Bias Correction for reanalyses

Alison Fowler, Keith Haines, Ross Bannister, Sukun Cheng, Yumeng Chen (NCEO, University of Reading)
With thanks to colleagues from CMCC and CNR, Italy

Reanalyses of the Earth system are vital for monitoring climate change, supporting climate mitigation and adaptation policies, understanding extreme events, validating models, and training AI-based forecasting models. However, biases in the dynamical models used to generate multi-decadal reanalyses, along with changes in the observing systems available to constrain these models, can lead to erroneous artefacts in the reanalyses. This issue is particularly pronounced in the ocean, which has historically been under-observed. Each new observing system introduced (such as eXpendable BathyThermographs (XBTs) in the 1960s and altimetry data in the 1990s) has significantly increased the volume of assimilated observations.

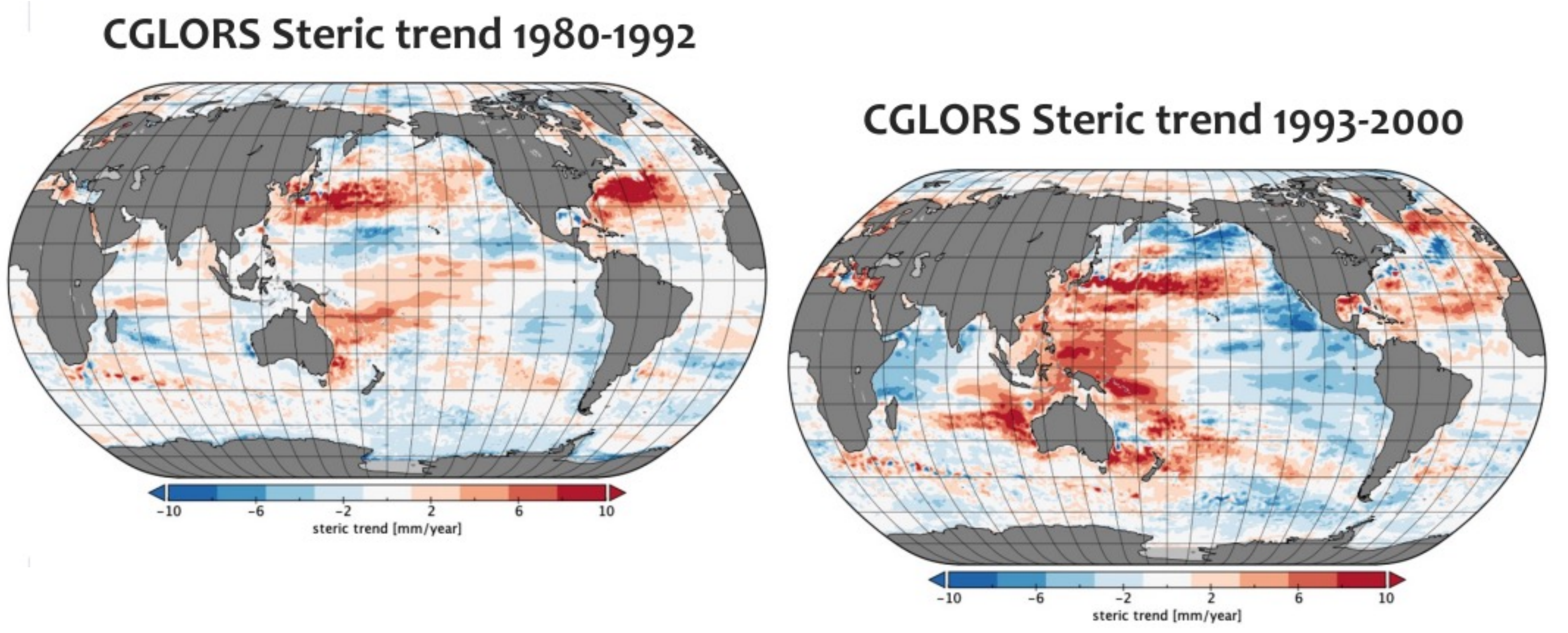

Figure: Change in Steric trend estimated from the CGLORS ocean reanalysis before and after the introduction of Satellite altimetry data in 1993. How can we know when trends are real?

Bias correction using a post-processing smoother

Biases in the model can be diagnosed from systematic analysis increments, but how do we translate this to a bias correction of the reanalysis?

A simple smoother was proposed by Dong et al. 2021 to remove unrealistic variability in the reanalysis time series, produced using filtering techniques, by incorporating information from future observations.

$$\mathbf{x}_\mathrm{a}^\mathrm{s}(t_i) = \mathbf{x}_\mathrm{a}(t_i) + \gamma\delta\mathbf{x}_\mathrm{a}(t_{i+1}) + \gamma^2\delta\mathbf{x}_\mathrm{a}(t_{i+2}) + \gamma^3\delta\mathbf{x}_\mathrm{a}(t_{i+3}) + \cdots,$$

where $\delta\mathbf{x}_\mathrm{a}(t_{i+j}) = \mathbf{x}_\mathrm{a}(t_{i+j}) - \mathbf{x}_\mathrm{b}(t_{i+j})$ are the analysis increments at time t_{i+j}.

How to choose the smoothing parameter, γ, so that the smoothed analysis has zero bias?

If we assume the observations are unbiased and the background bias, $\mathbf{b}$, is constant in time, then the bias in the smoothed analysis is

$$\mathbf{b}_\mathrm{a}^\mathrm{s} = \mathbf{b} - (1 + \gamma + \gamma^2 + \cdots)\mathbf{KHb} \approx \left(\mathbf{I} - \frac{1}{1-\gamma}\mathbf{KH}\right)\mathbf{b},$$

where $\mathbf{K}$ and $\mathbf{H}$ are the Kalman gain and observation operator matrices.

If we further assume the background bias is homogenous then

$$\mathbf{b}_\mathrm{a}^\mathrm{s} \approx 0 \text{ when } \gamma = \frac{1}{n}trace(\mathbf{I} - \mathbf{KH}) = 1 - \frac{1}{n}DFS$$

The DFS (Degrees of Freedom for Signal) is a metric of the observation influence on the analysis and can be computed in any reanalysis system (Fowler et al. 2020). It is function of the number and precision of the assimilated observations and the relative precision of the background information.

This choice of γ, therefore, results in a bias-correction method that is aware of the changing observation constraint.

Illustration using the 40-variable Lorenz 96 model with EnKF

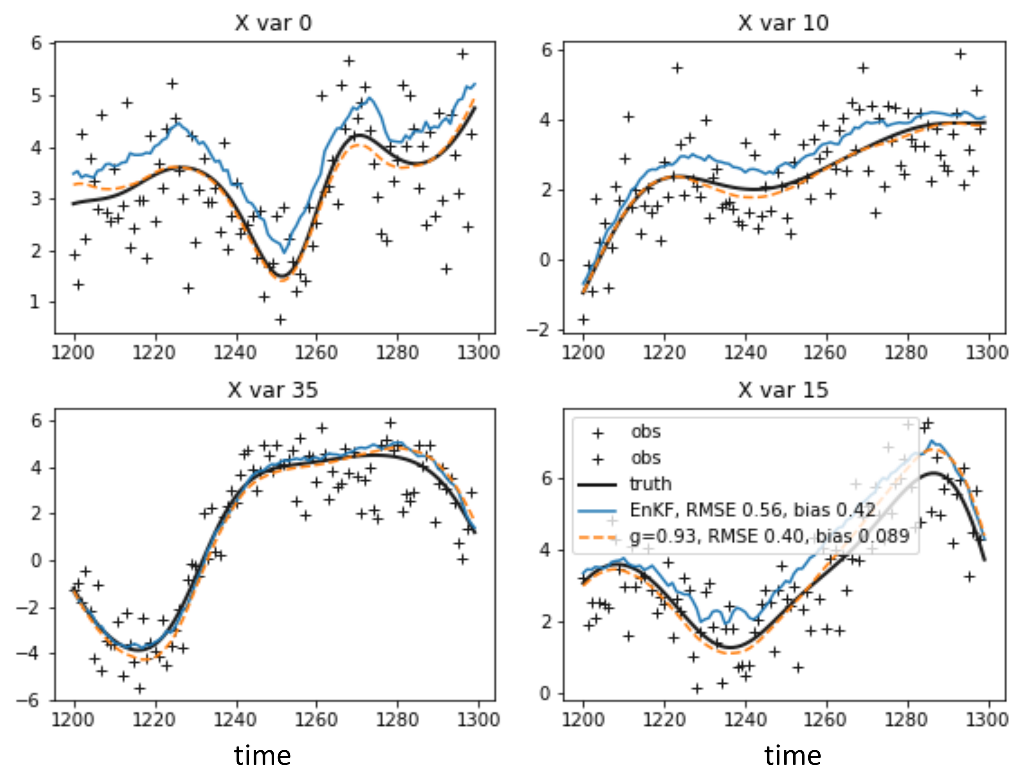

Figure: Illustration of the smoother applied to analysis time series for 4 variables. The EnKF assimilates observations using a model with too large forcing, resulting in a positive bias of 0.42. The bias in the smoothed analysis using $\gamma = 1 - \frac{1}{n}DFS$ is reduced by almost 80%.

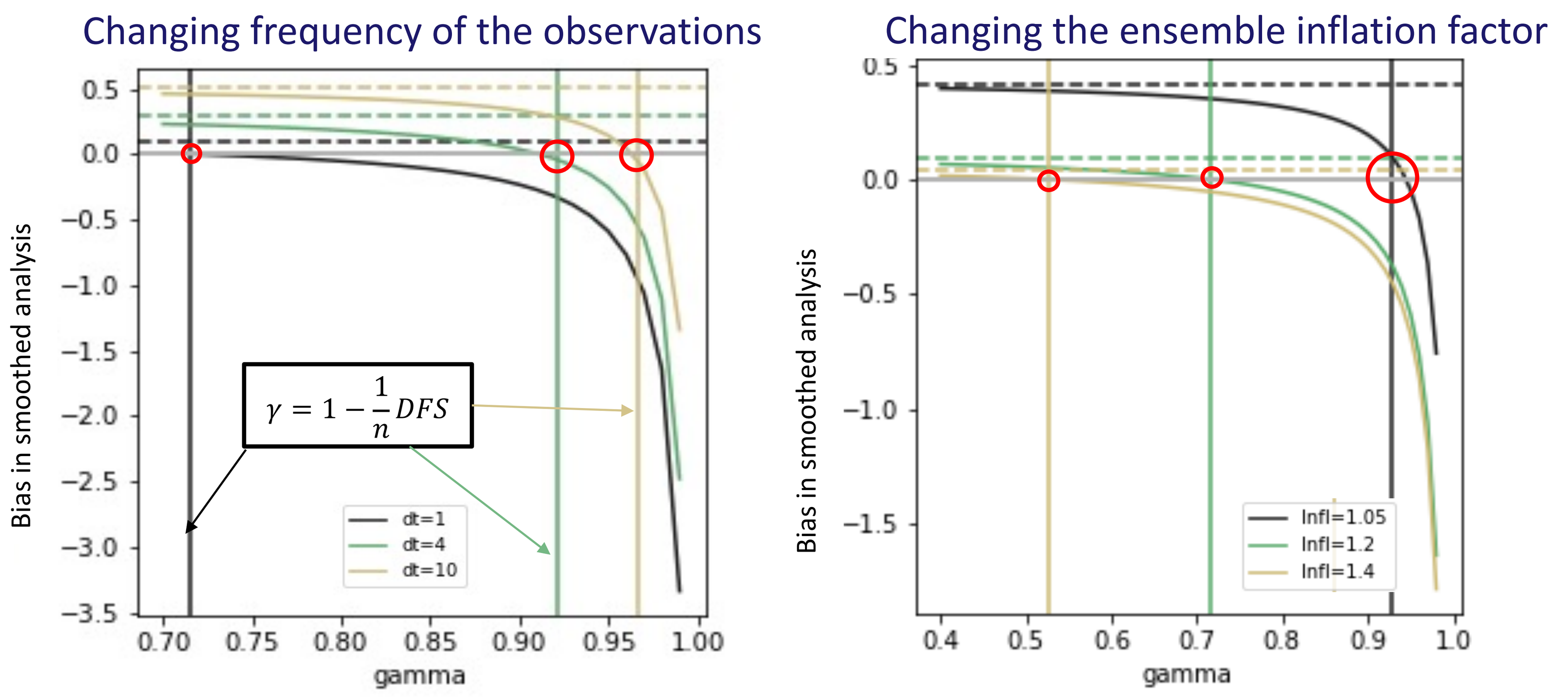

Figure: For each of the 6 experiments, the bias in the smoothed reanalysis is highly sensitive to the choice of γ; however, using DFS to define γ results in a reanalysis with nearly zero bias.

Summary and future work

- Reanalyses are often treated as the truth, but can have misleading artefacts due to changing observing systems constraints on model biases
- This work shows how the use of the 'simple smoother' can be used to correct for biases in the reanaylsis in a way that adapts to the changing observing system.
- Ongoing work is to apply the bias correction to ocean reanalysis and assess the impact on trends spanning the introduction of new observing systems.

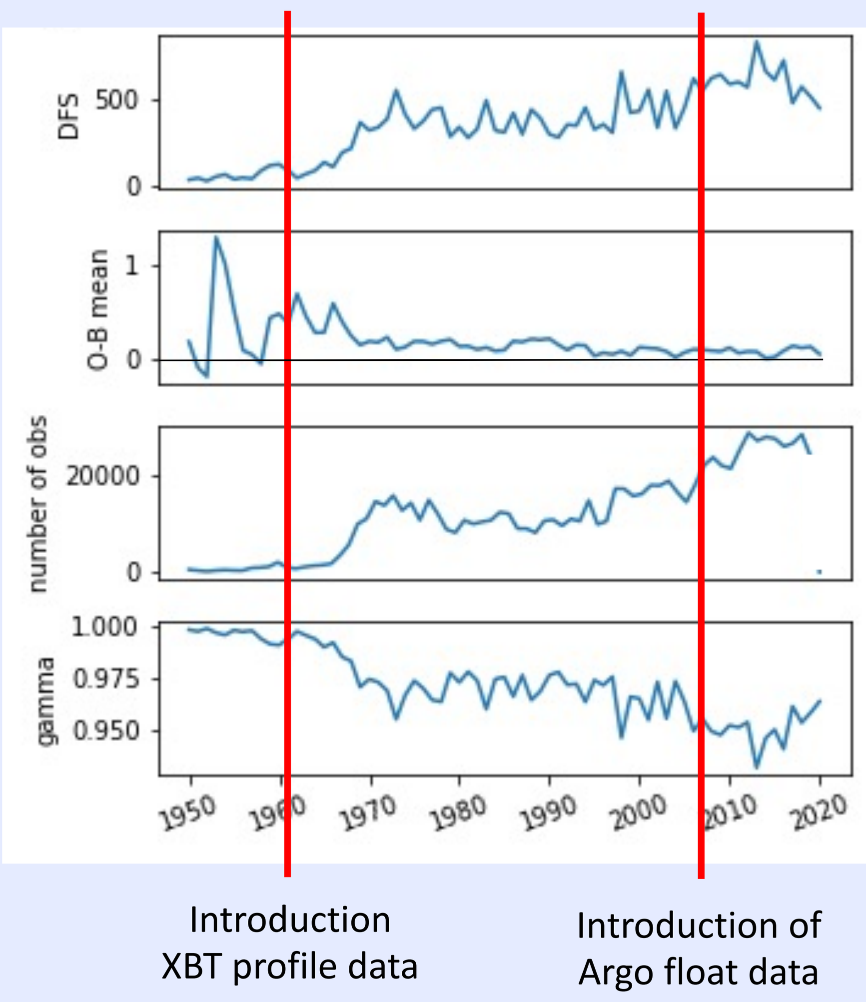

Figure: Application of bias-correction theory to the CIGAR reanalysis of temperature in the North Atlantic at depths of 300-800m for the period 1950-2020. From Top to Bottom: Estimate of degrees of freedom for signal, DFS; mean O-B as an indicator of bias in the reanalysis; number of observations; and the proposed smoothing parameter that will lead to a bias-corrected reanalysis.

References Dong et al. 2021: Improved high resolution ocean reanalyses using a simple smoother algorithm. *JAMES*
Fowler et al. 2020: Measuring theoretical and actual observation influence in the Met Office UKV: Application to Doppler radial winds. *GRL*

Contact information
a.m.fowler@reading.ac.uk