School of Mathematical, Physical and Computational Sciences

University of Reading

Treating sample covariances for use in strongly coupled atmosphere-ocean data assimilation

Polly J. Smith | Amos S. Lawless | Nancy K. Nichols

Aims

- Strongly coupled variational data assimilation for coupled systems requires the specification of cross-domain forecast error covariances;
- ensembles can be used to estimate these, but sample covariances are typically rank deficient and/or ill-conditioned.
- ***How can we obtain a well-conditioned matrix that retains important cross-covariance information?***

Matrix modification

We recall that the condition number of a symmetric, positive definite matrix **S** is defined by the ratio of the largest to smallest eigenvalue

$$\kappa(\mathbf{S}) = \lambda_{\max}(\mathbf{S})/\lambda_{\min}(\mathbf{S})$$

We consider two methods to improve the conditioning of sample covariance matrices:

1. Matrix reconditioning

Specify a required condition number κ_{tol} and increment all eigenvalues by a fixed amount λ_{inc} such that

$$\frac{\lambda_{\max} + \lambda_{\mathrm{inc}}}{\lambda_{\min} + \lambda_{\mathrm{inc}}} = \kappa_{\mathrm{tol}}$$

Note that reconditioning the covariance matrix in this way is not the same as reconditioning the correlation matrix.

2. Localization

Form Schur product of a localization matrix and the ensemble covariance matrix. For coupled assimilation we need to think carefully how to apply this to the cross-domain blocks and their sub-matrices.

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{\mathbf{AA}} & \mathbf{C}_{\mathbf{AO}} \\ \mathbf{C}_{\mathbf{AO}}^{T} & \mathbf{C}_{\mathbf{OO}} \end{pmatrix}$$

Here we apply localization separately to each sub-matrix. Localization can be applied to the covariance or correlation matrix with the same effect.

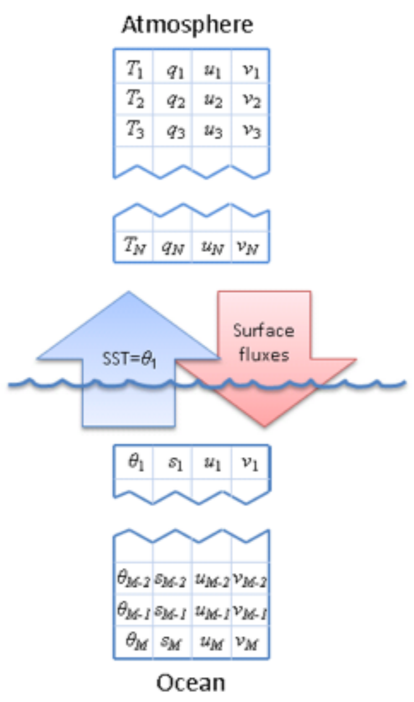
Single-column model

Atmosphere

- simplified version of the ECMWF single column model adiabatic component + vertical diffusion (no convection)
- 4 state variables on 60 model levels (surface to ~0.1hPa)
- forced by large scale horizontal advection

Ocean

- K-Profile Parameterisation (KPP) mixed-layer model based on the scheme of Large et al.
- 4 state variables on 35 levels (1-250m)
- forced by short and long wave radiation at surface



coupled via SST and surface fluxes of heat, moisture & momentum

Covariance generation

Sample covariances are generated using an ensemble of strongly-coupled 4D-Var identical twin experiments, as in Smith et al. (2017).

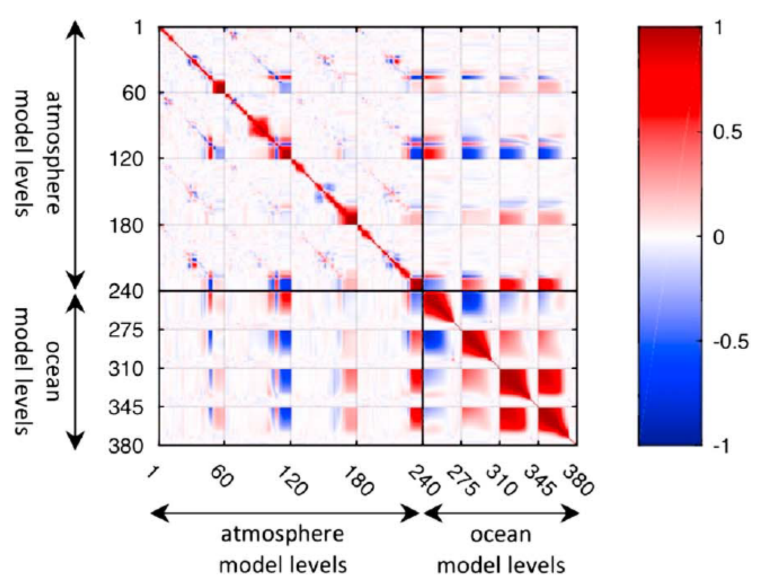


Figure 1: Raw sample correlation matrix from a 500-member ensemble.

Results are shown for a matrix derived from a 500-member ensemble averaged over 4 different assimilation cycles, for a point in the NW Pacific in December 2013. The raw ensemble matrix is shown in fig. 1.

Contact information

- Email: p.j.smith@reading.ac.uk
- Department of Meteorology, University of Reading, Whiteknights, RG6 6BB, UK

Results

Matrix reconditioning

We recondition the matrix to a target condition number of 10^4:

- reconditioning the correlation matrix reduces the condition number, but retains sample noise, fig. 2(a).
- reconditioning the covariance matrix destroys ocean and atmosphere-ocean cross correlations associated with smallest eigenvalues, fig 2(b).

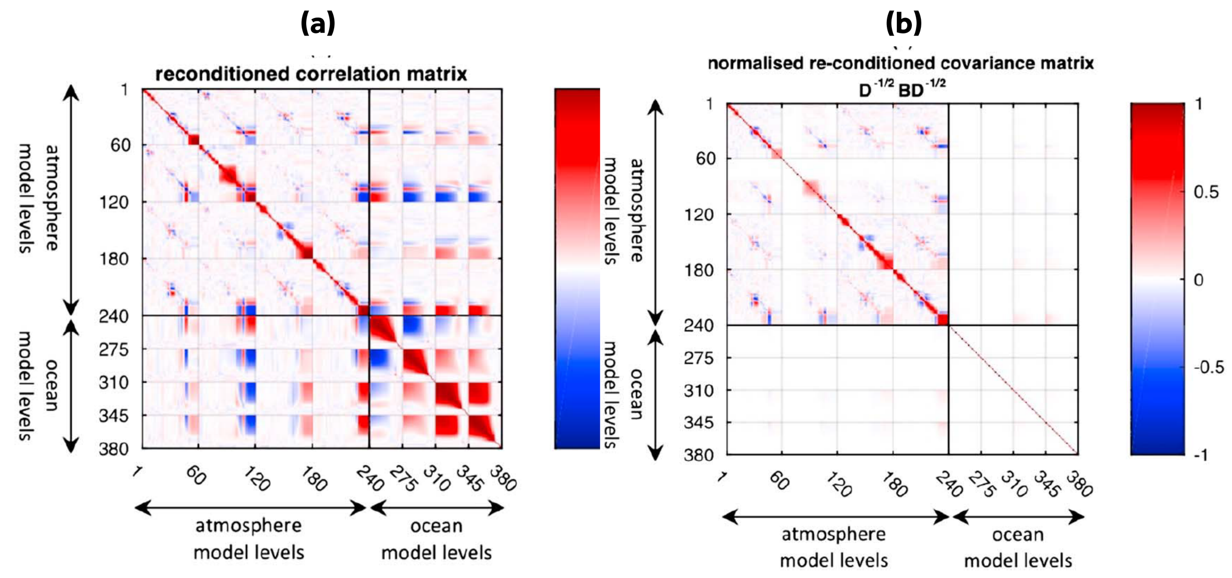


Figure 2: Reconditioned (a) correlation matrix and (b) covariance matrix.

Localization

We define a scaled distance between an atmosphere and ocean point, similar to Frolov et al. (2016):

$$\hat{d}(z_a(i), z_o(j)) = \left(\frac{z_a(i)}{L_a} + \frac{z_o(j)}{L_o}\right)$$

- localization reduces sampling error, but retains high condition number of 10^9, fig 3(a).
- condition number only reduced to 10^4 if use very short lengthscales, which destroys correlations, fig 3(b).

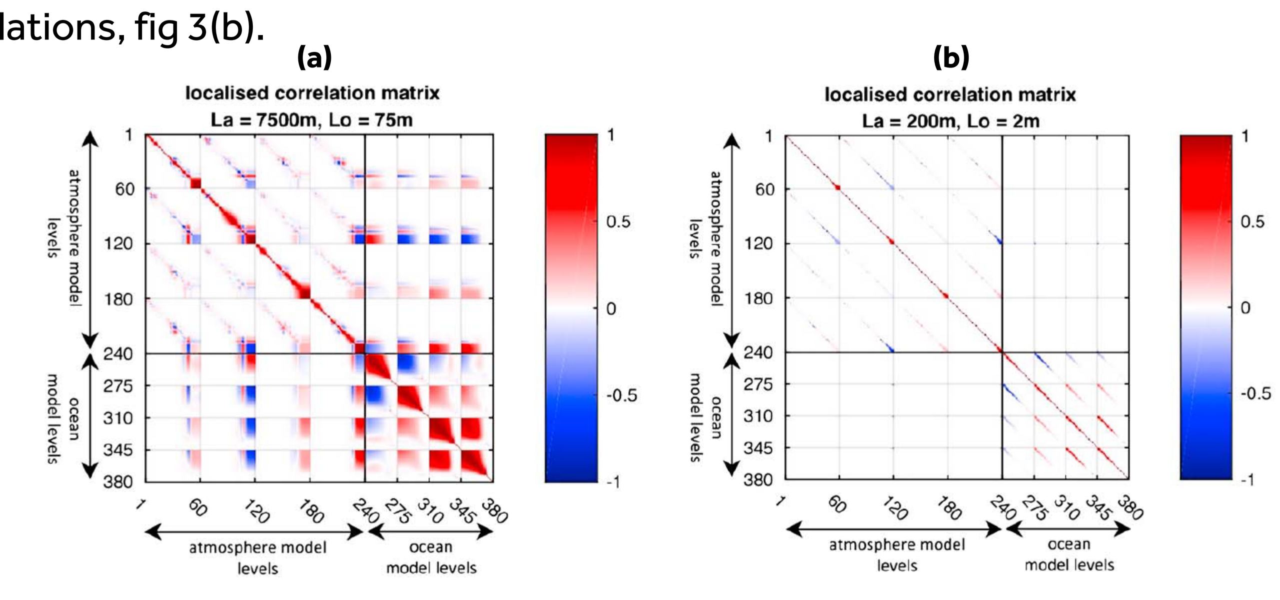


Figure 3: Localized correlation matrix with (a) long lengthscales and high condition number; (b) short lengthscales and low condition number.

Best of both worlds?

We first recondition and then localize (fig. 4):

- ✓ Sampling noise is removed.
- ✓ Cross-correlation signals are retained.
- ✓ The matrix is well-conditioned.

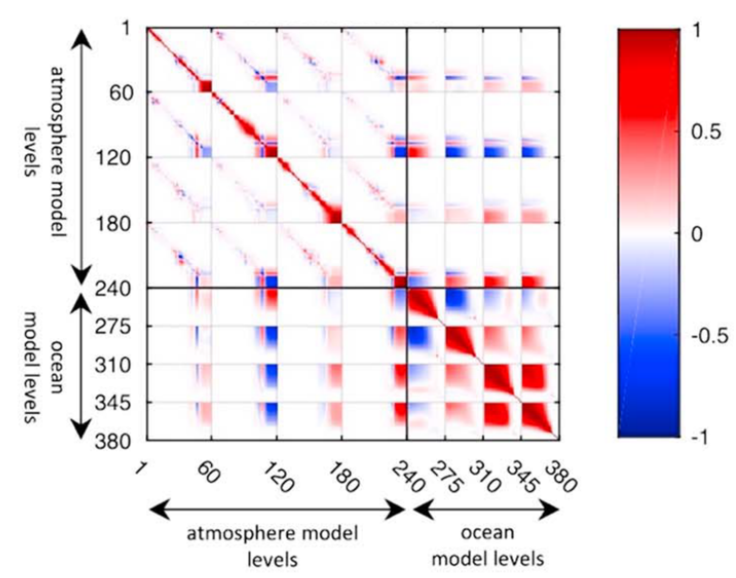


Figure 4: Error correlation matrix reconditioned with $\kappa_{\mathrm{tol}} = 10^4$ and then localized using long lengthscales as fig. 3(a).

Summary

- Reconditioning of correlation matrix can reduce the condition number, but sampling noise is retained.
- Important to treat the correlation matrix rather than covariance matrix, so as not to lose important signals.
- Localization can reduce sampling error, but the matrix still ill-conditioned.
- Combination of reconditioning and localization leads to a well-conditioned matrix, with cross-correlations retained and sampling error removed.

References

- Smith et al. (2018), Geophys. Res. Lett., doi: 10.1002/2017GL075534
- Smith et al. (2017), Mon. Wea. Rev., doi: 10.1175/MWR-D-16-0284.1
- Frolov et al. (2016), Mon. Wea. Rev., doi: 10.1175/MWR-D-15-0041.1

NERC SCIENCE OF THE ENVIRONMENT
National Centre for Earth Observation NATURAL ENVIRONMENT RESEARCH COUNCIL
Data Assimilation Research Centre