Reconstructing ocean variability using ARGO data and a data-driven method

E. Oulhen[1] ; N. Kolodziejcyk[1] ; P. Tandeo[2] ; B. Blanke[1] ; F. Sévellec[1]

[1] Laboratoire d'Océanographie Physique et Spatiale – IUEM, Université de Bretagne Occidentale, Plouzané, France

[2] IMT Atlantique, Lab Sticc, Plouzané, France

Motivation

- It is hard to disentangle anthropogenic and natural variability in the ocean.
- Reconstructing regional variability during the pre-ARGO period remains challenging due to sparse data coverage.
- Aim : Reanalyse pre-ARGO temperature observations with a data-driven method extracting information from the ARGO period.

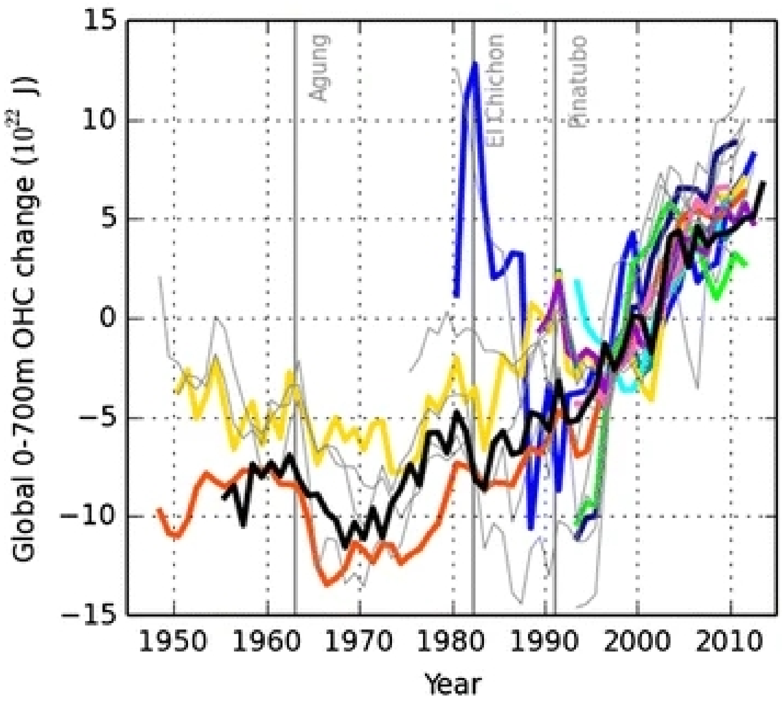


Figure 1 : Time series of the change in the integrated ocean global heat content, relative to a 1993–2007 baseline, for different analysis methods. From Palmer et al. (2017).

Data

- Temperature statistics from the ARGO period are extracted using ISAS 20 (Kolodziejcyk et al., 2021), a monthly optimal interpolation product method involving only in situ measurements.
- The analysis is done on quality-controlled in situ temperature observations from 1950 to 1999, from the EN.4.2.2 dataset (Good et al. 2013).
- Good et al. (2013) constructed their own optimal interpolation reanalysis using the same observations. This product, here named EN4, is used for comparison.
- The Niño 3.4 SST index is calculated from the HadISST1 product (Rayner et al., 2003).

Method : **Reduced Space** Data Assimilation Analogue Forecast (RedAnDA)

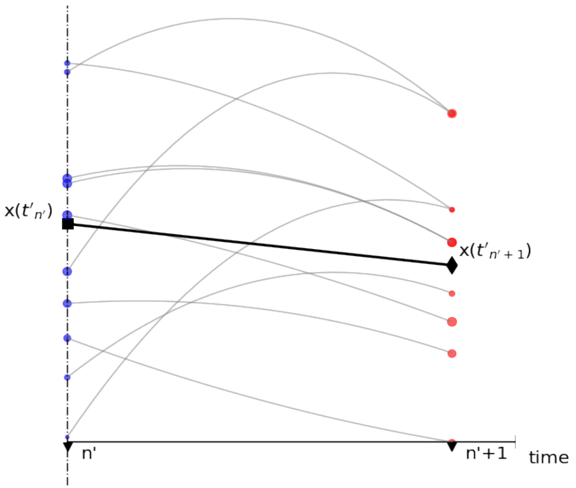
RedAnDA : Combination of reduced-space analysis (Kaplan et al., 1997) and ensemble analogue forecast data assimilation (Tandeo et al., 2014b ; Lguensat et al., 2017).

<u>step 1/2</u> : Potential temperature anomalies **T**(t) are computed for the ISAS product but also for the EN4 observations : $\mathbf{T}^{o}(t)$. Spatial empirical orthogonal functions **E** (EOFs) are constructed, with the associated time components **α**(t), using ISAS.

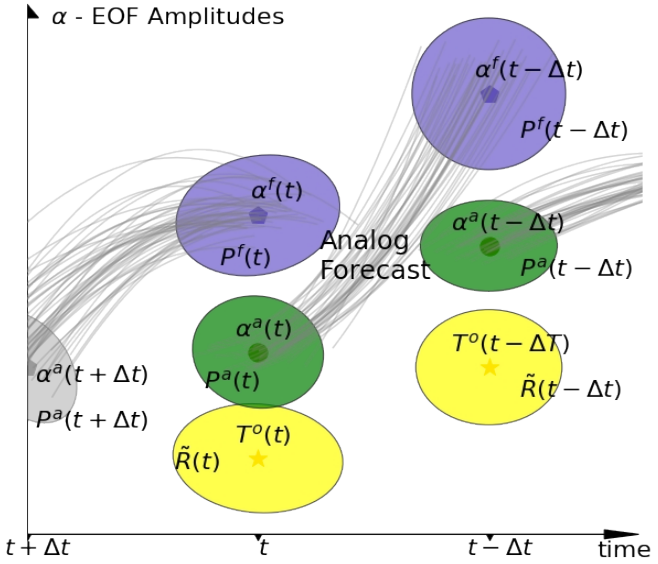
$$\begin{cases} \boldsymbol{T}(t) = \boldsymbol{E}\,\boldsymbol{\alpha}(t) + \boldsymbol{\epsilon}^{r}(t) \\ \mathbf{T}^{o}(t) = \mathbf{H}(t)(\mathbf{E}\,\boldsymbol{\alpha}(t) + \boldsymbol{\epsilon}^{r}) + \boldsymbol{\epsilon}^{o}(t) = \boldsymbol{\mathcal{H}}(t)\boldsymbol{\alpha}(t) + \tilde{\boldsymbol{\epsilon}}^{o}(t) \\ \boldsymbol{\alpha}(t) = \mathcal{A}\{\boldsymbol{\alpha}(t+\Delta t), \boldsymbol{\epsilon}^{f}(t)\} \end{cases}$$

Method : Reduced Space **Data Assimilation Analogue Forecast** (RedAnDA)

<u>step 2/2</u> : The goal is to recover **α**(t) between 1950 and 1999, by extracting dynamical information from the ARGO period. The analysis is an analogue forecast and a data assimilation scheme, as in figure 3.



<u>Figure 2</u> : Analogue forecast : For a given state (square), the closest analogues (blue dots) are selected from ISAS, with their successors (red dots). This selection allows to predict the next state (diamond), using a local linear regression. If the current state is persistent in time, the forecast tends to be accurate.



<u>Figure 3 :</u> Data assimilation : The different shapes represent the spread of uncertainty around the analogue forecast results (in blue). When combined with the observations (in yellow), they give the analysis (in green).

Spatial reconstruction and efficiency of the El Niño signal retrieval

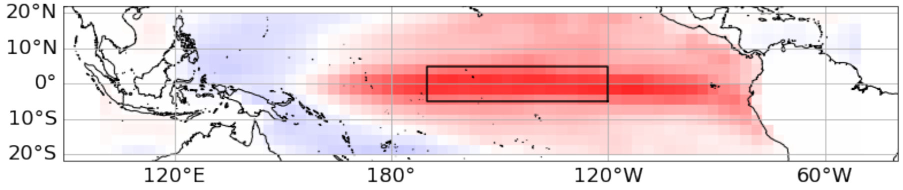


Figure 4 : First EOF, in the Pacific Ocean. The index is the average of SST in the enboxed region.

- The first reconstruction focuses on the Pacific basin.
- Figure 5 shows a satisfying reconstruction of the index.
- The EN4 95% confidence interval covers most of the signal, because of its sensitivity to the sparse sampling.
- The RedAnDA error's estimate is thinner. It carries information on both the sampling evolution and the dynamic statistics exported from the ARGO period.

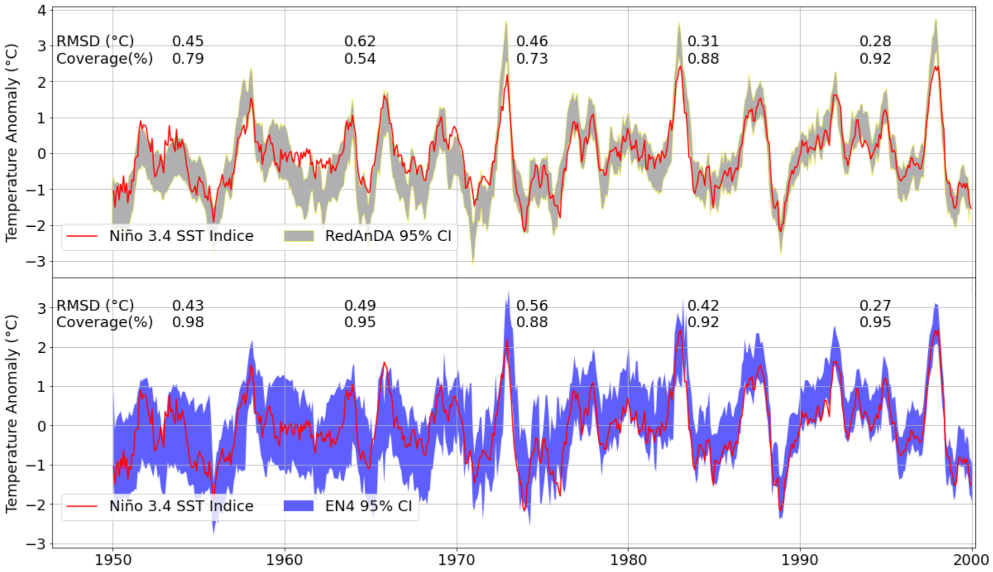


Figure 5 : Niño 3.4 index computed with RedAnDA (grey) and HadlSST1 (red) in the first panel, and for EN4 (blue) and HadlSST1 (red again). The decadal RMSD is calculated with the differences of each method with HadlSST1. Decadal Coverage is the percent presence of the HadlSST1 signal within confidence intervals

Conclusions :

- A new method of assimilation (RedAnDA) was presented using a new data-driven approach combining reduced-space analysis, analogue forecast and data assimilation.
- RedAnDA successfully reanalyses the upper Tropical Pacific Ocean temperature.
- The uncertainty estimate is improved with RedAnDA in comparison with the EN4 method
- However, the analogue forecast may be limited by the small size of the ISAS record

Perspectives :

- Further validation is still needed (in another oceanic basin, with cross validation).
- Sensitivity tests will be performed using numerical model outputs.
- The potential of the method for reanalysing salinity or other ocean variables is to be assessed.

References :

Good et al., (2013) https://doi.org/10.1002/2013JC009067 ; Kaplan et al., (1997) 10.1029/97JC01734 ; Kolodziejcyk et al., (2021) https://doi.org/10.17882/52367 ; Lguensat et al., (2017) 10.1175/MWR-D-16-0441.1 ; Palmer et al., (2018) 10.3389/fmars.2019.00416 ; Rayner et al., (2003) https://doi.org/10.1029/2002JD002670 ; Tandeo et al., (2014b) https://hal.archives-ouvertes.fr/hal-01188825