Strategy to generate perturbations of the Drakkar Forcing Set 5.2 (DFS5.2)

F. Massonnet – 11 Feb. 2016

The problem, the constraints and the approach

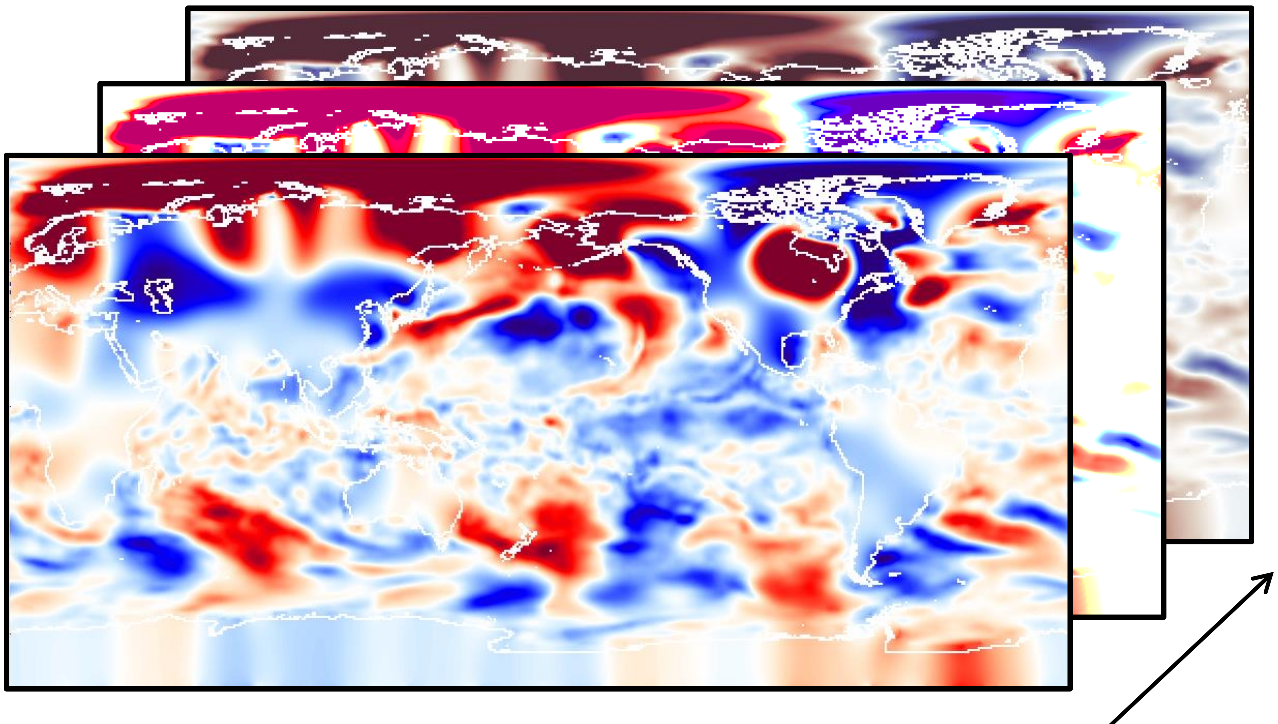
- We need to run *ensembles* for the data assimilation but there is only one forcing available.
- Atmospheric fields have some spatio-temporal **coherence**: we can’t simply add white noise to generate perturbations
- We have to be able to create **as many members as possibe**.

- The solution proposed here is to create an arbitrary number of perturbations that have the same statistics as some reference

We want to create *perturbations* to the actual forcing. For this we need to base ourselves on some examples, from which we can then estimate statistics

Temp. 2-m, 1959 minus 1958

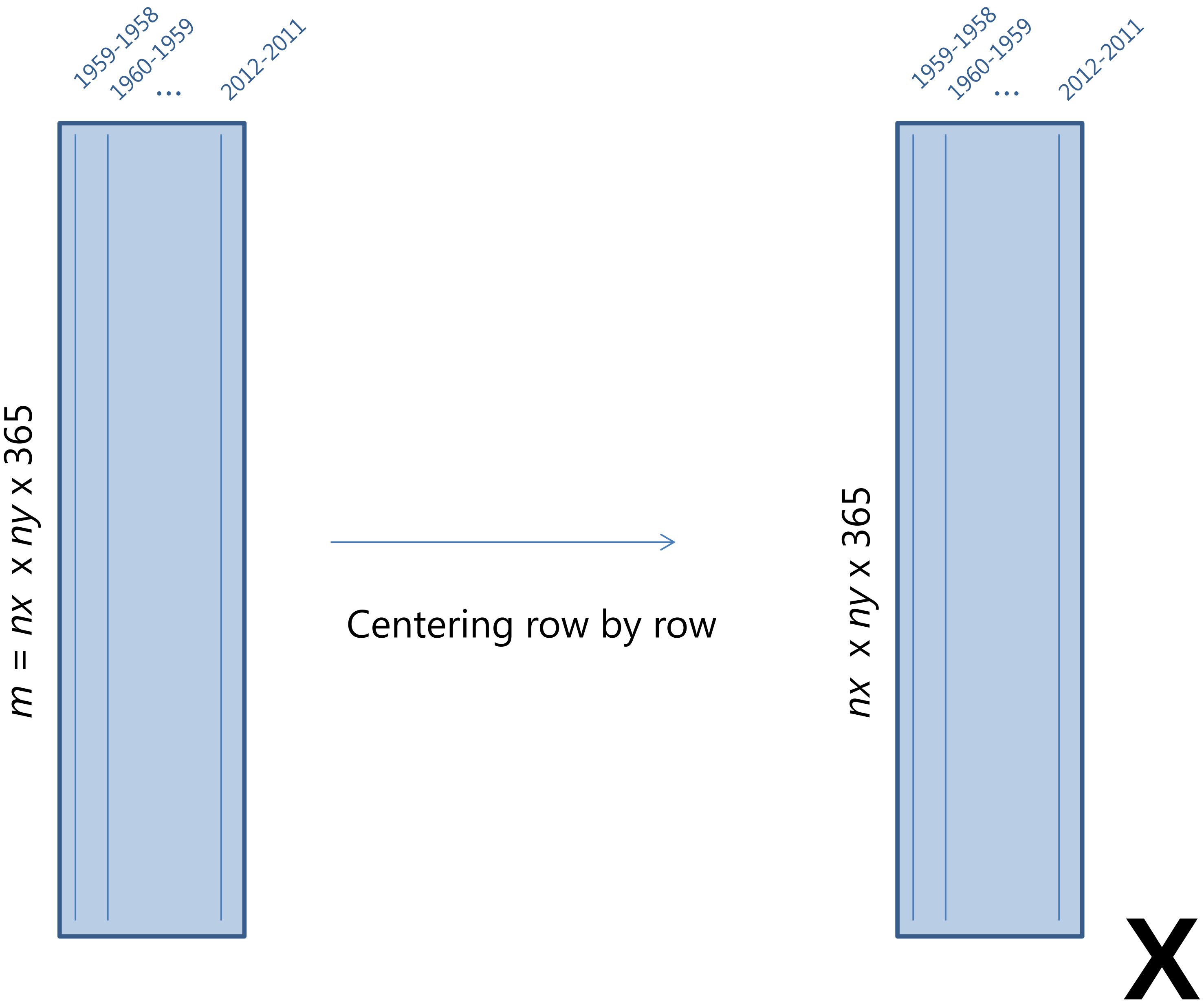
nx = 256



ny = 512

nt = 365 x 8 = 2920

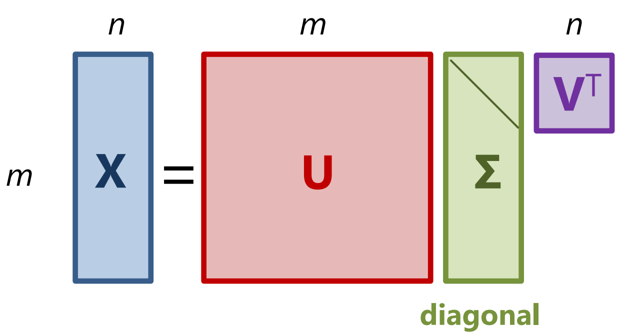
However **R** can't handle a 256 x 512 x 2920 vector, but well a 256 x 512 x 365. Hence daily averages are computed first.



1959-1958
1960-1959
...
2012-2011
m = nx x ny x 365
Centering row by row
1959-1958
1960-1959
...
2012-2011
nx x ny x 365
X

The covariance matrix of the sample is $\mathbf{C} = \mathbf{X}\mathbf{X}^T / (n-1)$

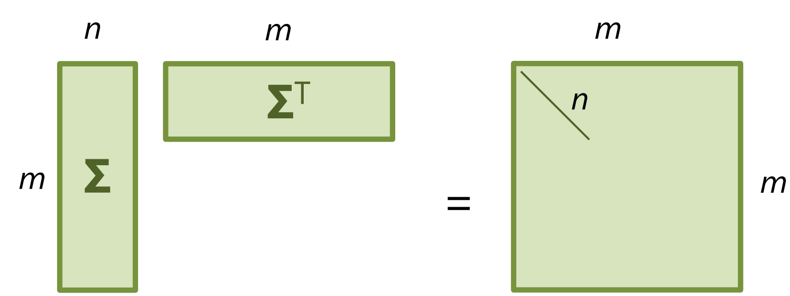
This matrix is *impossible* to handle with *R*, so we need to find a work-around.



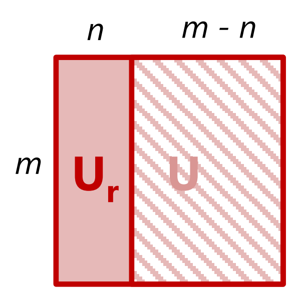
Singular Value Decomposition

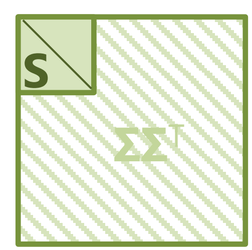
Then $\mathbf{X}\mathbf{X}^T = \mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\mathbf{U}^T$

But $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T$ is a square matrix with only *n* non-zero entries, lying on the diagonal, and equal to the square of singular values

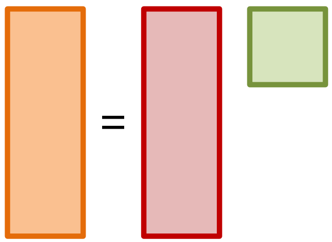


In the product $\mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\mathbf{U}^T$ we can therefore ignore the columns *n+1* to *m* of **U**. **We need only *n* left singular vectors of X**. Let's denote by $\mathbf{U_r}$ the reduced version of **U**, and $\mathbf{S}^{1/2}$ the square root of the reduced version of $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T$





Finally, define $\mathbf{R} = \mathbf{U_r}\, \mathbf{S}^{1/2} \big/ (n\text{-}1)^{1/2}$



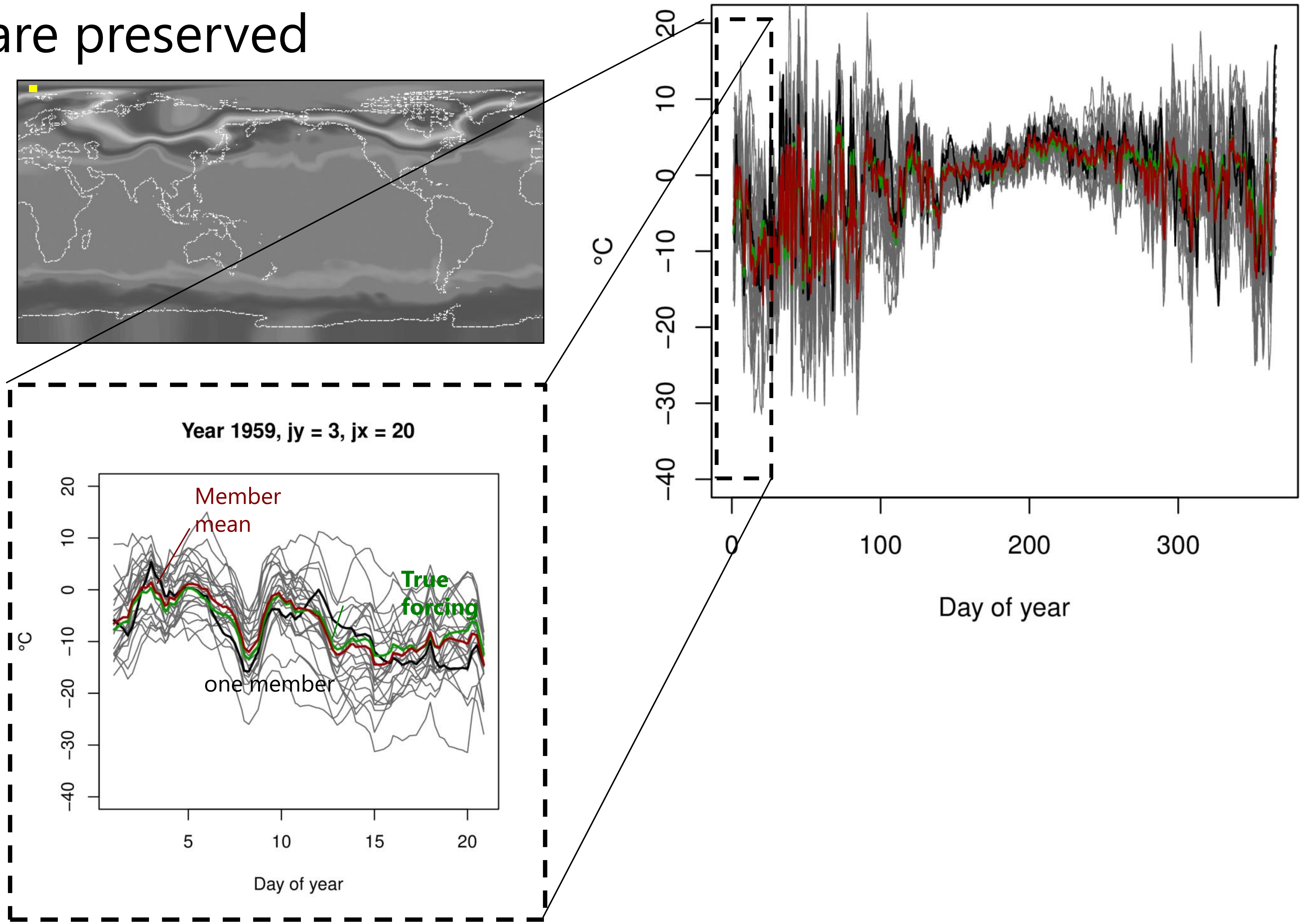
R can generate as many random variables with zero mean and identity covariance matrix. Let's call that **z**. Then $\mathbf{Y} = \mathbf{R} \,.\, \mathbf{z}$ has <u>exactly</u> the same covariance matrix as our sample **X** !

Indeed:

$$\begin{aligned}
\boldsymbol{E}(\mathbf{R} \,.\, \mathbf{z}) &= \mathbf{R} \,.\, \boldsymbol{E}(\mathbf{z}) = \mathbf{0} \\
\boldsymbol{E}(\mathbf{Y}\ \mathbf{Y}^T) &= \mathbf{R}\mathbf{R}^T/(n\text{-}1) = \mathbf{U_r}\ \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\ \mathbf{U_r}^T/(n\text{-}1) \\
&= \mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\mathbf{U}^T/(n\text{-}1) = \mathbf{X}\mathbf{X}^T/(n\text{-}1) = \mathbf{C}
\end{aligned}$$

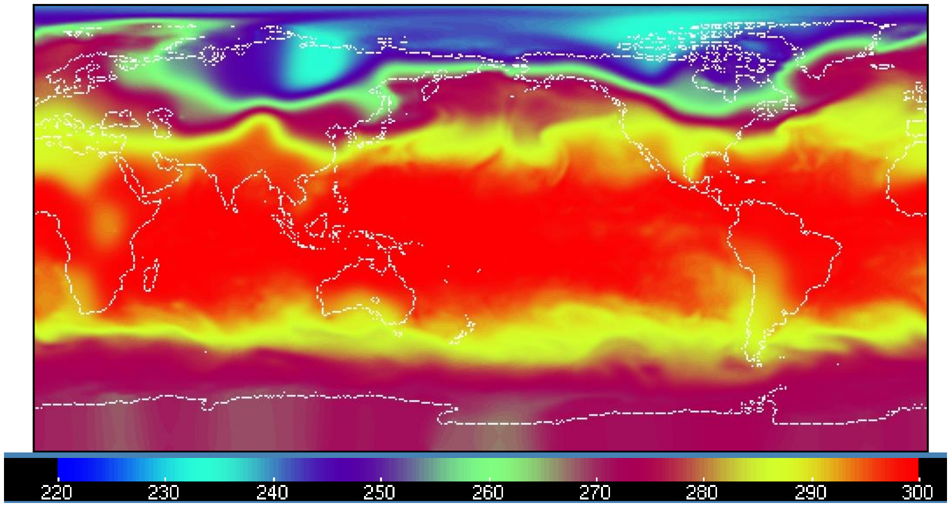
The perturbation created is finally regridded, interpolated to 8-daily and added to the original forcing

Temporal characteristics are preserved

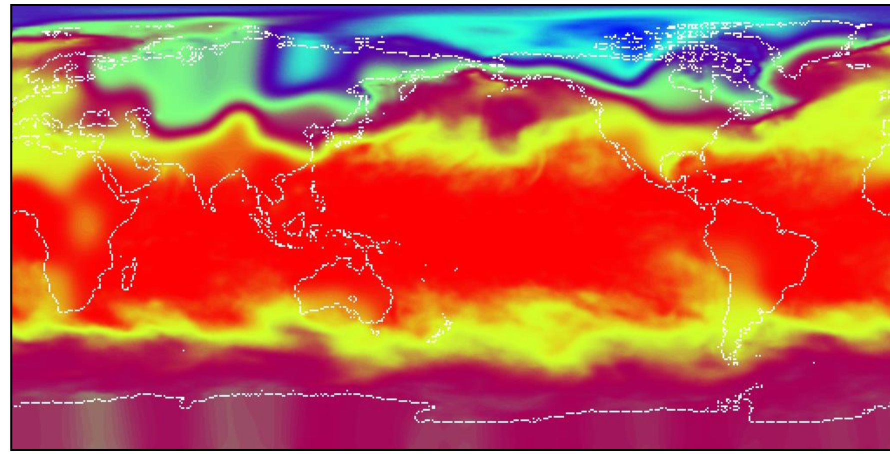


Spatial characteristics are also preserved

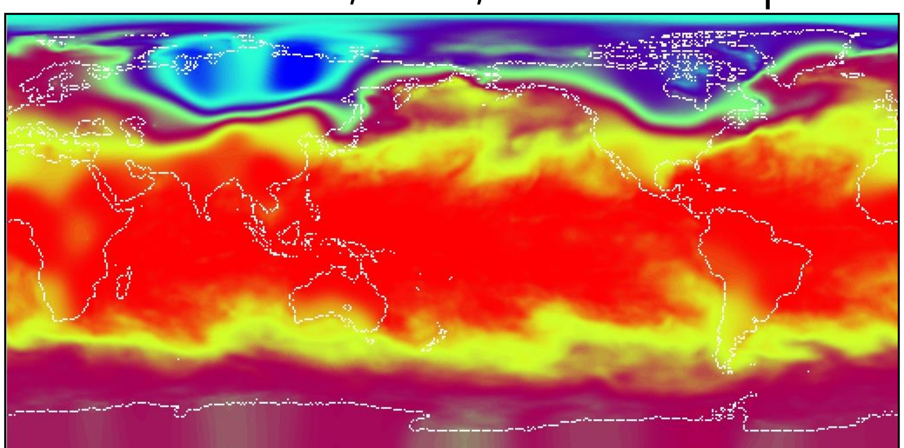
True forcing, 1959, first time step



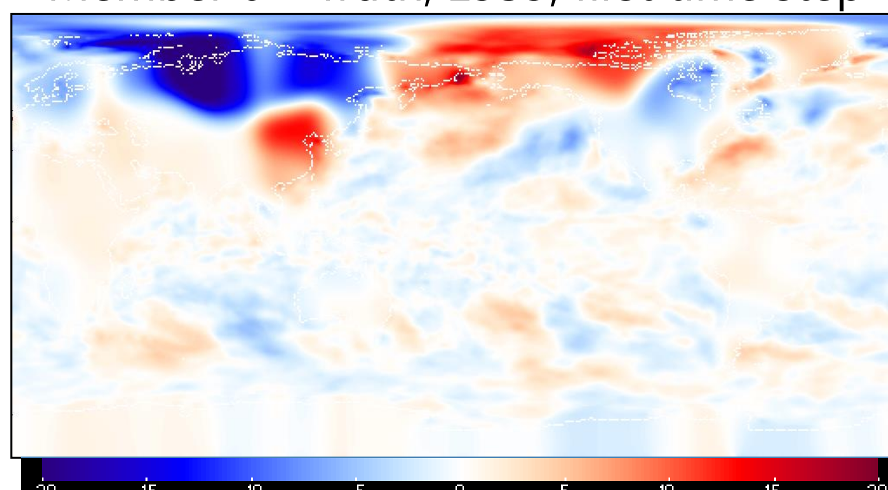
Member 1, 1959, first time step



Member 0, 1959, first time step



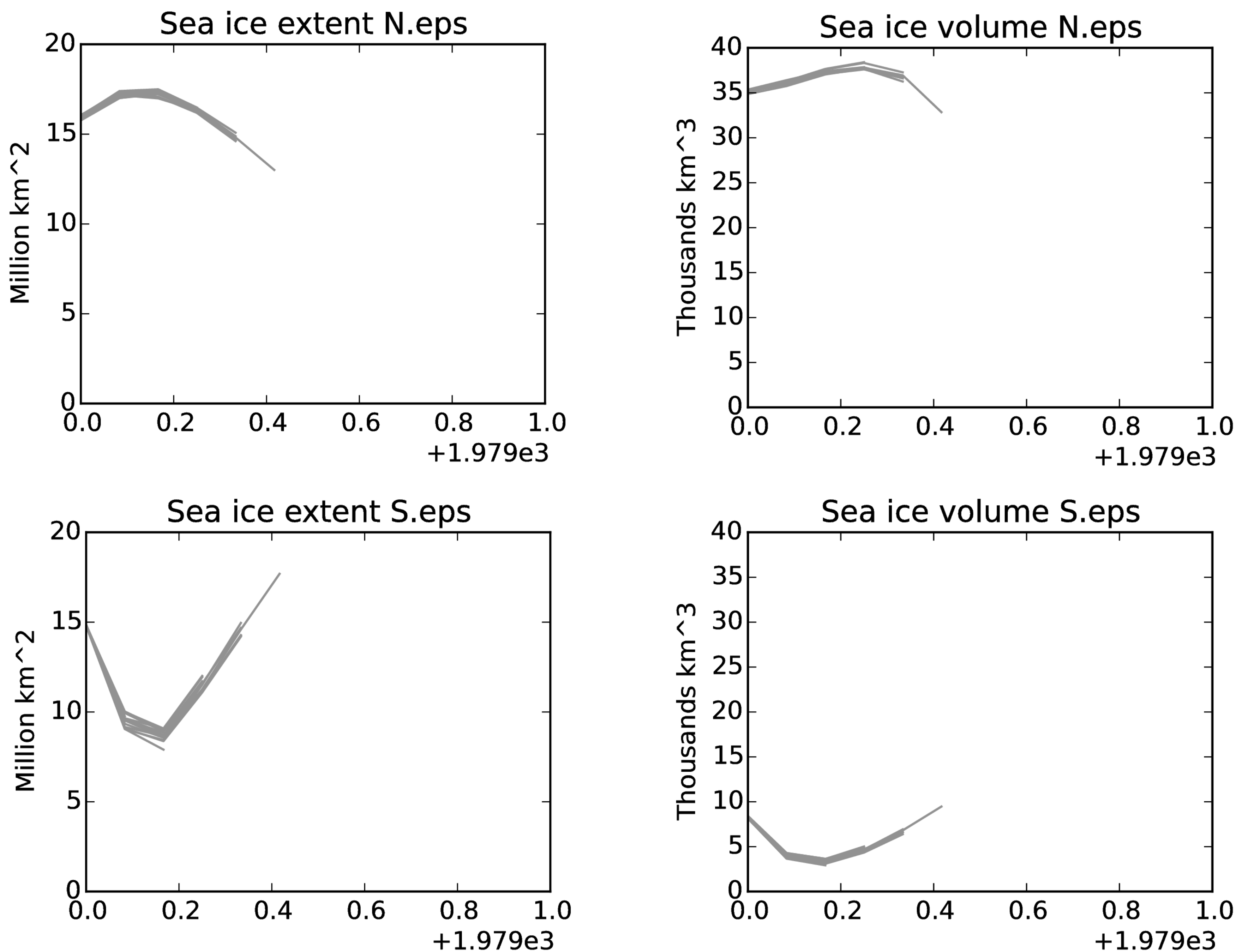
Member 0 – Truth, 1959, first time step



The perturbed forcing is now in
`/esnas/releases/fg/ocean/DFS5.2_perturbed`

(25 perturbations per year, produced with a seed → reproducibility)

First tests are now done with ORCA1.


Sea ice extent N.eps
Million km^2
Sea ice volume N.eps
Thousands km^3
Sea ice extent S.eps
Million km^2
Sea ice volume S.eps
Thousands km^3
0.0
0.2
0.4
0.6
0.8
1.0
+1.979e3