Coupled atmosphere-ocean data assimilation

Polly Smith

Data Assimilation Research Centre

p.j.smith@reading.ac.uk

Outline

- motivation
- coupled DA in theory vs. coupled DA in practice
- examples - idealised system experiments

Why coupled DA?

Traditionally, initial conditions for coupled atmosphere-ocean model forecasts are provided by combining analyses from independent (uncoupled) data assimilation systems:

- ignores interactions between systems
- imbalance at initial time (initialisation shock)
- near surface data not fully utilised, e.g. SST, scatterometer winds

Coupled data assimilation (CDA) treats the atmosphere and ocean as a single coherent system:

- improved use of near-surface observations
- greater balance between atmosphere and ocean analysis fields
  - reduction of initialisation shocks in coupled forecasts
- generation of a consistent system state for the initialisation of coupled forecasts across all timescales
  - from short range to decadal
- better re-analysis of past climate and climate variation
- greater understanding & representation of air-sea exchange processes
- diagnosing & understanding coupled model errors

Terminology

strongly coupled

- all components analysed within a single seamless assimilation framework - single fully coupled analysis
- observations in one domain directly influence analysis in the other

weakly coupled

- analysis computed independently for each model component
- model-observation misfits (the `innovations') measured against the coupled model forecast state
- immediate impact of observations limited to domain in which they reside

Also many intermediate or 'quasi' approaches.

Incremental 4D-Var

Solve iteratively

set $\mathbf{x}_0^{(0)} = \mathbf{x}_b$

outer loop: for k = 0, … , Nouter

compute $\mathbf{d}_i^{(k)} = \mathbf{y}_i - h(\mathbf{x}_i^{(k)})$, where $\mathbf{x}_i^{(k)} = m(t_i, t_0, \mathbf{x}_0^{(k)})$

inner loop: minimise

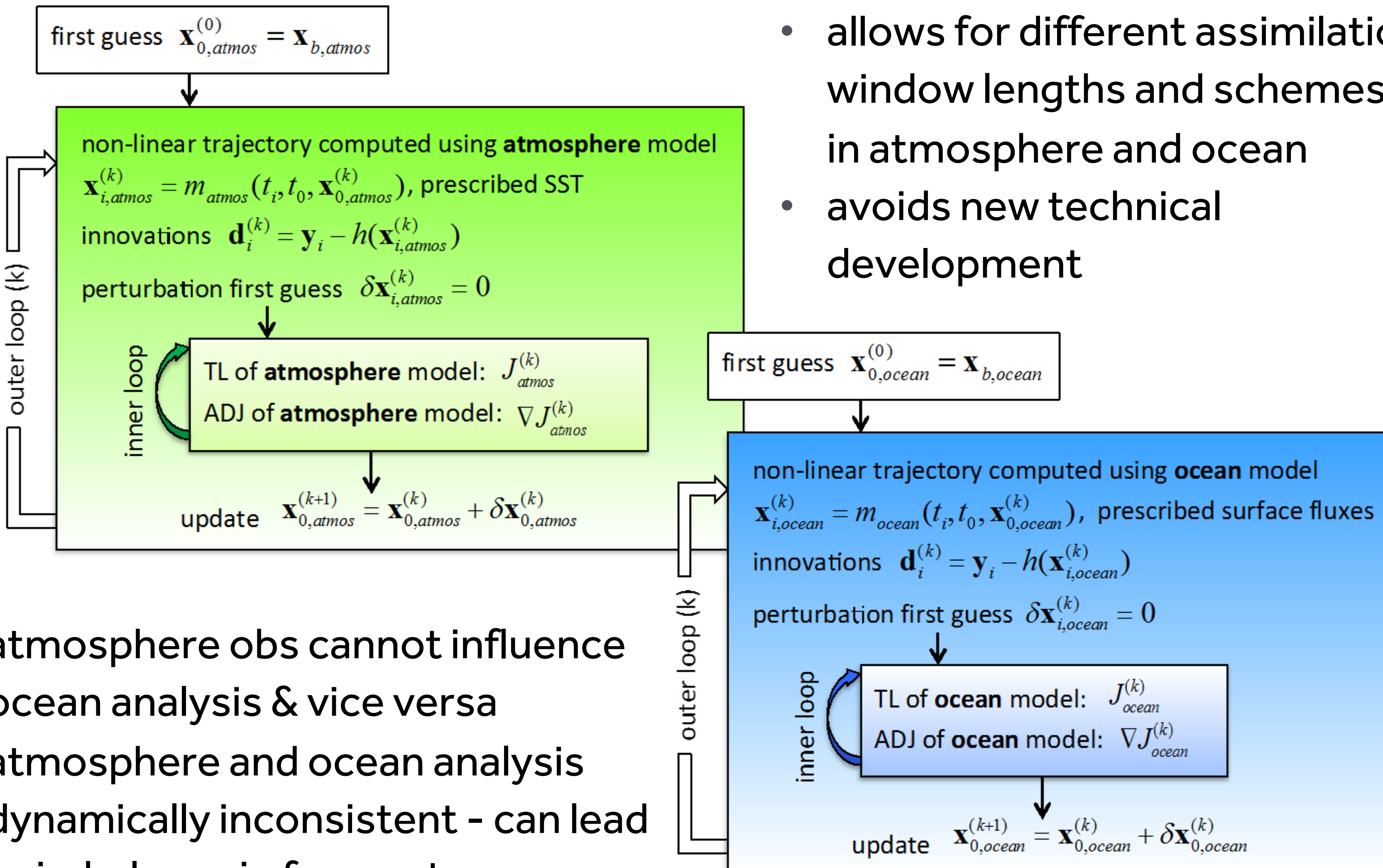
$$\mathrm{J}^{(k)}\left(\delta\mathbf{x}_0^{(k)}\right) = \frac{1}{2}\left(\delta\mathbf{x}_0^{(k)} - \left(\mathbf{x}_b - \mathbf{x}_0^{(k)}\right)\right)^T \mathbf{B}^{-1}\left(\delta\mathbf{x}_0^{(k)} - (\mathbf{x}_b - \mathbf{x}_0^{(k)})\right) + \frac{1}{2}\sum_{i=0}^{n}\left(\mathbf{H}_i\delta\mathbf{x}_i^{(k)} - \mathbf{d}_i^{(k)}\right)^T \mathbf{R}_i^{-1}\left(\mathbf{H}_i\delta\mathbf{x}_i^{(k)} - \mathbf{d}_i^{(k)}\right)$$

subject to $\delta\mathbf{x}_i^{(k)} = \mathbf{M}(t_i, t_0, \mathbf{x}^{(k)})\delta\mathbf{x}_0^{(k)}$

update $\mathbf{x}_0^{(k+1)} = \mathbf{x}_0^{(k)} + \delta\mathbf{x}_0^{(k)}$

Uncoupled incremental 4D-Var

University of Reading



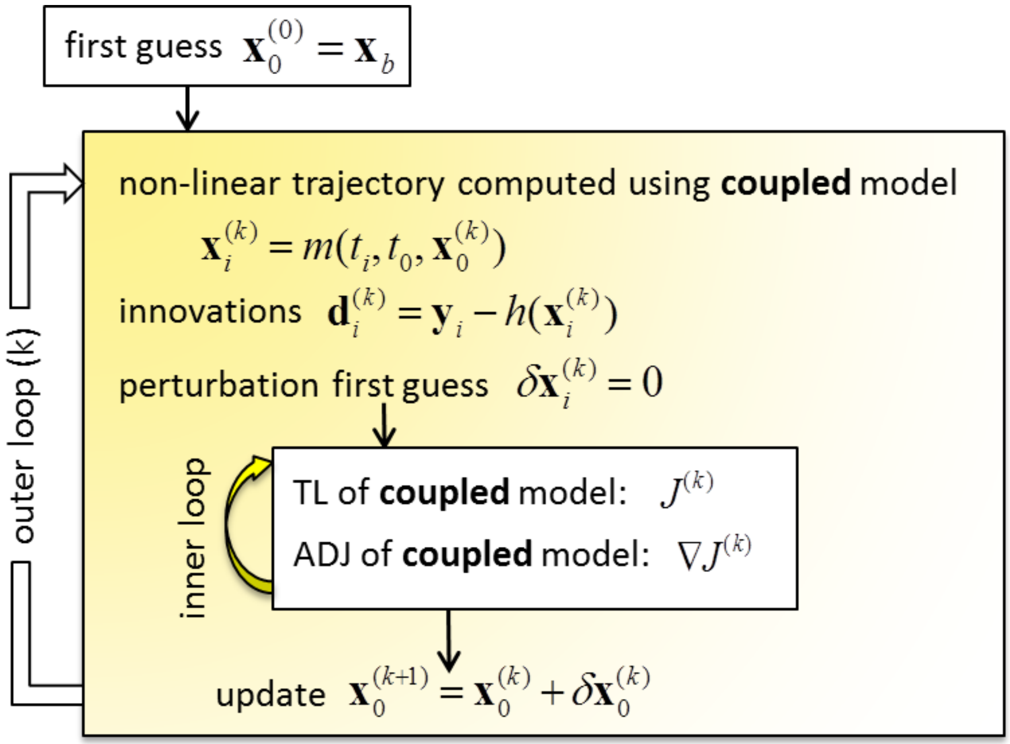
- allows for different assimilation window lengths and schemes in atmosphere and ocean
- avoids new technical development

- atmosphere obs cannot influence ocean analysis & vice versa
- atmosphere and ocean analysis dynamically inconsistent - can lead to imbalance in forecast

Strongly coupled incremental 4D-Var

- control vector contains both atmosphere & ocean model variables
- fully coupled tangent linear & adjoint models
- allows for cross-domain error covariances between atmosphere & ocean forecast errors

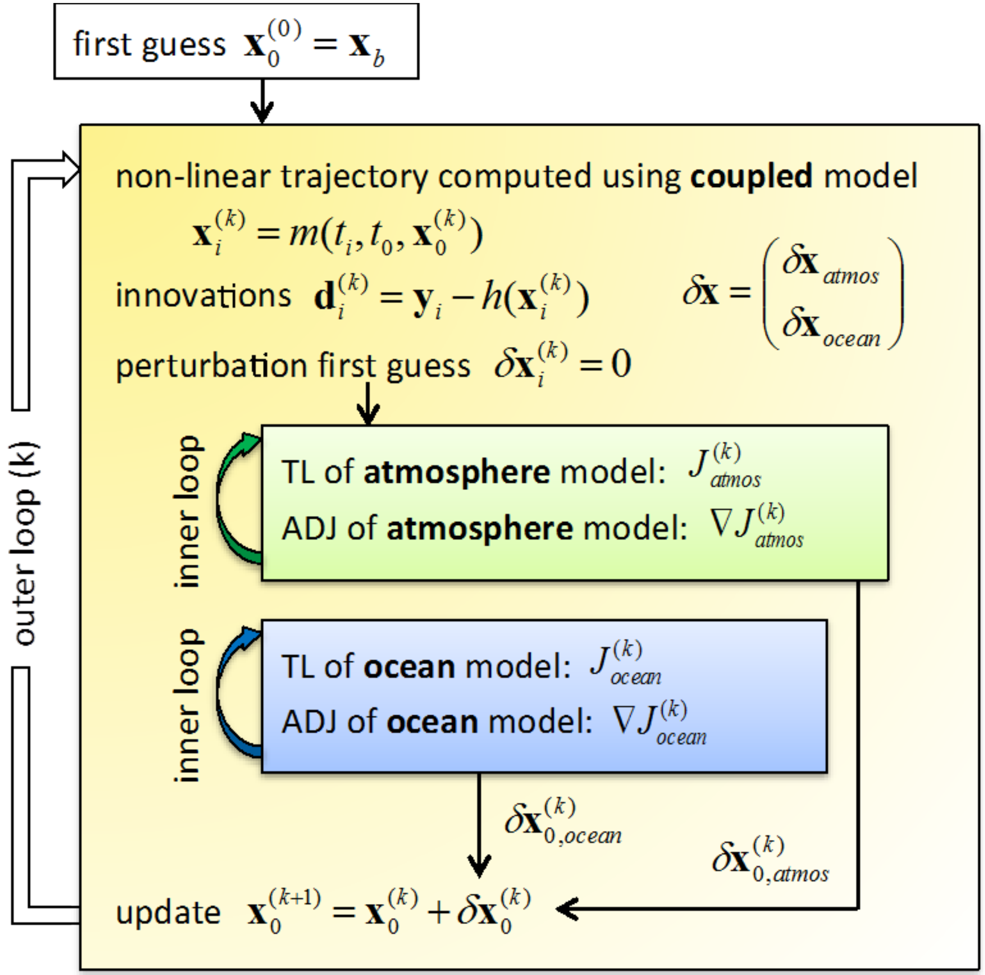
$$\mathbf{B}_0 = \begin{pmatrix} \mathbf{B}_{\mathrm{AA}} & \mathbf{B}_{\mathrm{AO}} \\ \mathbf{B}_{\mathrm{AO}}^{\mathrm{T}} & \mathbf{B}_{\mathrm{OO}} \end{pmatrix}$$



- atmosphere obs can influence ocean analysis & vice versa

Weakly coupled incremental 4D-Var

- coupled model used in outer loop
- separate inner loop cost functions
- no explicit cross-domain error covariances
- atmosphere (ocean) observations can only influence ocean (atmosphere) analysis if multiple outer loops used
- limits amount of new technical development
- allows for different assimilation window lengths and schemes in ocean and atmosphere



Practical challenges

Strongly coupled DA is simple in theory, but in practice ...

- increased dimension of state vector
- increased complexity of interactions between variables
- increased computational cost
- potentially different spatio-temporal scales
- requires specification of the relationship between the errors in the atmosphere and ocean model forecasts (Var)
- requires coupled localization techniques (ensemble/ hybrid methods)

Idealised system experiments

- comparison of uncoupled, weakly coupled & fully coupled DA strategies
  - can coupled DA reduce or eliminate initialisation shock & improve balance at initial time?
- coupled forecast error covariance estimation
  - analysis-ensemble method
- coupled forecast error covariance implementation
  - reconditioning/ localization
  - single & double observation experiments (making greater use of near-surface observations)

Idealised system

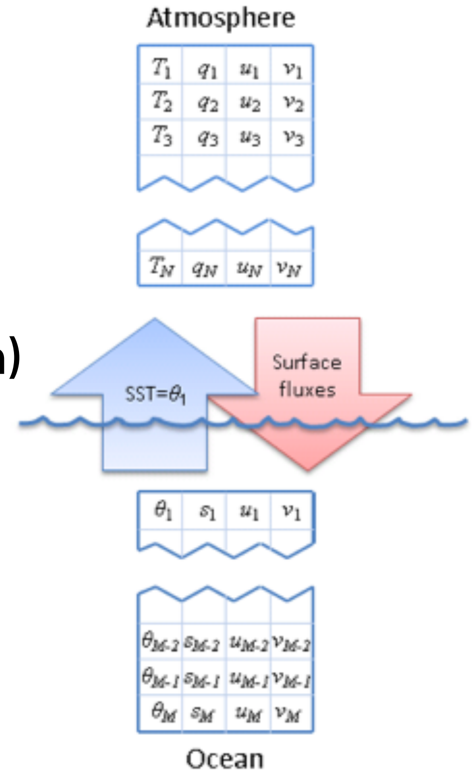
single-column, coupled atmosphere-ocean model

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
- 4 state variables on 35 model levels (1-250m)
- forced by short and long wave radiation at surface



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

Smith et al 2015, doi:10.3402/tellusa.v67.27025

Experiment set-up

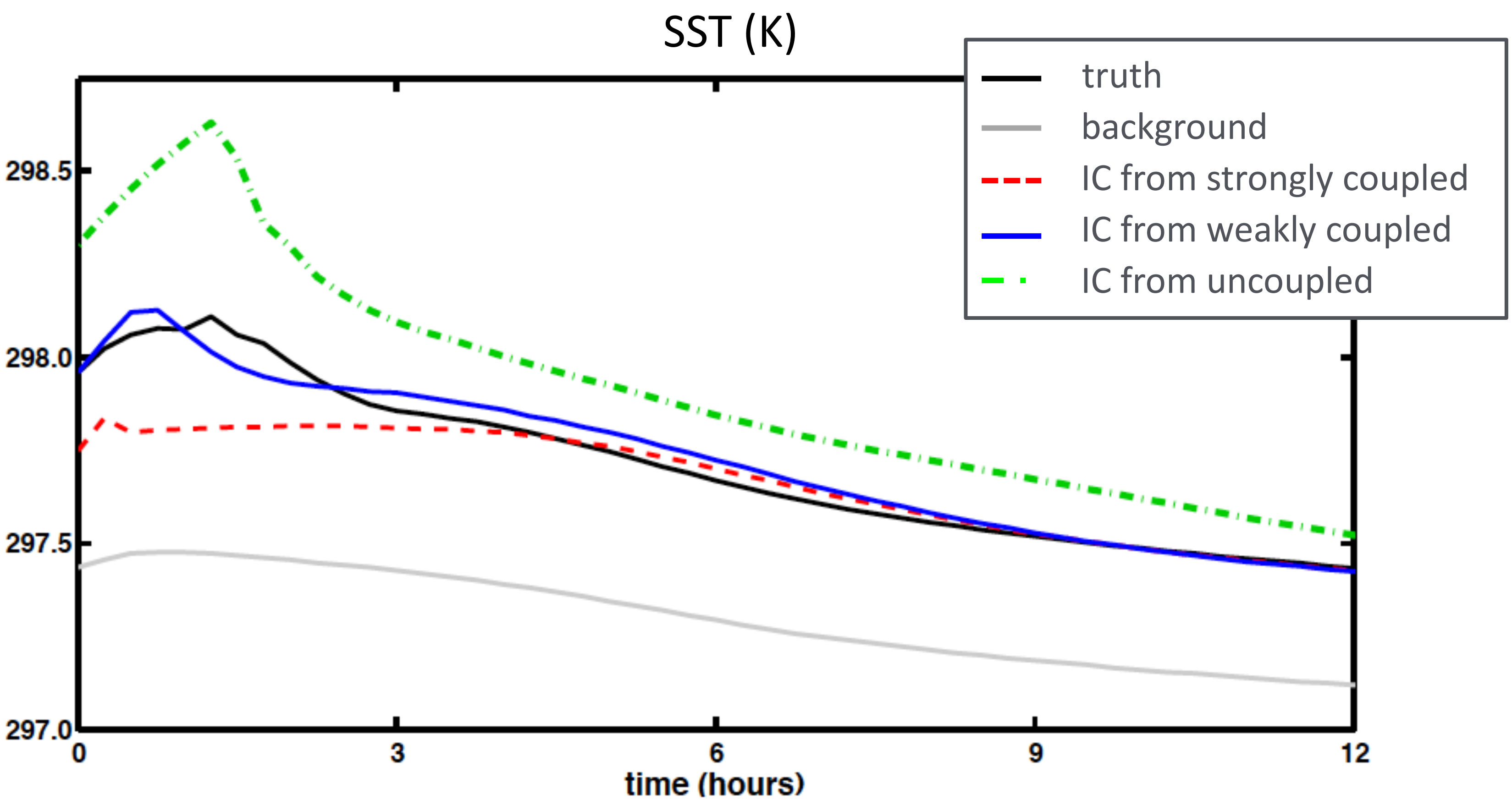
- identical twin (coupled non-linear model assumed to be perfect)
- 12 hour assimilation window, 3 outer loops
- data from June or December 2013, for point in NW Pacific Ocean
- 'true' initial state is coupled model forecast initialised using ERA Interim and Mercator Ocean data
- initial background state is a perturbed coupled model forecast
- 3 hourly observations are generated by adding random noise to 'truth'
- uncoupled assimilations - SST & surface fluxes from ERA interim

Idealised system experiments

- comparison of uncoupled, weakly coupled & fully coupled DA strategies
  - can coupled DA reduce or eliminate initialisation shock & improve balance at initial time?
- coupled forecast error covariance estimation
  - analysis-ensemble method
- coupled forecast error covariance implementation
  - reconditioning/ localization
  - single & double observation experiments (making greater use of near-surface observations)

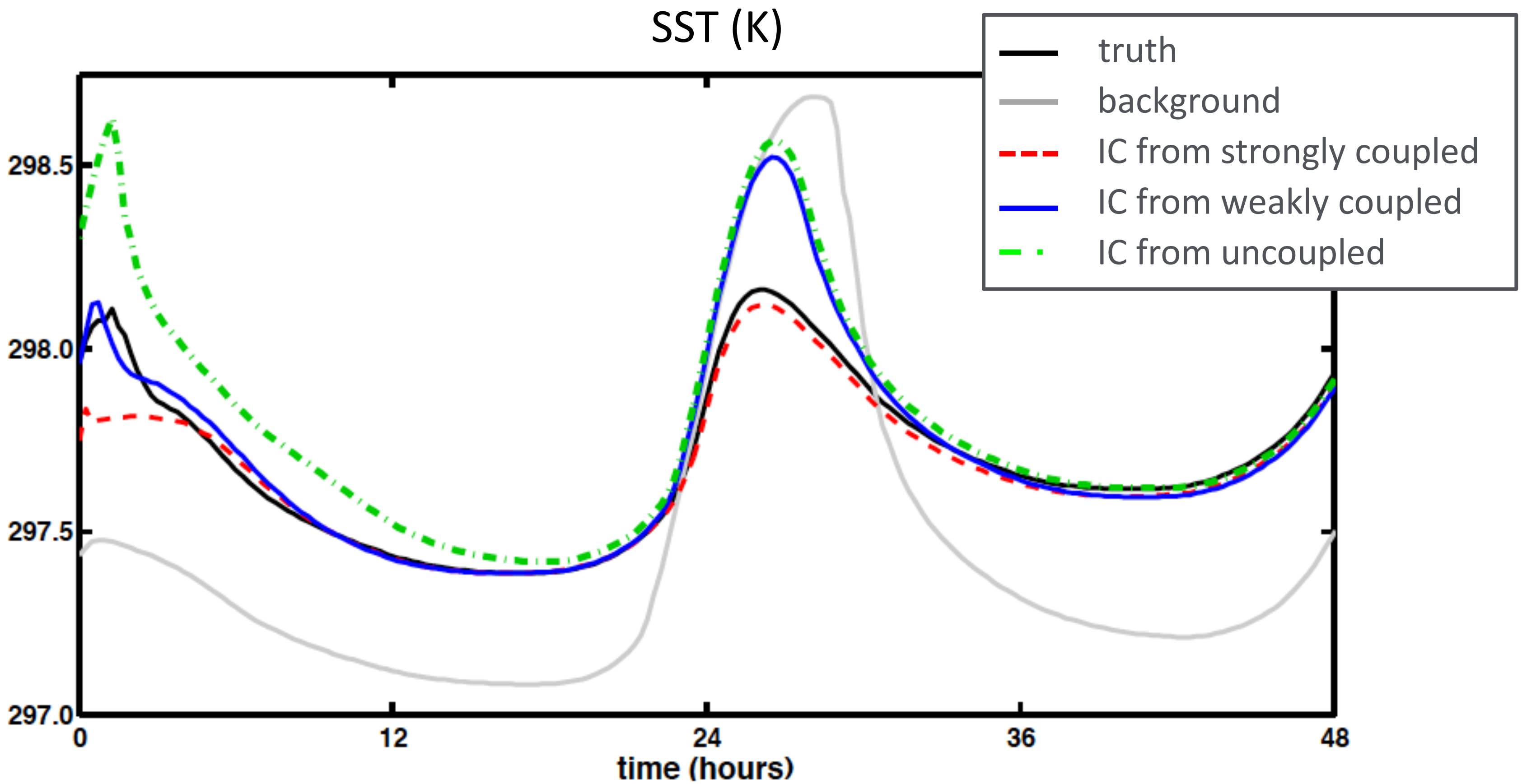
Initialisation shock (1)

coupled forecast initialised from t_0 analyses – first 12 hours



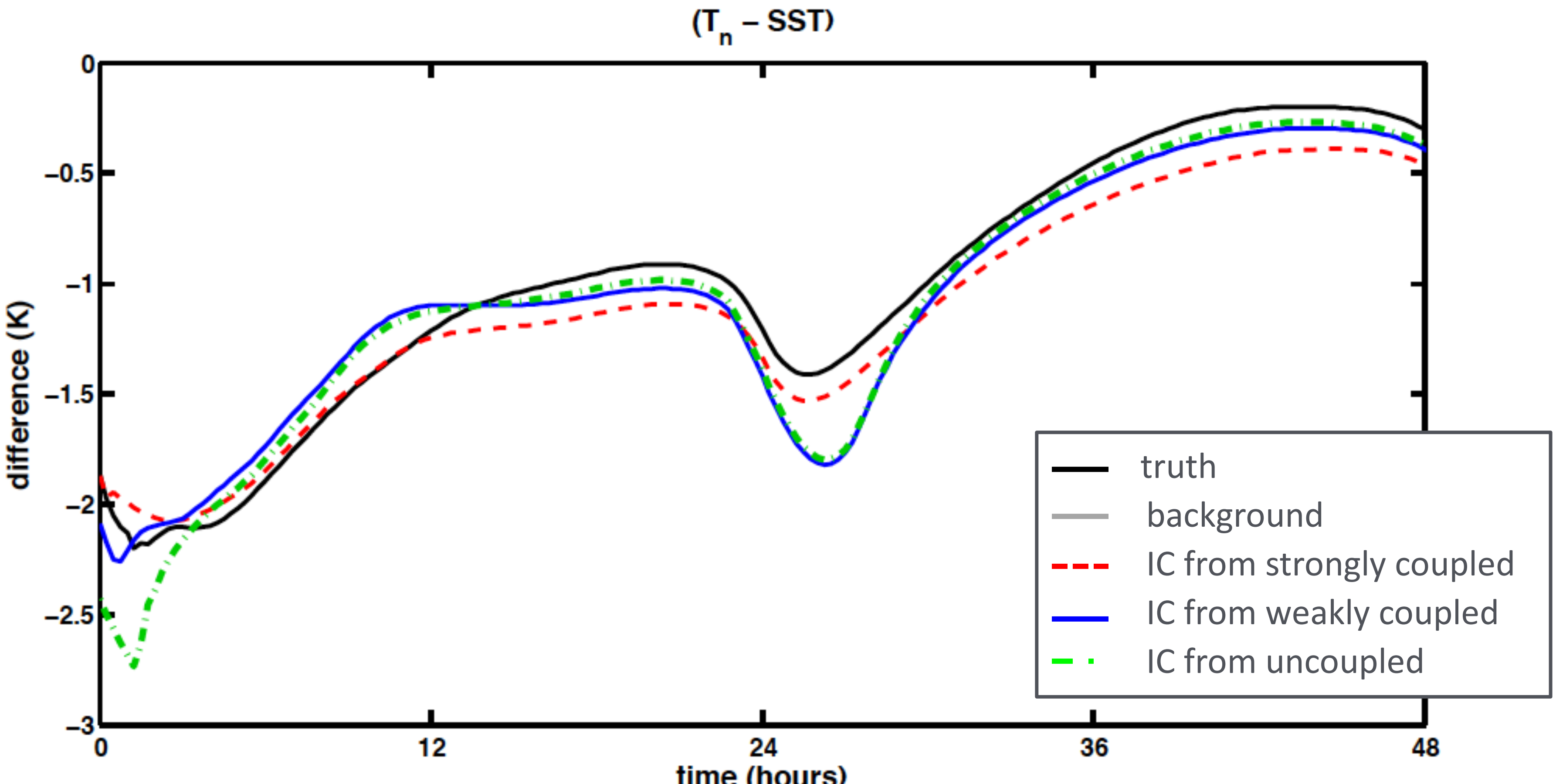
Initialisation shock (2)

coupled forecast initialised from t_0 analyses - up to 48 hours



Air-sea balance

atmosphere-ocean temperature difference



Key points

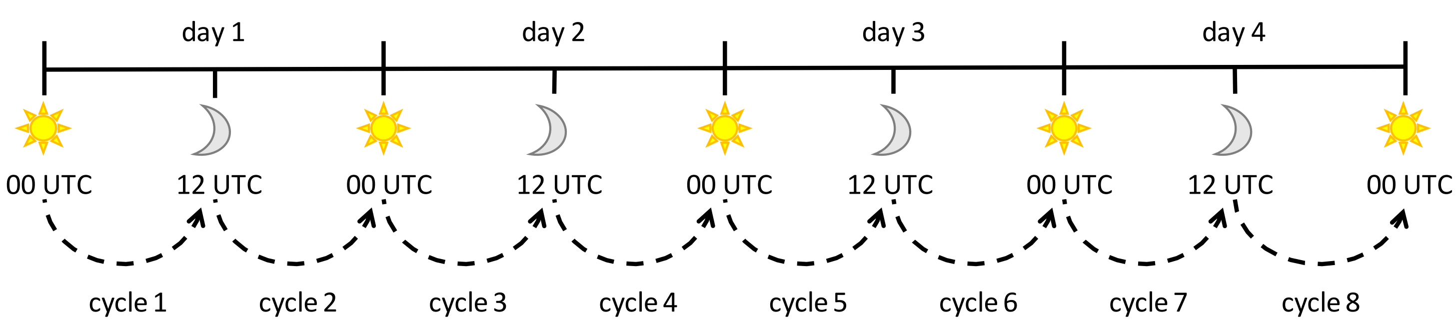
- compared to uncoupled assimilation, coupled assimilation has positive impact on initial balance (initialisation shock reduced)
- weakly coupled systems are more sensitive to the frequency and number of observations, but still offers benefits over uncoupled systems
- greater transfer of information in weakly-coupled assimilation if both systems are observed.

Idealised system experiments

- comparison of uncoupled, weakly coupled & fully coupled DA strategies
  - can coupled DA reduce or eliminate initialisation shock & improve balance at initial time?
- coupled forecast error covariance estimation
  - analysis-ensemble method
- coupled forecast error covariance implementation
  - reconditioning/ localization
  - single & double observation experiments (making greater use of near-surface observations)

Ensemble error correlations

- estimate background error covariance from a 500 member ensemble of perturbed strongly coupled 4D-Var analyses
- average over a several assimilation cycles to increase effective ensemble size.
- 8 cycles, each uses 12 hour assimilation window
- each cycle starts at either 12 UTC or 00 UTC which corresponds to the early hours of the morning and early afternoon local time.
- allows comparison of day-night plus summer-winter error correlations



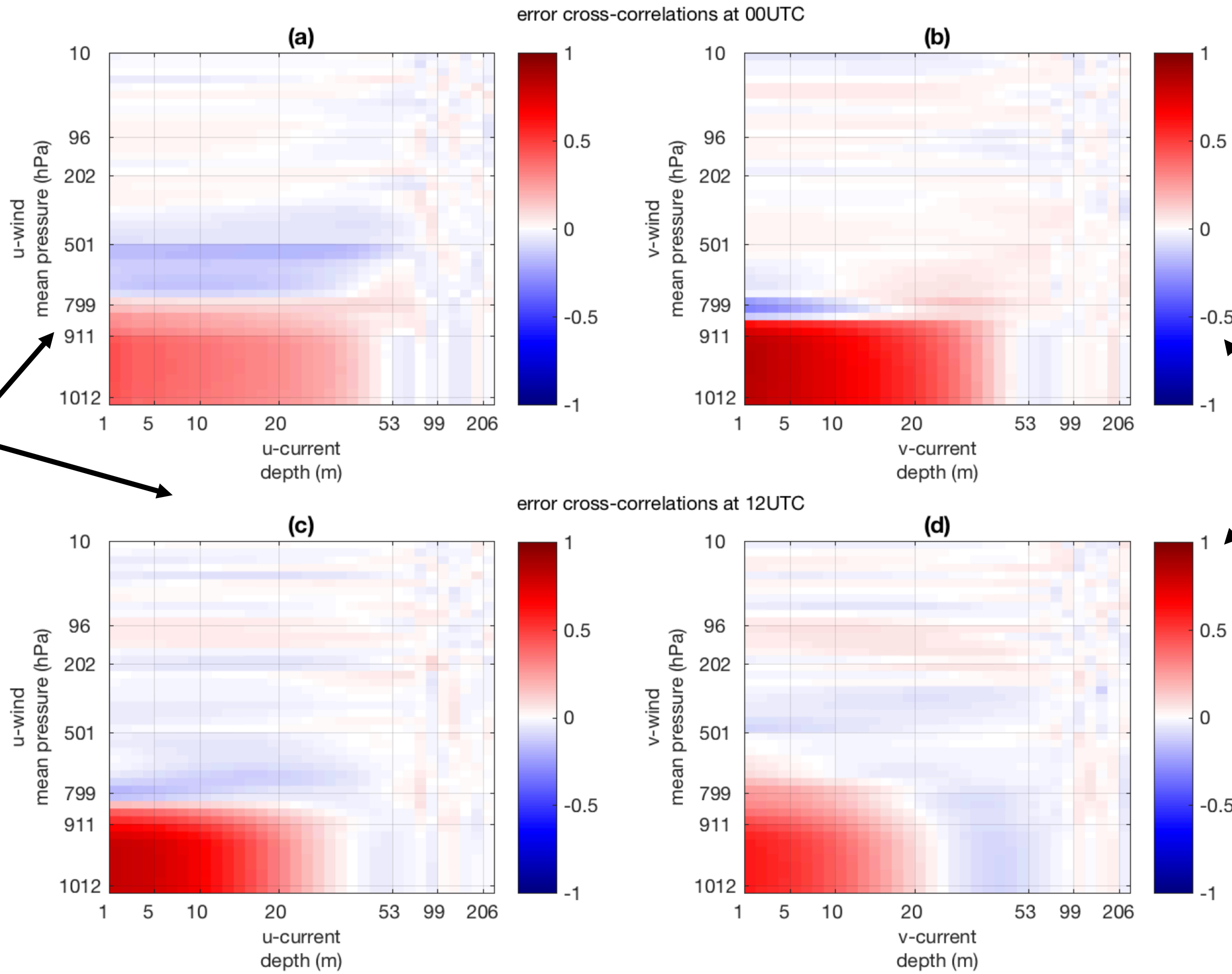
Questions

- where are the atmosphere-ocean error cross-correlations strongest?
- how do the atmosphere-ocean error cross correlation structures vary between summer and winter, and between day and night?
- can we explain our results by considering the underlying model physics, forcing and known atmosphere-ocean feedback mechanisms?

December case:

University of Reading

u-wind
&
u-current



day

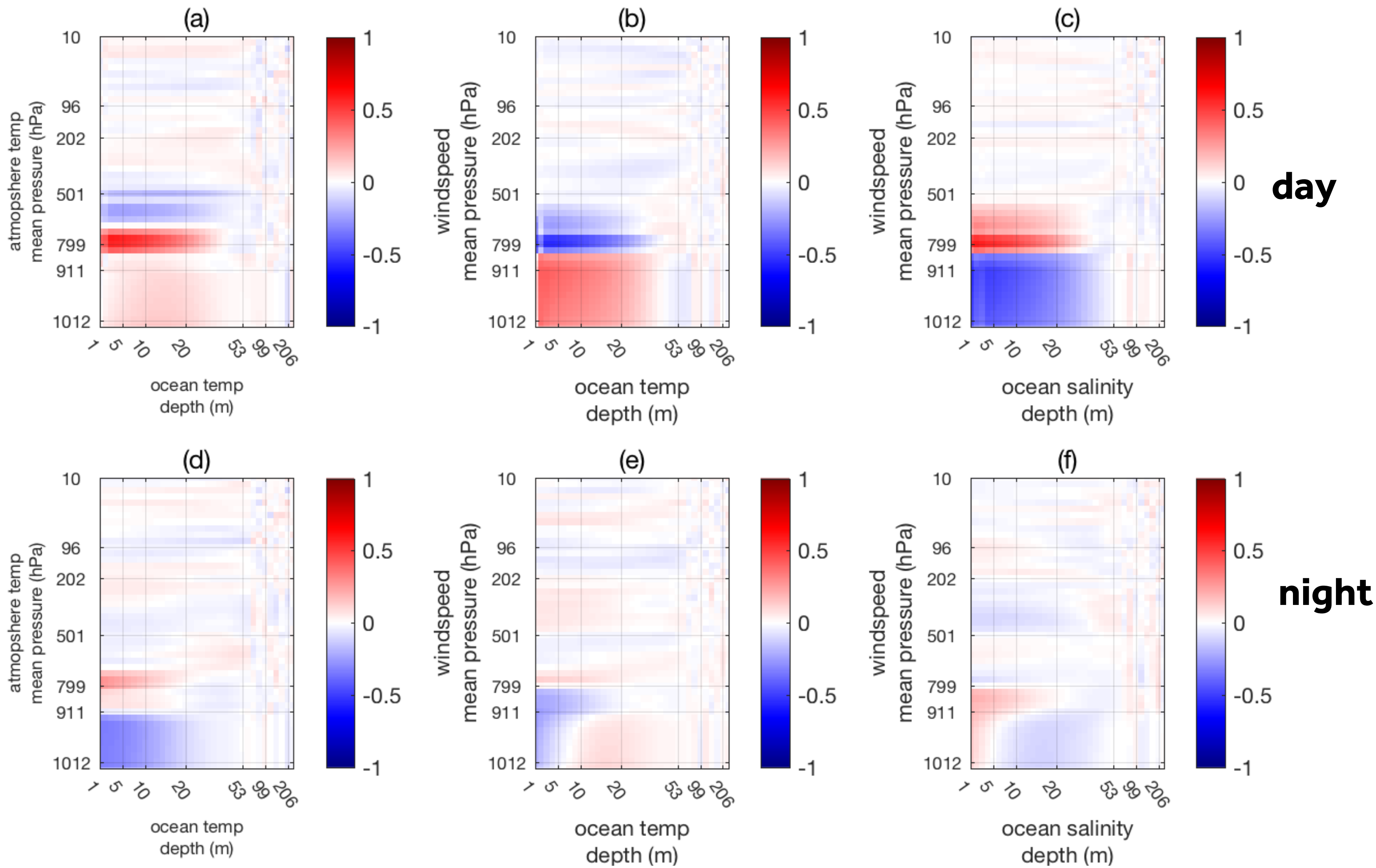
v-wind
&
v-current

night

atmosphere wind-ocean current error cross-correlations

December case:

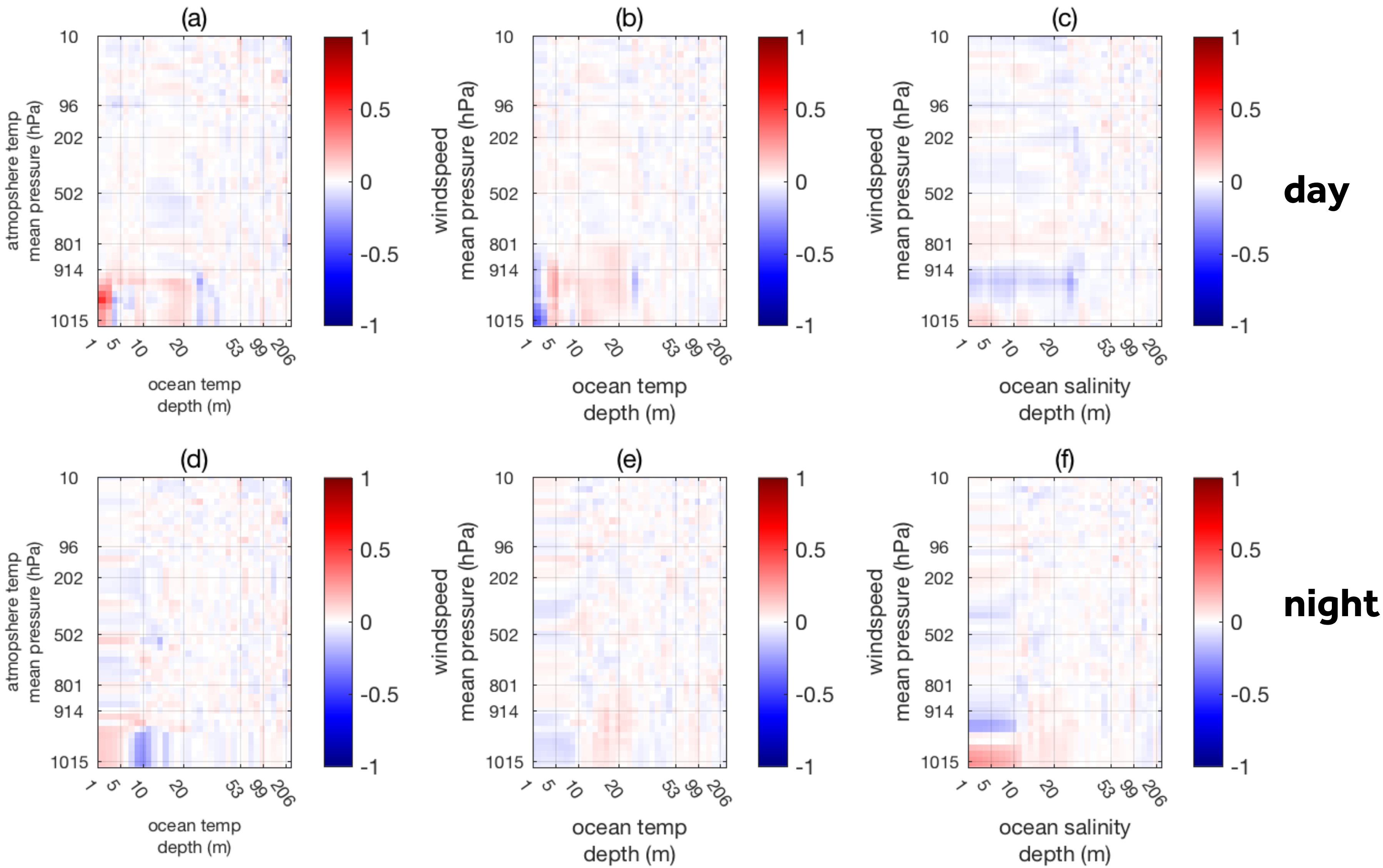
atmosphere-ocean error cross-correlations



left to right: atmosphere-ocean temp, wind speed-ocean temp, wind speed-ocean salinity

June case:

atmosphere-ocean error cross-correlations



left to right: atmosphere-ocean temp, wind speed-ocean temp, wind speed-ocean salinity

Key points

- correlations are between ***errors*** in different atmosphere and ocean forecast fields.
- errors in two different variables will not necessarily interact in the same way as the model variables themselves and this interaction may not be linear.
- strongest error cross-correlations are in near surface atmosphere-ocean boundary, beyond this atmosphere-ocean errors appear to be mostly uncorrelated.
- error correlation structures are most distinct in the winter case: effect of solar insolation on ocean stability is reduced, surface winds are high and the atmosphere-ocean surface temperature difference is large; these combine to produce turbulent heat fluxes of greater magnitude so that air-sea coupling is strong.

Idealised system experiments

- comparison of uncoupled, weakly coupled & fully coupled DA strategies
  - can coupled DA reduce or eliminate initialisation shock & improve balance at initial time?
- coupled forecast error covariance estimation
  - analysis-ensemble method
- coupled forecast error covariance implementation
  - reconditioning/ localization
  - single & double observation experiments (making greater use of near-surface observations)

- have shown that ensembles can be used to estimate required cross-domain forecast error covariances, but …
- sample covariances are typically rank deficient/ and or ill conditioned and marred by sampling noise
- How can we obtain a well-conditioned matrix that retains important covariance information?

Condition number

For a symmetric positive definite matrix **S** the condition number is given by

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

Matrix modification methods

1. Matrix reconditioning

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

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

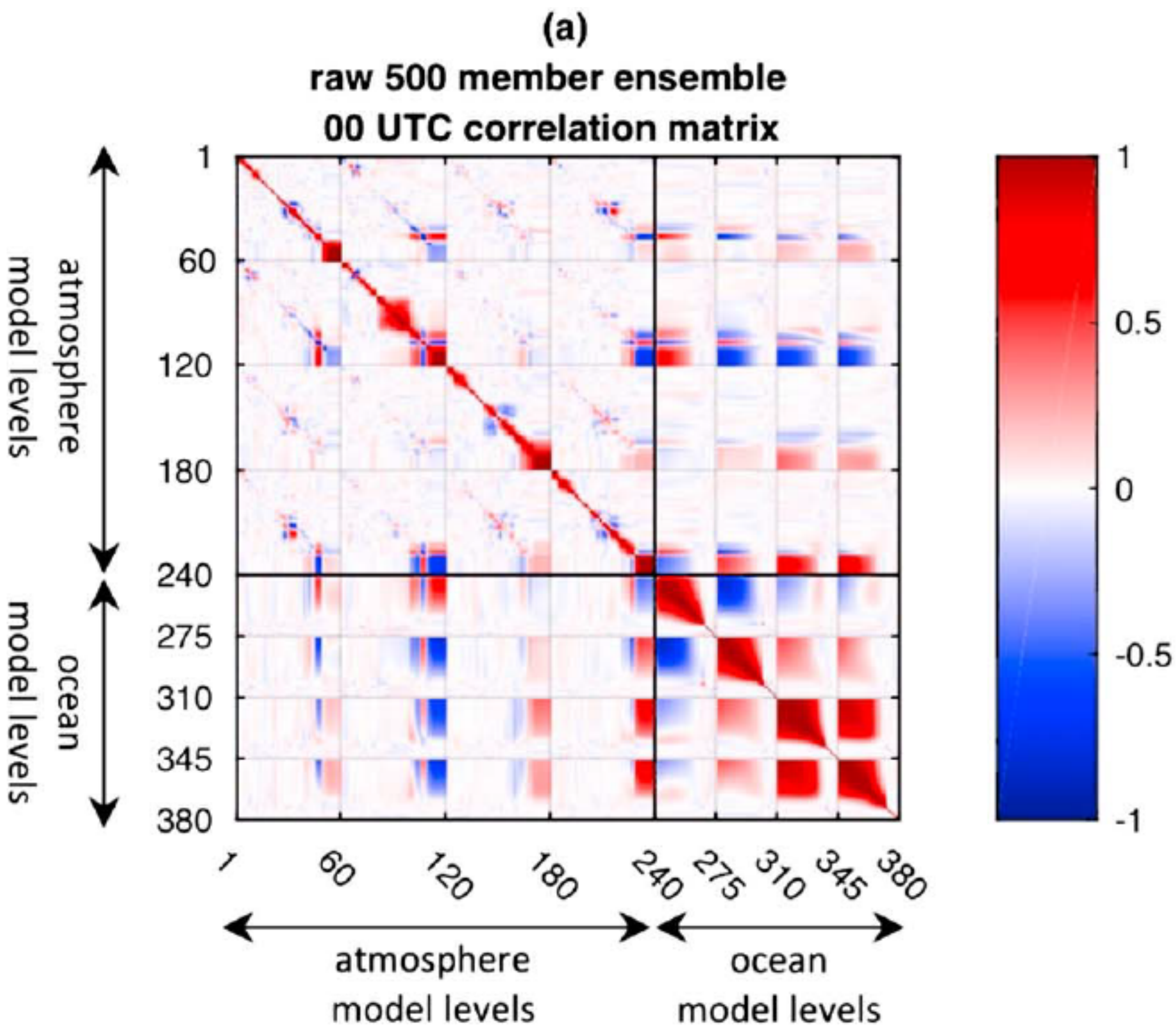
- *Note*: reconditioning the covariance matrix in this way is **not** the same as reconditioning the correlation matrix.

Matrix modification methods

2. Localization

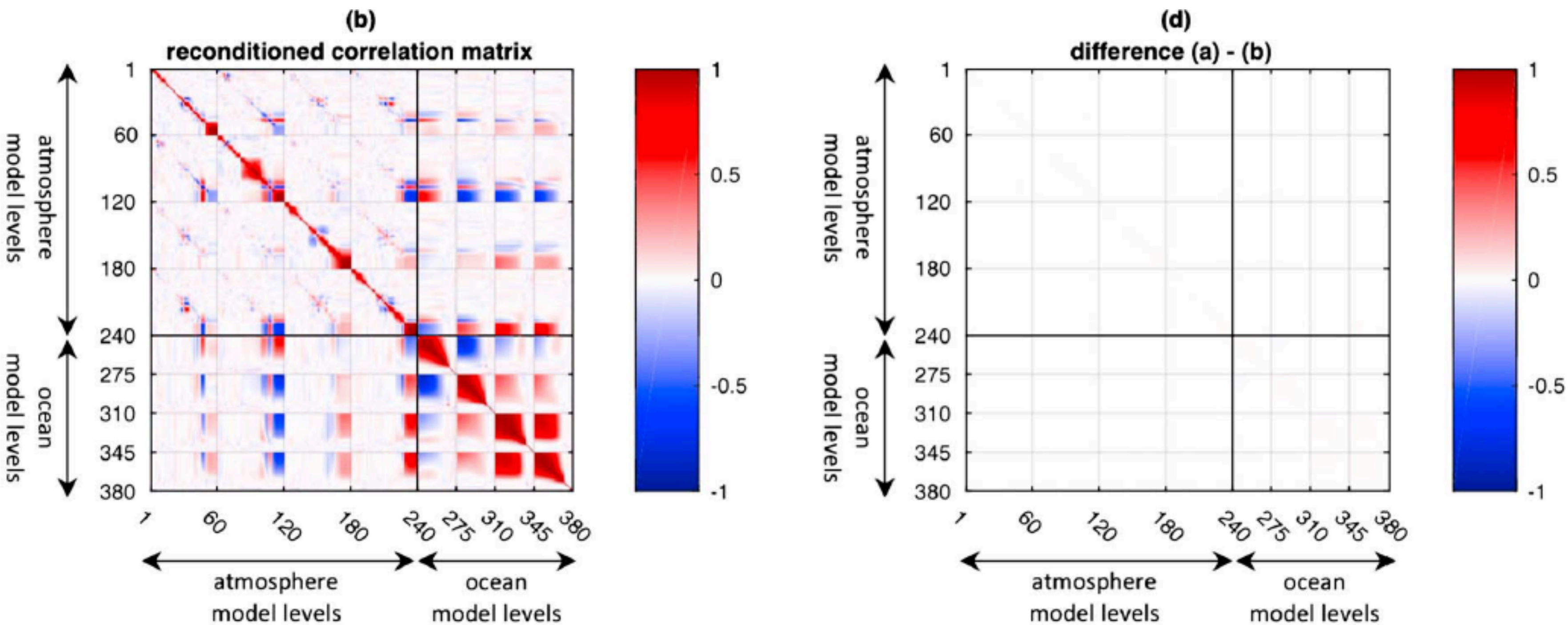
- specify a positive definite localization matrix $\boldsymbol{\rho}$ and form Schur product with the sampled ensemble correlation matrix $\mathbf{C}$
- since $\boldsymbol{\rho}$ is positive definite then the Schur product is positive definite and full rank
- the condition number will depend on the localization length scale
- localization can be applied to the covariance or correlation matrix with the same effect

Raw error correlation matrix



Reconditioned correlation matrix

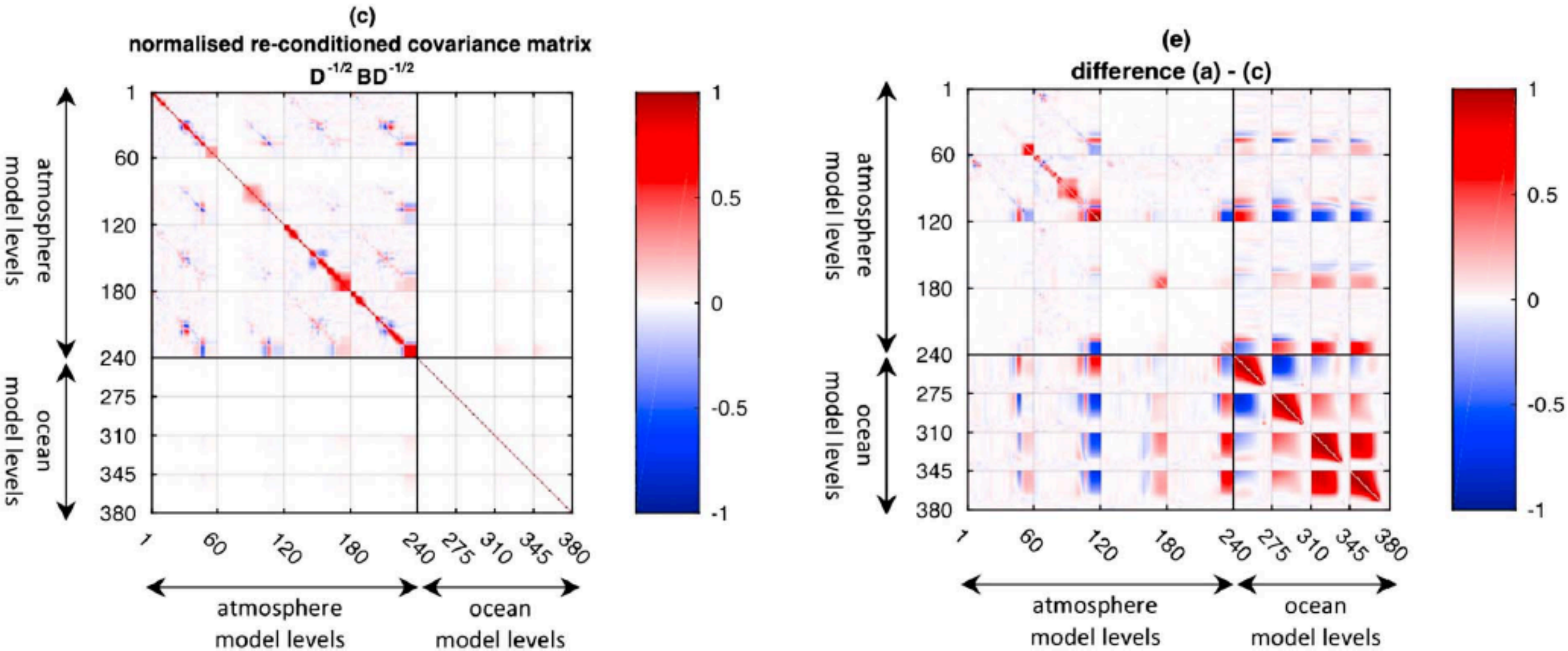
We recondition the correlation matrix to a target condition number of $\kappa_{\mathrm{tol}} = 10^4$



Reconditioning the correlation matrix reduces the condition number and keeps the correlation structure, but retains sample noise.

Reconditioned covariance matrix

We recondition the covariance matrix to a target condition number of $\kappa_{\mathrm{tol}} = 10^4$



Reconditioning the covariance matrix destroys correlation structure associated with smallest eigenvalues

Localization

- for coupled DA need to think carefully how to apply to 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

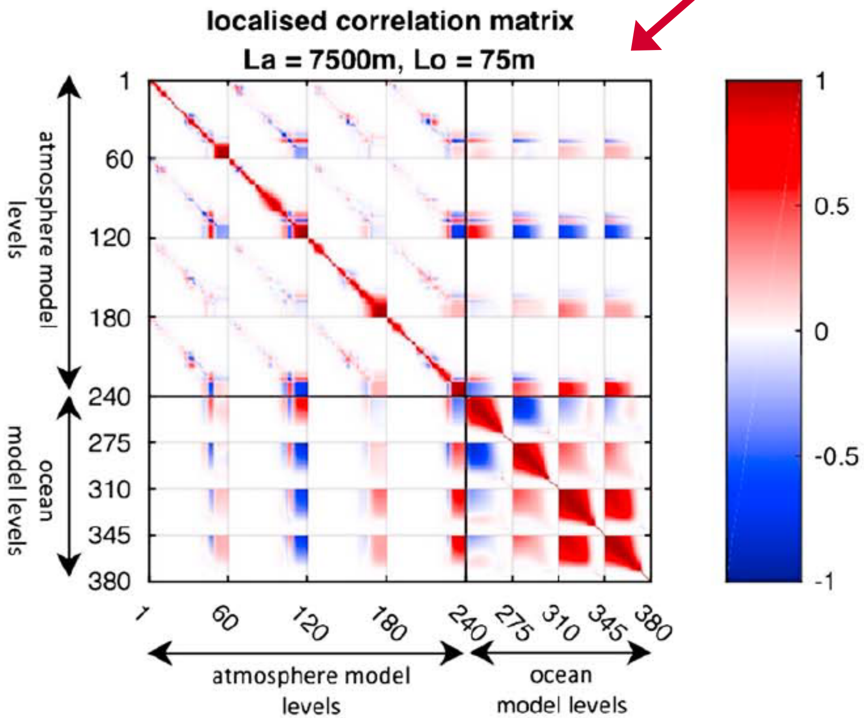
- we define a scaled distance between an atmosphere and ocean point, similar to Frolov et al. (2016)

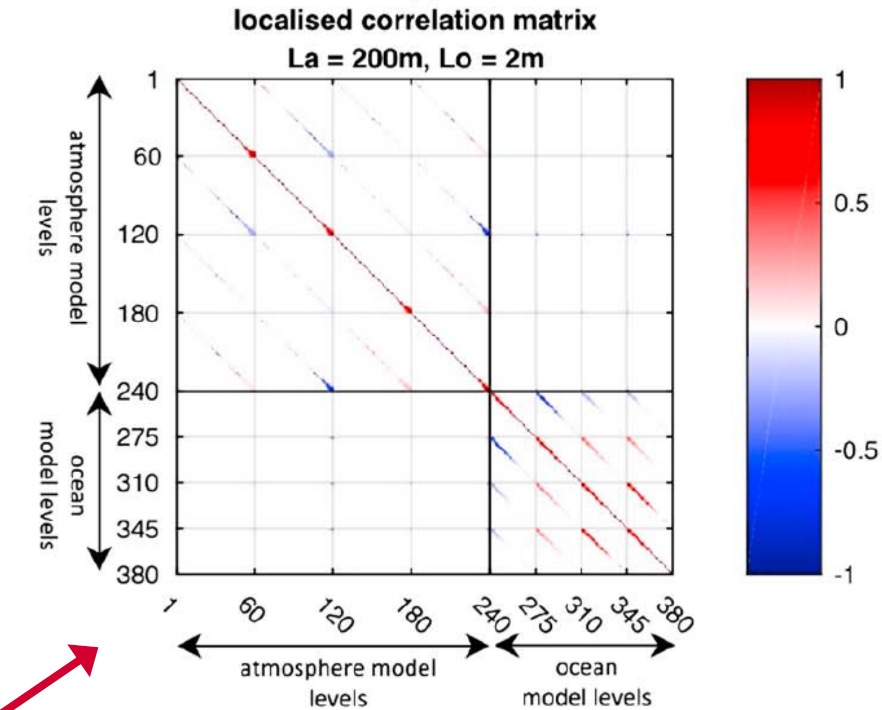
$$\hat{d}(z_a(i), z_o(j)) = \left(\frac{z_a(i)}{L_a} + \frac{z_o(j)}{L_o} \right)$$

- the same localization length scale is used for each sub-matrix in the atmosphere and same in the ocean; this helps keep the matrix positive definite

Results - localization

localization reduces sampling error, but retains high condition number of O(10^9).



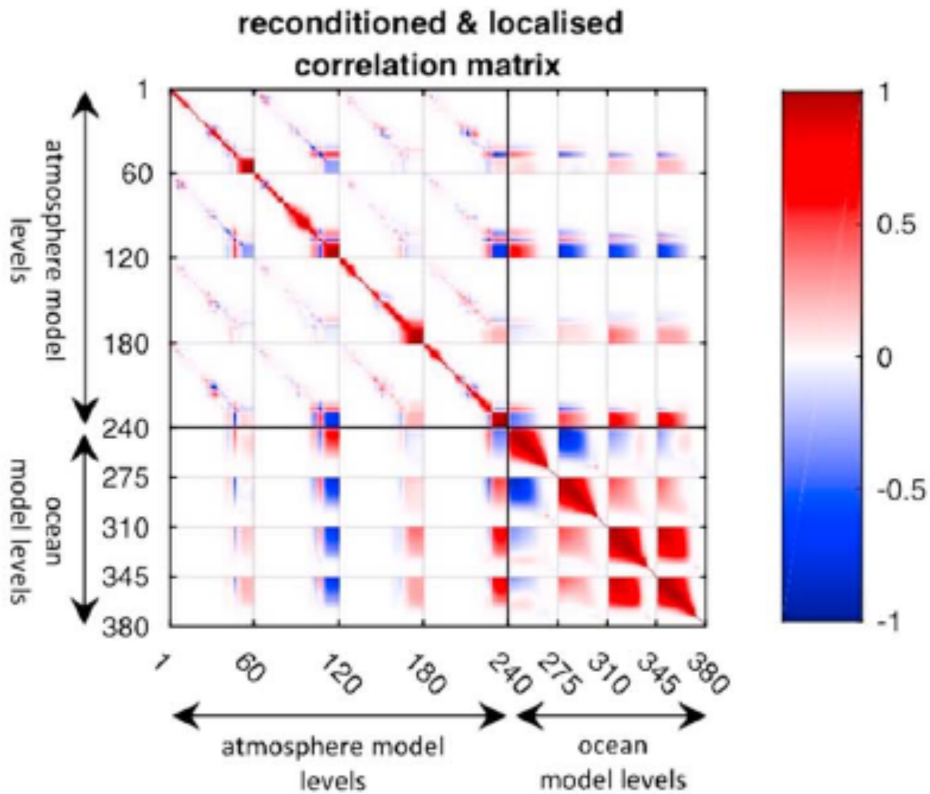


to obtain condition number O(10^4) we need to make lengthscales so short that correlations are destroyed.

Best of both worlds?

we first recondition using κ_{tol} = 10^4 and then localize with length-scales, L_a = 7500m and L_o=75m

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



Key points

- With a limited ensemble size sample forecast error correlations will be noisy, rank deficient and/or ill-conditioned.
- Reconditioning the 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.

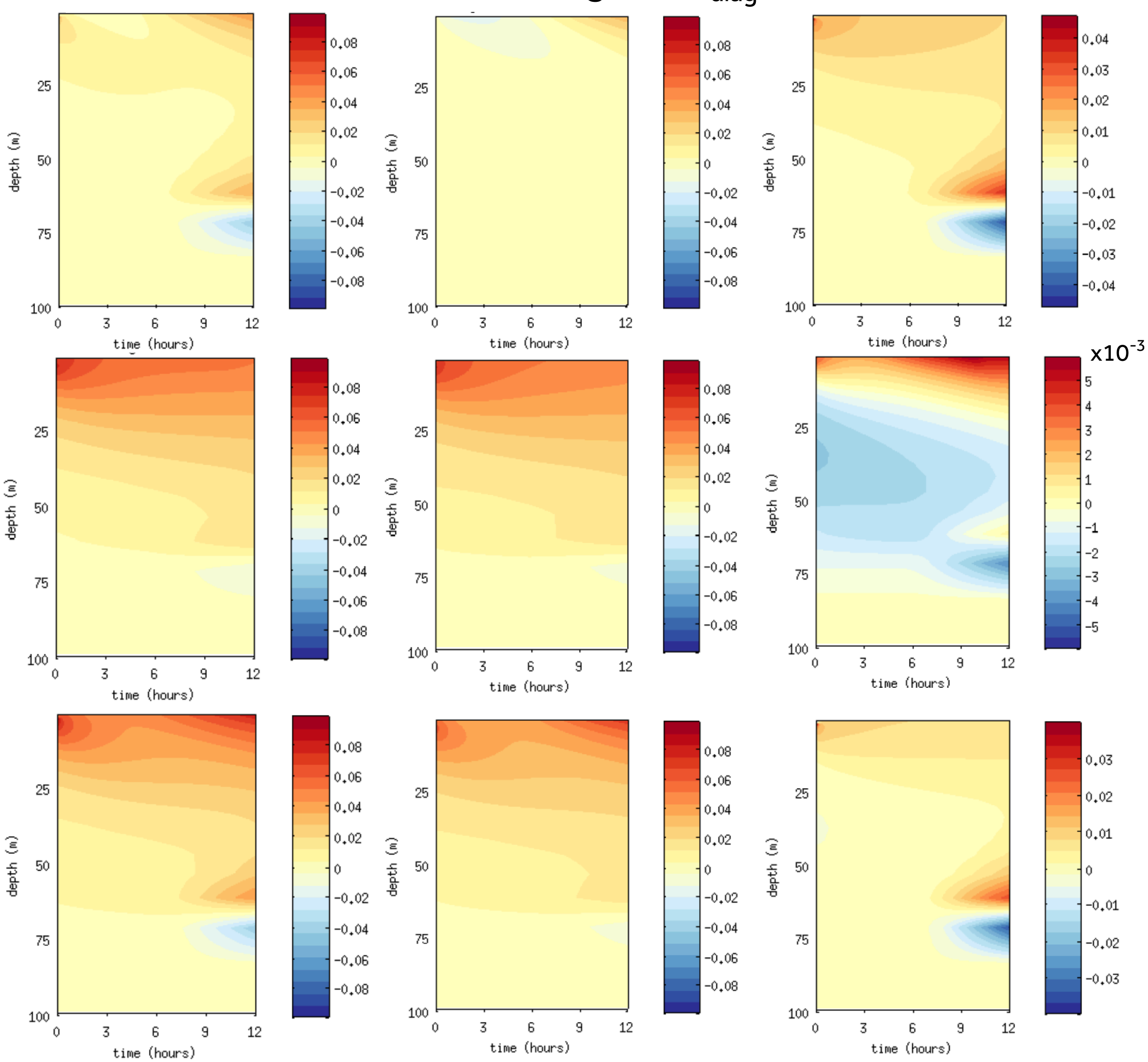
Idealised system experiments

- comparison of uncoupled, weakly coupled & fully coupled DA strategies
  - can coupled DA reduce or eliminate initialisation shock & improve balance at initial time?
- coupled forecast error covariance estimation
  - analysis-ensemble method
- coupled forecast error covariance implementation
  - reconditioning/ localization
  - single & double observation experiments (making greater use of near-surface observations)

Single & double observation exp

analysis increments: ocean temperature

(a) full **B** | (b) block diagonal $\mathbf{B}_{diag}$ | difference (a)-(b)



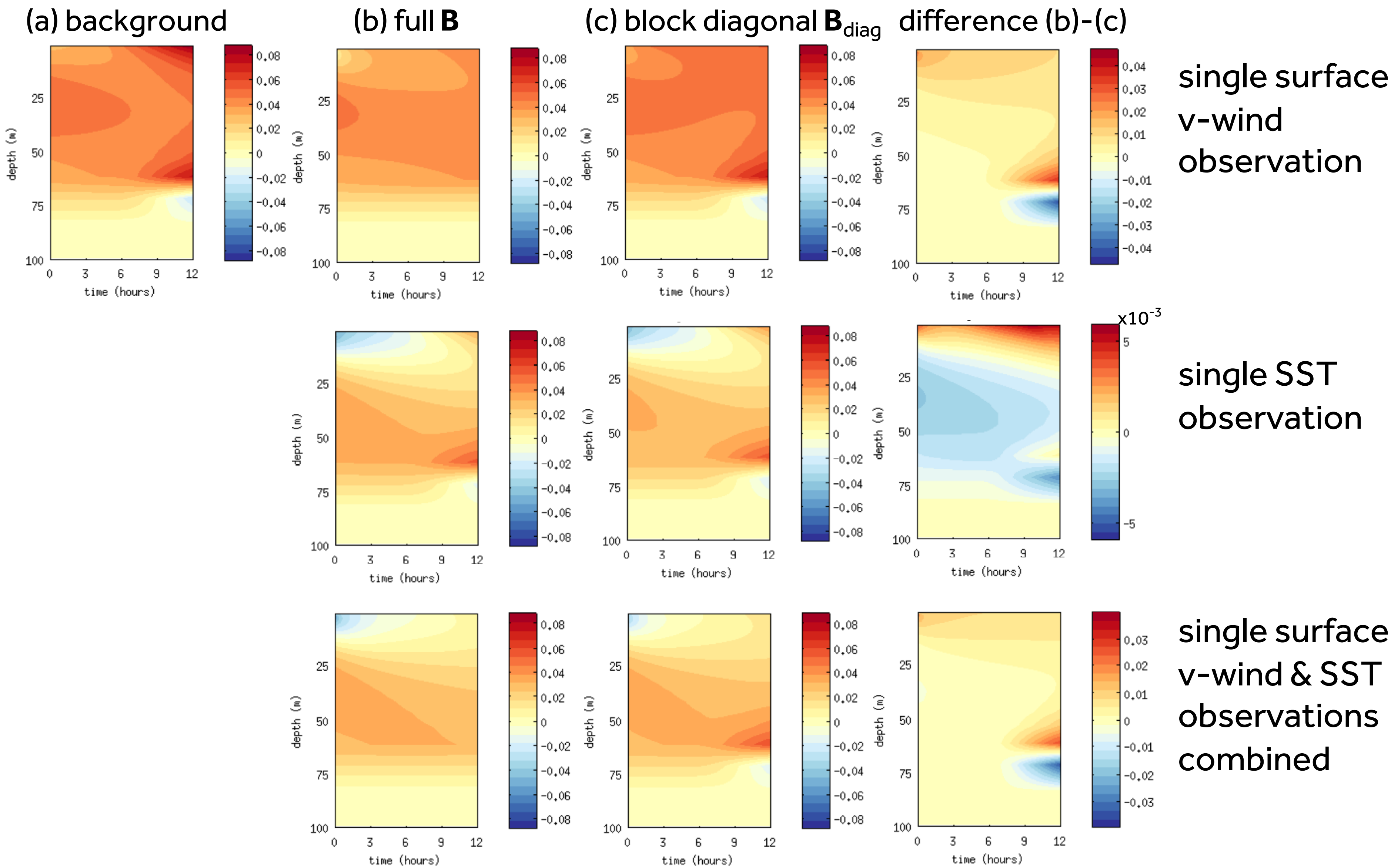
single surface v-wind observation (at end of 12hr window)

single SST observation

single surface v-wind & SST observations combined

Single & double observation exp

analysis errors: ocean temperature



Key points

- including explicit *a priori* cross-domain covariances:
  - improves information exchange across the modelled air-sea interface, leading to greater use of near-surface data
  - enables the ocean to have a direct influence on the structure of the initial atmosphere increments when only the atmosphere is observed (& vice versa)
  - when both atmosphere & ocean observations are assimilated, a full **B** allows the obs to work together to reduce errors and improve the consistency of the coupled analysis state
- when only one domain is observed, including explicit cross-domain forecast error covariances mostly impacts the unobserved domain
- unless the true cross-domain covariances are small using a fully coupled matrix **B** will always offer the greatest potential gains.

Summary

- used an idealised 1D system to explore some of the expected benefits of coupled atmosphere-ocean DA
- simple experiments show coupled DA is able to produce more balanced initial analysis fields
  - reduces initialisation shock and its impact on the subsequent forecast.
- In strongly coupled 4D-Var information exchange across the air-sea interface can be maximised by specification of *a priori* cross-domain forecast error covariances
  - we can use information from ensembles for this, but ...
  - sample covariances require modification before they can be incorporated into a standard assimilation framework
- it will take time for operational coupled DA systems to mature but our simple experiments confirm that there are clear benefits to be gained from a fully coupled approach.

University of Reading

References

1. Smith et al. (2018), Geophys. Res. Lett., doi: 10.1002/2017GL075534
2. Smith et al. (2017), Mon. Wea. Rev., doi: 10.1175/MWR-D-16-0284.1
3. Smith et al (2015), Tellus A, doi: 10.3402/tellusa.v67.27025