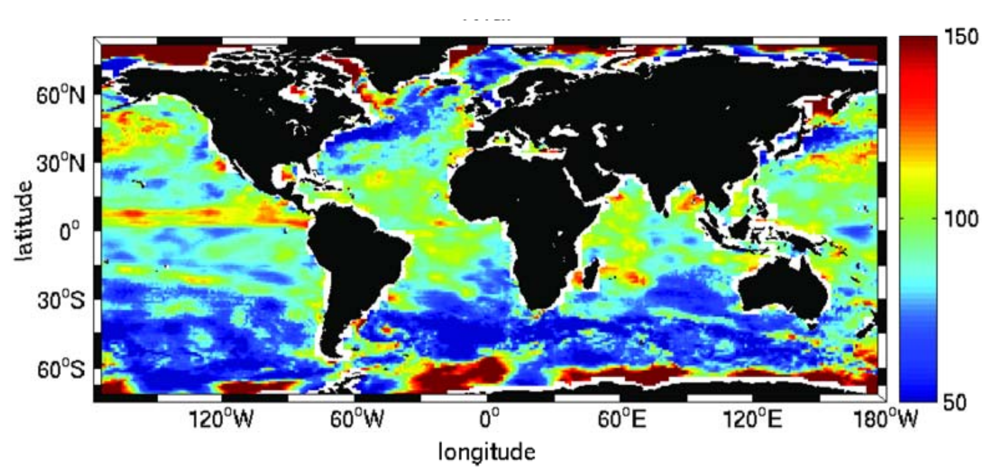
RMS



Rvar



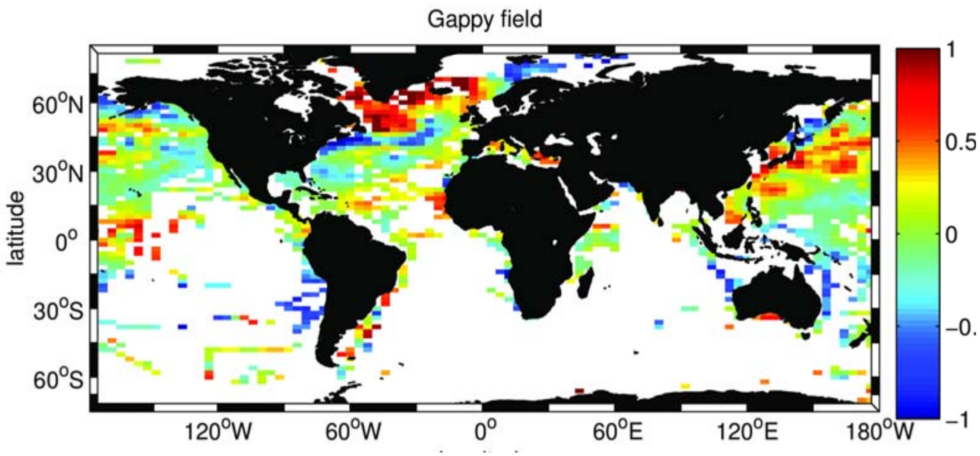
Rvar



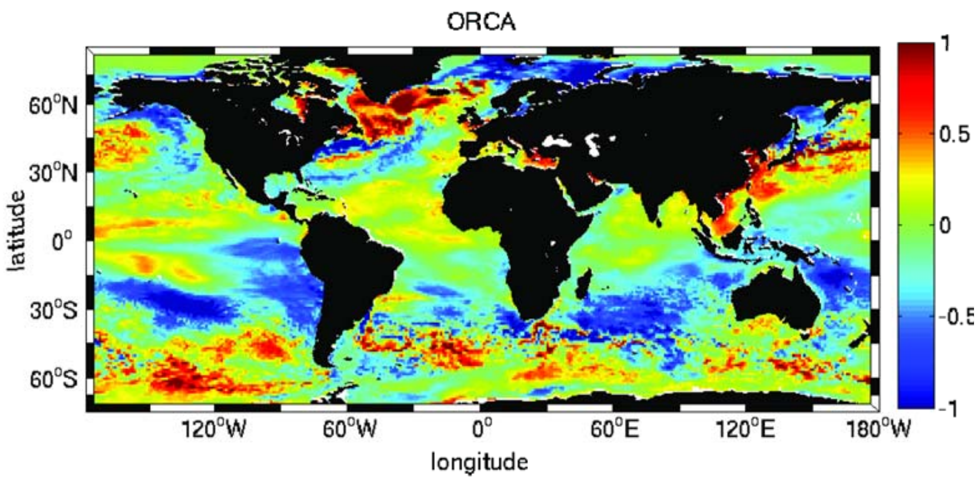
RMS (°C) and variance ratio (%) between the interpolated fields and ORCA fields at the surface. On the left the interpolated field has been reconstructed with all boxes and the right with only the robust boxes.

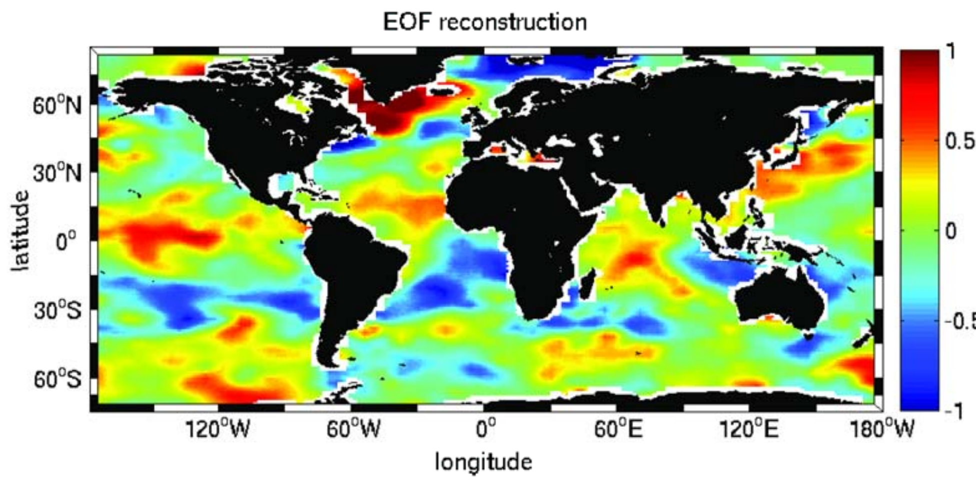
EOF reconstruction

ORCA025 - Example at 10m depth



The iterative process computes EOFs and fills large areas without data.

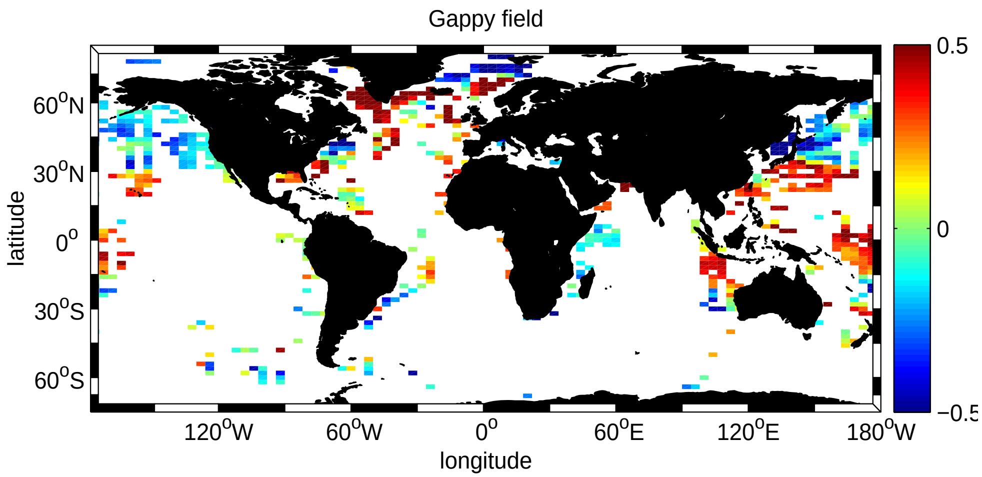




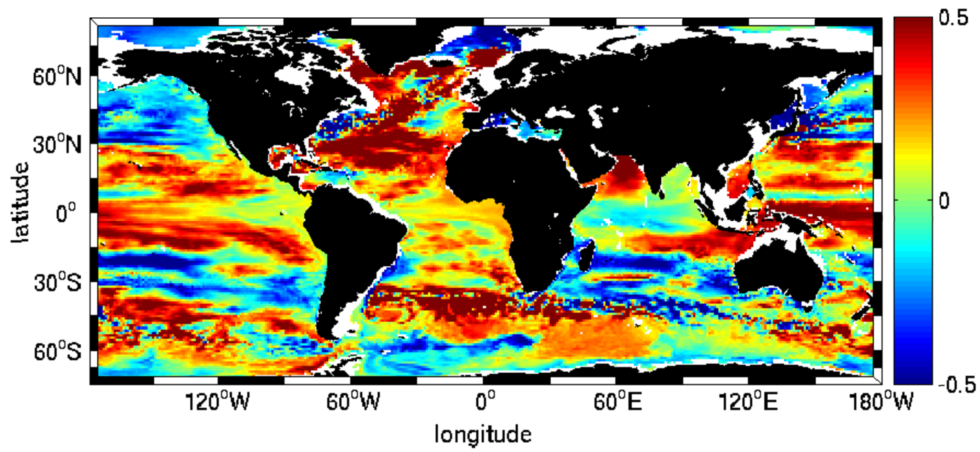
Example of reconstruction at 10m depth (1966). On the top, the initial gappy field, on the bottom right, the ORCA field and on the bottom left, the interpolated field.

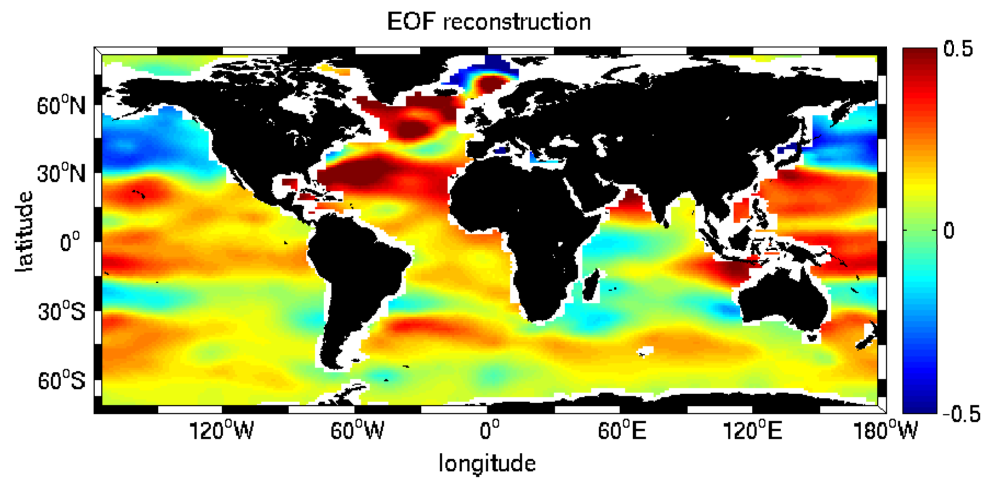
EOF reconstruction

ORCA025 - Example at 500m depth



The iterative process computes EOFs and fills large areas without data.



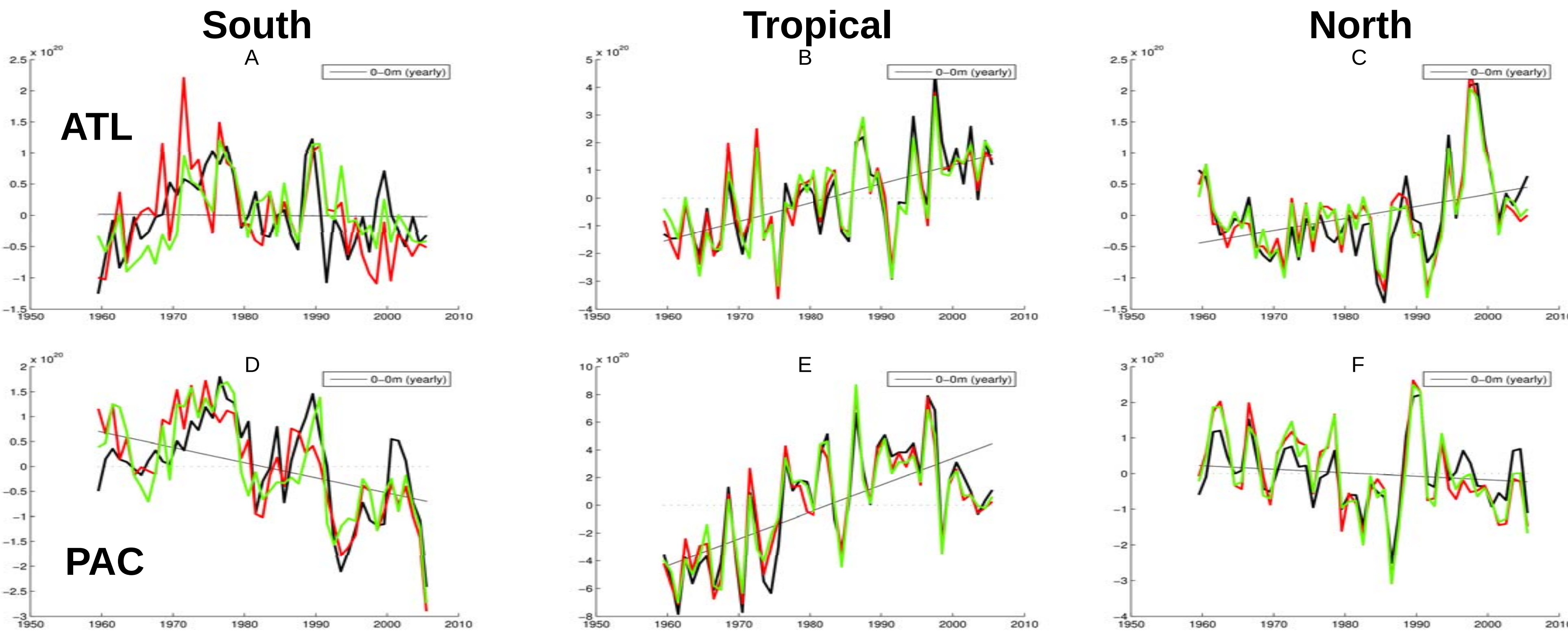


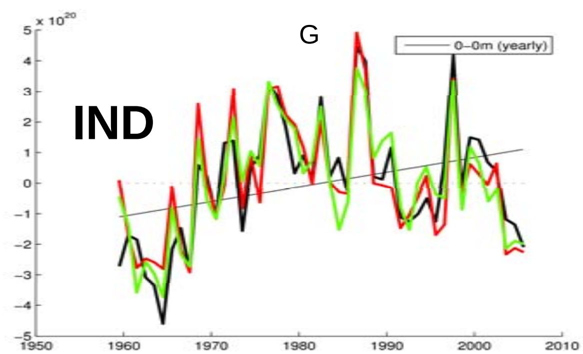
Example of reconstruction at 500m depth (1966). On the top, the initial gappy field, on the bottom right, the ORCA field and on the bottom left, the interpolated field.

EOF reconstruction

ORCA025 – Heat content at 10m depth

- Robustness of our interpolated fields for integrated values analysis?



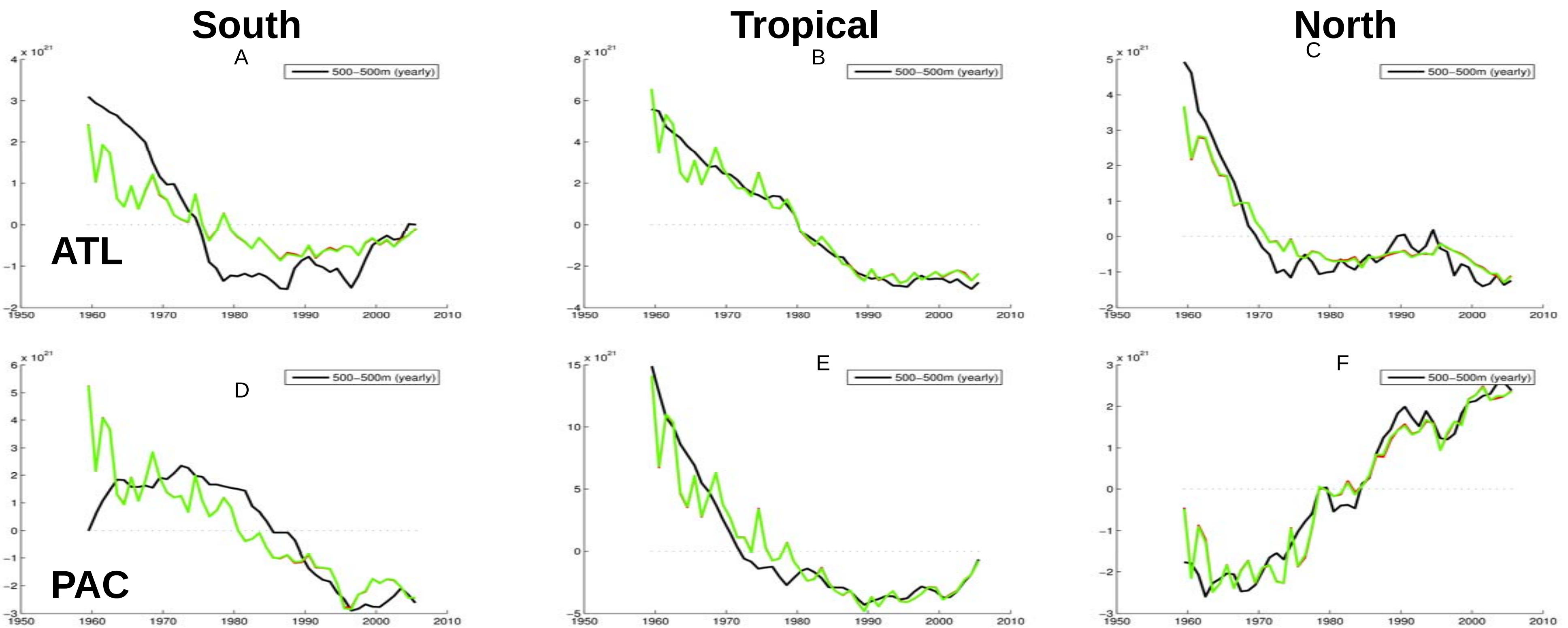


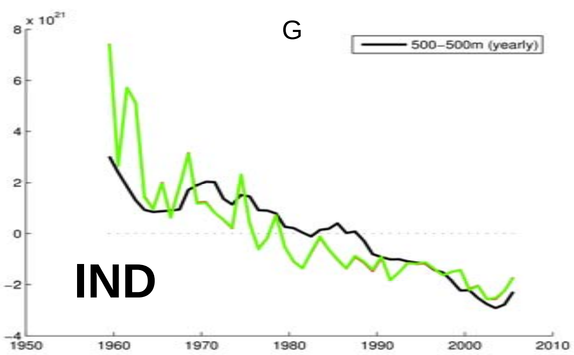
Heat content at 10m depth calculated with the ORCA fields (black) and the interpolated fields (all boxes [red] and robust boxes [green]). A,B,C are for the Atlantic basin (60°S/30°S, 30°S/30°N and 30°N/60°N). D,E,F are for the Pacific basin (same latitude bands) and G for the Indian basin (60°S/30°N).

EOF reconstruction

ORCA025 – Heat content at 500m depth

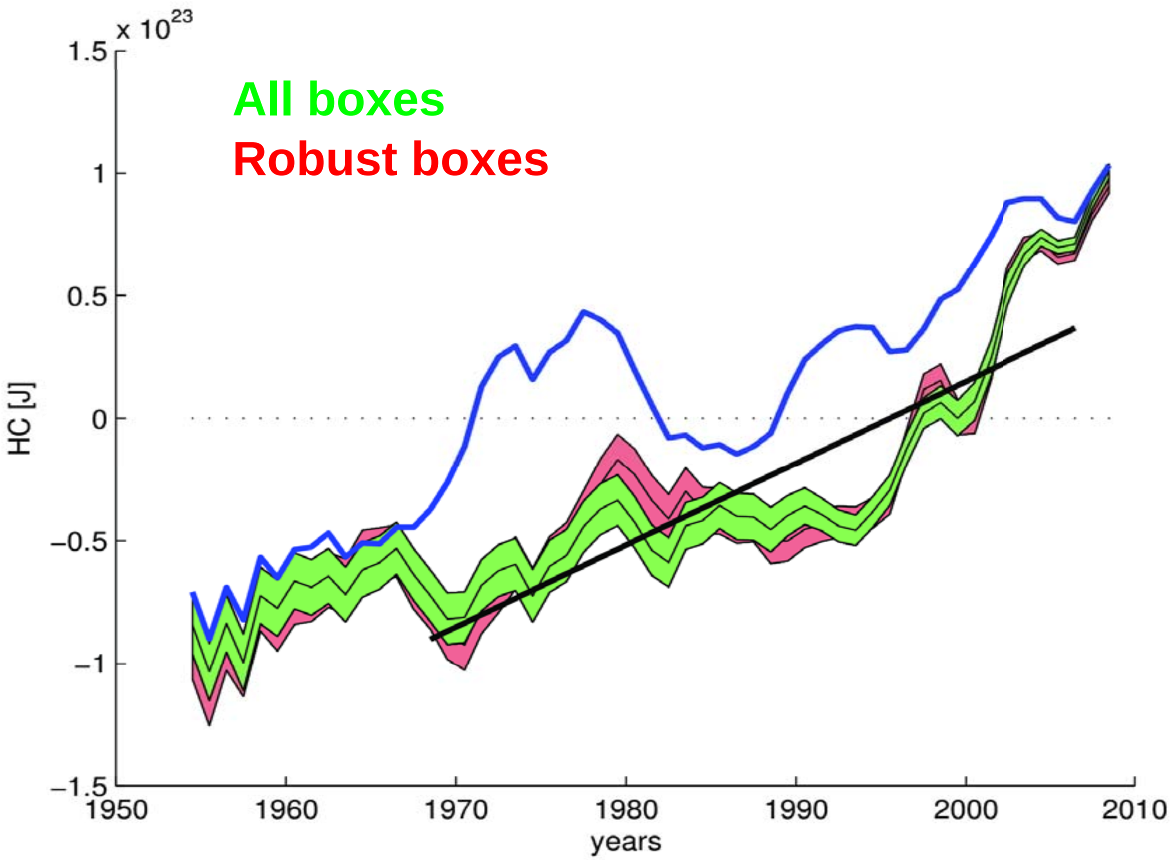
- Robustness of our interpolated fields for integrated values analysis?





Heat content between at 500m depth calculated with the ORCA fields (black) and the interpolated fields (all boxes [red] and robust boxes [green]).
A,B,C are for the Atlantic basin (60°S/30°S, 30°S/30°N and 30°N/60°N). D,E,F are for the Pacific basin (same latitude bands) and G for the Indian basin (60°S/30°N).

Results



Evolution of the 0-700m oceanic heat content calculated with interpolated fields. In blue, fields have been interpolated from raw data (all boxes), in red, fields have been interpolated from corrected data (robust boxes) and in green, fields have been interpolated from corrected data (all boxes). The black line is the linear trend of the evolution of this signal.

- We estimate a new linear trend for the period 1969-2007 :

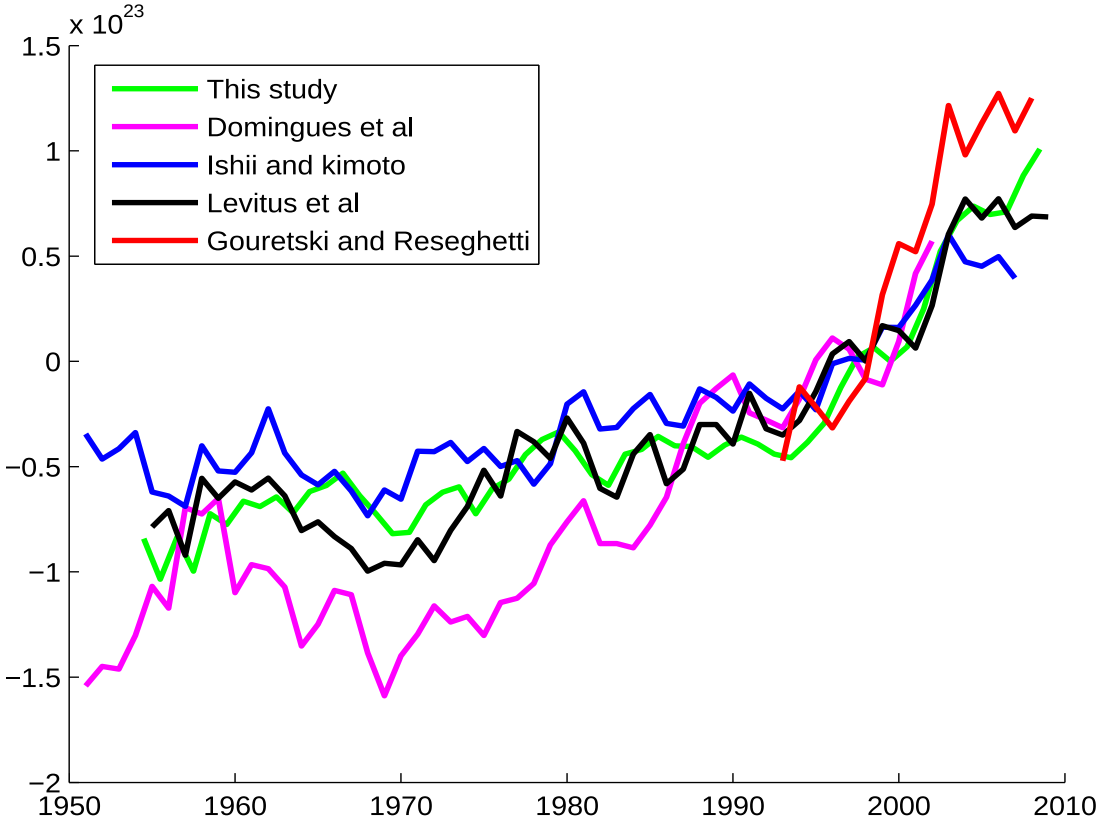
0.33e20 J/year

- Error bars of the red (robust boxes) and green (all boxes) curves overlap.

- Our OHC estimation is consistent with Domingues et al, 2008, Ishii and Kimoto, 2009; Levitus et al, 2009; Wijffels et al, 2008; Gouretski and Reseghetti 2010.

Ifremer LOCEAN Coriolis OPERATIONAL OCEANOGRAPHY

Results



Evolution of the 0-700m oceanic heat content.

- We estimate a new linear trend for the period 1969-2007 :

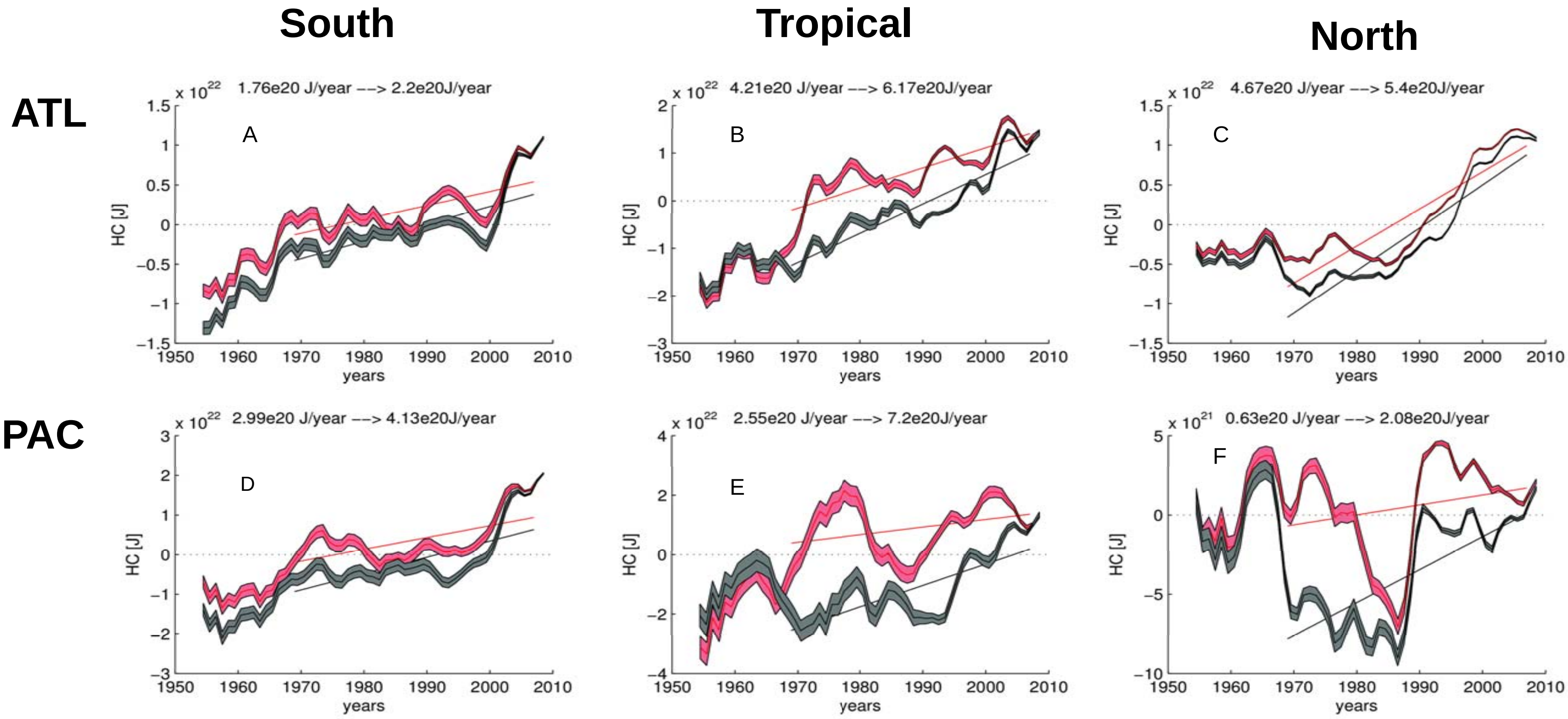
  0.33e20 J/year

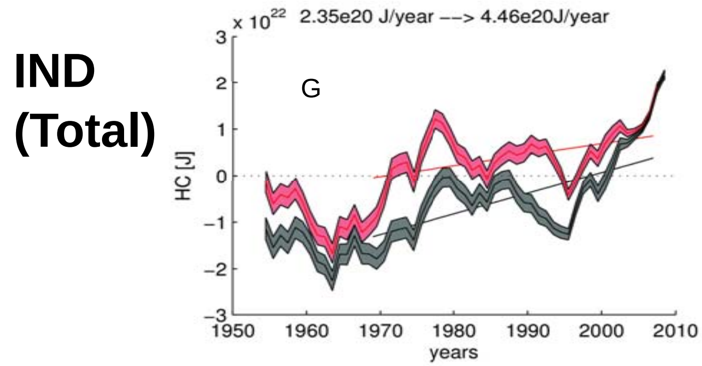
- Error bars of the red (robust boxes) and green (all boxes) curves overlap.

- Our OHC estimation is consistent with Domingues et al, 2008, Ishii and Kimoto, 2009; Levitus et al, 2009; Wijffels et al, 2008; Gouretski and Reseghetti 2010.

Ifremer LOCEAN Coriolis OPERATIONAL OCEANOGRAPHY

Results





0-700m integrated heat content calculated from the interpolated fields (raw data [red] and corrected data [black]).

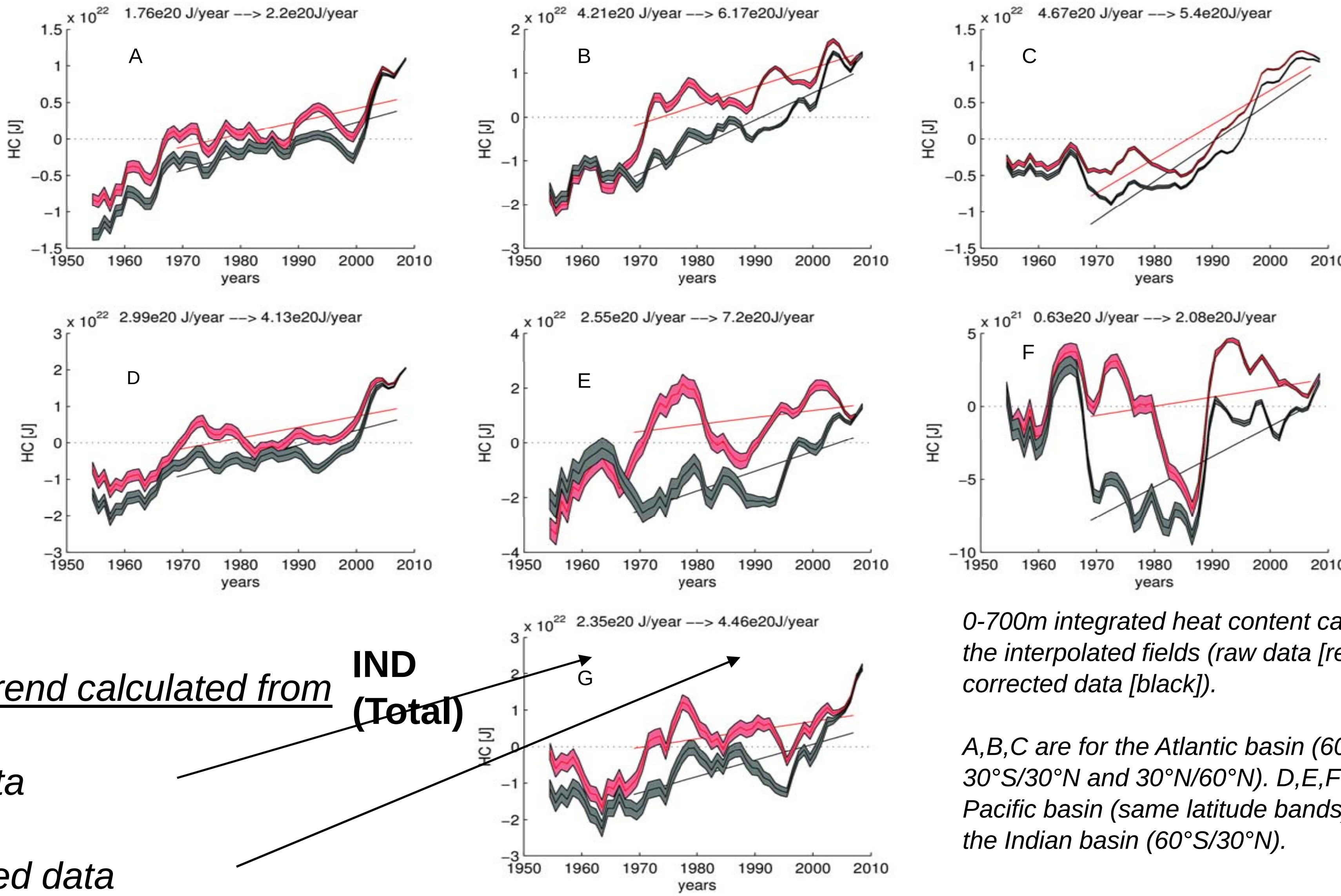
A,B,C are for the Atlantic basin (60°S/30°S, 30°S/30°N and 30°N/60°N). D,E,F are for the Pacific basin (same latitude bands) and G for the Indian basin (60°S/30°N).

Results

South **Tropical** **North**

ATL

PAC



Linear trend calculated from

Raw data

Corrected data

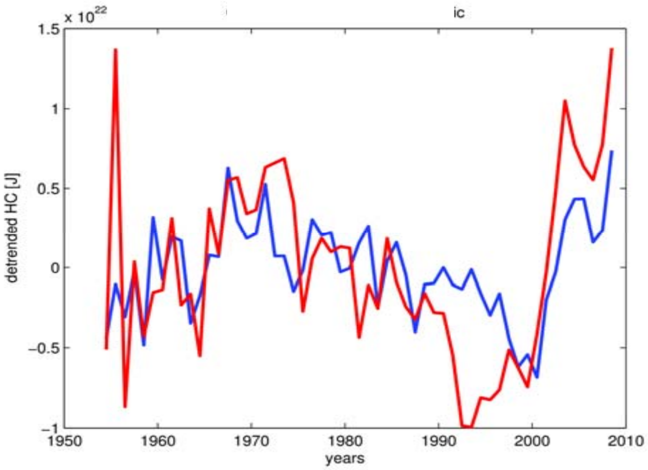
0-700m integrated heat content calculated from the interpolated fields (raw data [red] and corrected data [black]).

A,B,C are for the Atlantic basin (60°S/30°S, 30°S/30°N and 30°N/60°N). D,E,F are for the Pacific basin (same latitude bands) and G for the Indian basin (60°S/30°N).

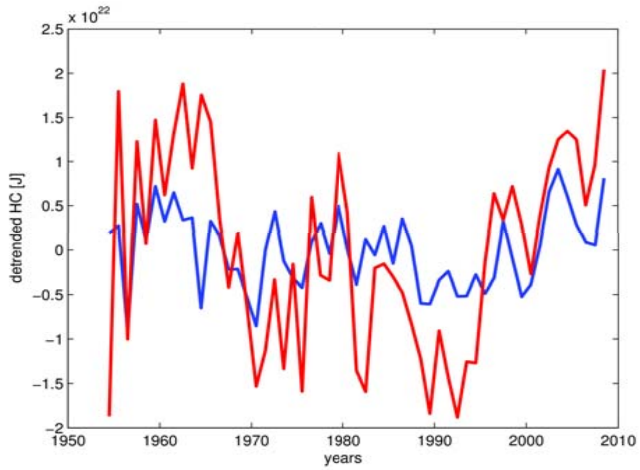
Results

regional heat content correlation

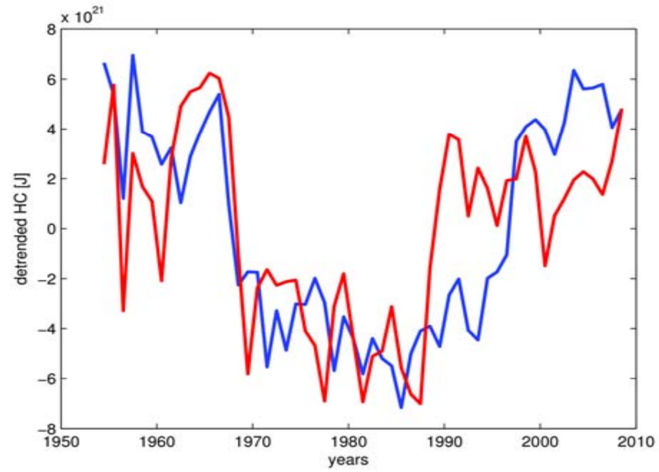
ATL PAC



55% → **66%**



61% → **63%**



45% → **70%**

Detrended heat content signal (0-700m) for Atlantic (blue) and Pacific (red) basins (South/Equatorial/North).

The correlation of the detrended heat content signal of Atlantic and Pacific basin increase after the XBT correction.

Conclusion

- According to W08, XBTs are subject to a depth bias varying with the year of deployment.
- However, our collocation method reveals that this bias should be better corrected with a second order function added to a thermal offset.
- Behavior of XBTS and XBTD are quite different and depends on the temperature of the sea water (Thadathil, 2002; GR10).
- All the parameters of our correction are time-dependant.
- Large residual biases induced us to treat separately XBTs launched in the West Pacific basin between 1968 and 1985 (also discussed in W08).
- The estimates are sensitive on the difference of bathymetry between collocated profiles.

Conclusion

• We have developed an interpolation tool using EOFs (based on DINEOF ; Beckers et al, 2003) to reconstruct temperature fields from in situ data.

• The tests with subsampling of ORCA imply that our OHC estimations are quite robust.

• We confirm that the maximum of heat content during the 70's can be explained by the XBT bias (Domingues et al, 2008, Ishii and Kimoto, 2009; Levitus et al, 2009; Wijffels et al, 2008; Gouretski and Reseghetti 2010).

• In addition, a linear trend of **0,33.10^{22} J/year** is apparent between 1969 and 2007 (consistent with the cited papers above).

•The linear trend of all basins increases due to the XBT correction.

• All correction provides similarities on OHC estimation but, what about the structure of the variability ?

Ifremer LOCEAN Coriolis OPERATIONAL OCEANOGRAPHY

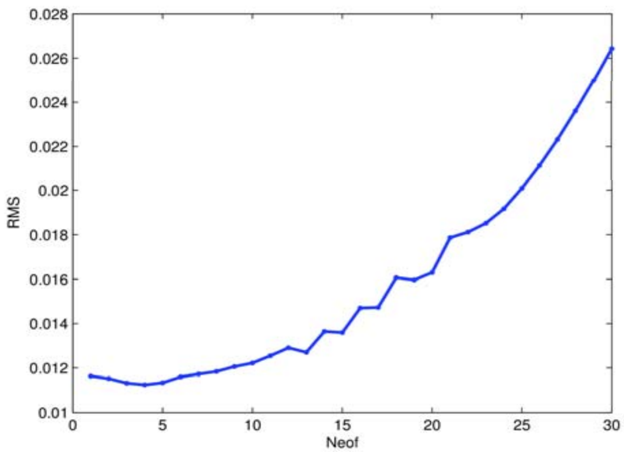
Thank You

Annexes

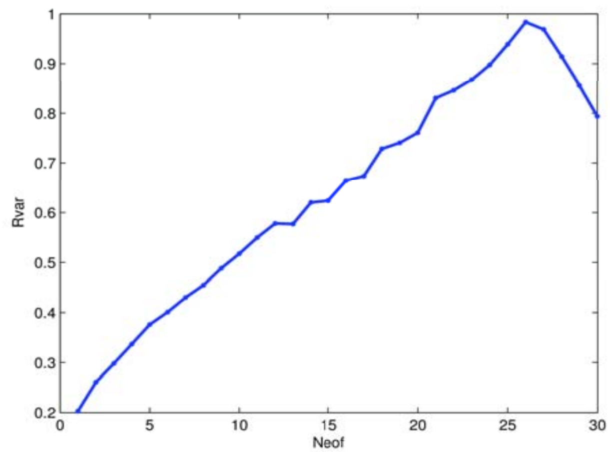
EOF reconstruction

Method

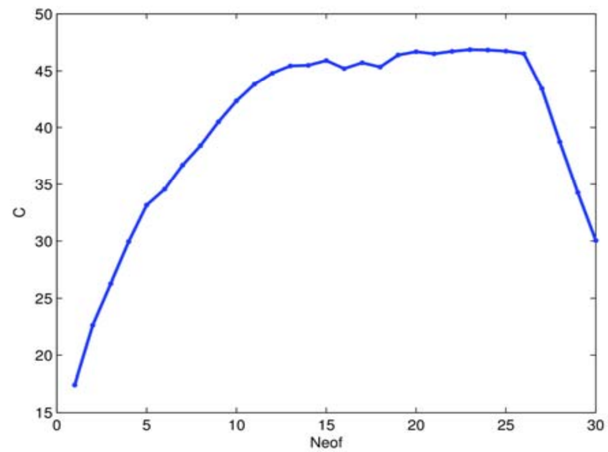
<u>Application to the in situ dataset</u>



RMS(Neof)



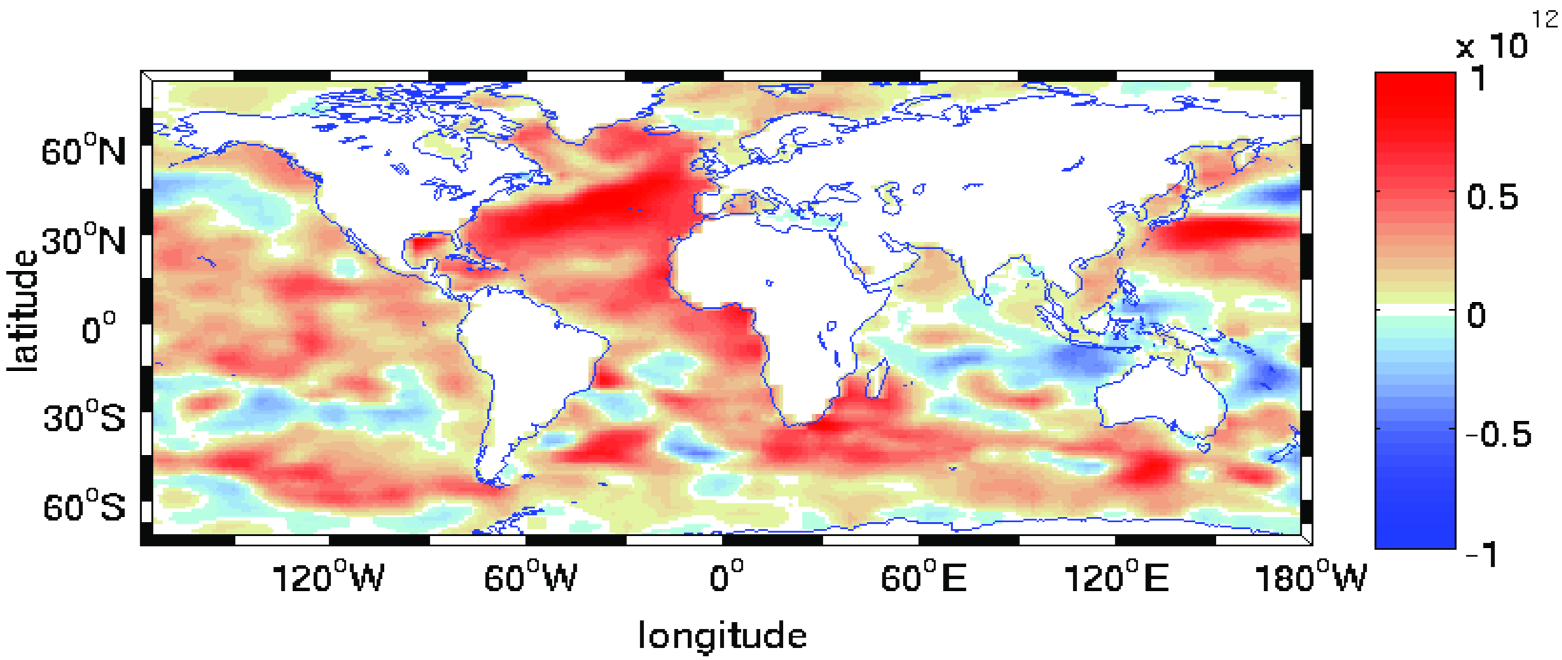
Rvar(Neof)



C(Neof)

RMS error, variance ratio between true initial values and interpolated ones of cross validation boxes and parameter C function of the number of EOFs retained for filling gaps at 700m depth.

Heat content trends



Linear trend of the heat content vertically integrated between the surface and 700m depth (J/year).