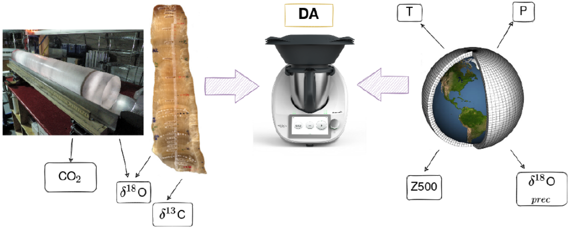
Paleoclimate Data Assimilation for the LM with $\delta^{18}O$

Extended Update

Mathurin Choblet
mchoblet@iup.uni-heidelberg.de

March 31, 2022



Paleoclimate DA Recipe

Following the Last Millennium Reanalysis Project [Hakim et al., 2016]

1. Get **proxy values**, their uncertainties and locations throughout time (can also be pseudoproxies)
2. Prepare **model prior**: Yearly averaged variables of interest over time for every gridbox. Sample static background ensemble (Offline DA)
3. Calculate **observations from model** (Proxy-System-Model, PSM)
4. For each year apply **Kalman Filter Analysis equations**
5. Evaluate results

→ LMR code publicly available, but a bit cumbersome and slow . . .

Center Piece: The Ensemble Kalman Filter equations

Kalman Filter equations

$$\mathrm{X}^a = \mathrm{X}^b + \mathrm{K}(z - \mathrm{H}x^b) \tag{1}$$

$$\mathrm{K} = \mathrm{P}^b H^T(\mathrm{H}P^b H^T + \mathrm{R})^{-1} \tag{2}$$

$$\mathrm{P}^a = (\mathrm{I} - \mathrm{K}H)\mathrm{P}^b \tag{3}$$

Center Piece: The Ensemble Kalman Filter equations

Kalman Filter equations

$$\mathrm{X}^{a} = \mathrm{X}^{b} + \mathrm{K}(z - \mathrm{H}x^{b}) \tag{1}$$

$$\mathrm{K} = \mathrm{P}^{b}H^{T}(\mathrm{H}P^{b}H^{T} + \mathrm{R})^{-1} \tag{2}$$

$$\mathrm{P}^{a} = (\mathrm{I} - \mathrm{K}H)\mathrm{P}^{b} \tag{3}$$

Posterior covariance square root form

$$\mathrm{P}^{a} = \frac{\mathrm{X}^{a'}(\mathrm{X}^{a'})^{T}}{N_e - 1} \tag{4}$$

$$= \mathrm{X}^{b'}T(\mathrm{X}^{b'}T)^{T} \tag{5}$$

Center Piece: The Ensemble Kalman Filter equations

Kalman Filter equations

$$X^a = X^b + K(z - Hx^b) \tag{1}$$

$$K = P^b H^T (H P^b H^T + R)^{-1} \tag{2}$$

$$P^a = (I - KH) P^b \tag{3}$$

Posterior covariance square root form

$$P^a = \frac{X^{a'} (X^{a'})^T}{N_e - 1} \tag{4}$$

$$= X^{b'} T (X^{b'} T)^T \tag{5}$$

Find the matrix T

- not uniquely defined
- Studied in oceanography, meteorology and apl. math for last 20 years
- Linear Algebra Tools: SVD, EVD ...
- Best solution depends on problem dimensions

Center Piece: The Ensemble Kalman Filter equations

Kalman Filter equations

$$\mathrm{X}^a = \mathrm{X}^b + \mathrm{K}(z - \mathrm{H}x^b) \quad (1)$$

$$\mathrm{K} = \mathrm{P}^b H^T (\mathrm{H}P^b H^T + \mathrm{R})^{-1} \quad (2)$$

$$\mathrm{P}^a = (\mathrm{I} - \mathrm{K}H)\mathrm{P}^b \quad (3)$$

Posterior covariance square root form

$$\mathrm{P}^a = \frac{\mathrm{X}^{a'}(\mathrm{X}^{a'})^T}{N_e - 1} \quad (4)$$

$$= \mathrm{X}^{b'} T (\mathrm{X}^{b'} T)^T \quad (5)$$

Find the matrix T

- not uniquely defined
- Studied in oceanography, meteorology and apl. math for last 20 years
- Linear Algebra Tools: SVD, EVD ...
- Best solution depends on problem dimensions

Problem dimensions

N_e	Ensemble members in prior
N_y	Number of proxies
N_x	State vector length (grid $\times$ vars)

Center Piece: The Ensemble Kalman Filter equations

STACY

Kalman Filter equations

$$\mathrm{X}^a = \mathrm{X}^b + \mathrm{K}(z - \mathrm{H}x^b) \tag{1}$$

$$\mathrm{K} = \mathrm{P}^b H^T (\mathrm{H}P^b H^T + \mathrm{R})^{-1} \tag{2}$$

$$\mathrm{P}^a = (\mathrm{I} - \mathrm{K}H)\mathrm{P}^b \tag{3}$$

Posterior covariance square root form

$$\mathrm{P}^a = \frac{\mathrm{X}^{a'}(\mathrm{X}^{a'})^T}{N_e - 1} \tag{4}$$

$$= \mathrm{X}^{b'} T (\mathrm{X}^{b'} T)^T \tag{5}$$

Find the matrix T

- not uniquely defined
- Studied in oceanography, meteorology and apl. math for last 20 years
- Linear Algebra Tools: SVD, EVD ...
- Best solution depends on problem dimensions

Problem dimensions

- N_e Ensemble members in prior
- N_y Number of proxies
- N_x State vector length (grid $\times$ vars)

$\rightarrow$ Overview paper by [Vetra-Carvalho et al., 2018] provides unified notation and pseudocode

Solutions for the square root formulation

following Vetra-Carvalho et al 2018

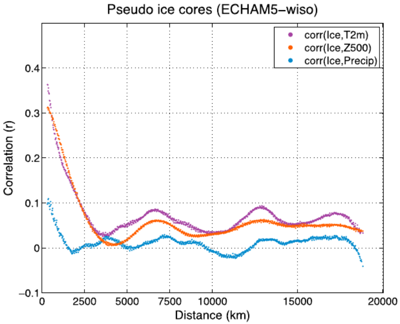
Applied to typical paleoclimate DA situation for reconstructing 1 year.

Method	$\sim$ Time [s]
Serialized Ensemble Square Root KF as used in [Hakim et al., 2016]	10
Direct Ensemble Square Root KF solver [Steiger et al., 2018]	0.5
Optimized EnSRF	0.12
Ensemble Transform KF	0.05
Error subspace Transform KF	0.05

- LMR Code inefficient
- Reconstruction over whole millennium + Monte Carlo techniques: Speed is relevant
- Fastest variants limited by inevitable large matrix multiplication

Range of correlations

Recall that gridpoints not covered by measurements are reconstructed via correlations
From [Steiger et al., 2017]



- How does this plot look like before and after DA for infiltration weighted δ^{18}O?
- Are there spurious long-range correlations in the static prior? ($\rightarrow$ covariance localization)

G. J. Hakim, J. Emile-Geay, E. J. Steig, D. Noone, D. M. Anderson, R. Tardif, N. Steiger, and W. A. Perkins. The last millennium climate reanalysis project: Framework and first results. *Journal of Geophysical Research: Atmospheres*, 121 (12):6745–6764, 2016.

N. J. Steiger, E. J. Steig, S. G. Dee, G. H. Roe, and G. J. Hakim. Climate reconstruction using data assimilation of water isotope ratios from ice cores. *Journal of Geophysical Research: Atmospheres*, 122(3):1545–1568, 2017.

N. J. Steiger, J. E. Smerdon, E. R. Cook, and B. I. Cook. A reconstruction of global hydroclimate and dynamical variables over the common era. *Scientific data*, 5(1): 1–15, 2018.

S. Vetra-Carvalho, P. J. van Leeuwen, L. Nerger, A. Barth, M. U. Altaf, P. Brasseur, P. Kirchgessner, and J.-M. Beckers. State-of-the-art stochastic data assimilation methods for high-dimensional non-gaussian problems. *Tellus A: Dynamic Meteorology and Oceanography*, 70(1):1–43, 2018. doi: 10.1080/16000870.2018.1445364. URL https://doi.org/10.1080/16000870.2018.1445364.