Mediating scale separation in Strongly Coupled Data Assimilation

Avneet Singh[1, 2, 3†], Alberto Carrassi[1, 2, 3], Francois Counillon[1, 2, 3]

[1] The Geophysical Institute, University of Bergen, Bergen 5007, Norway
[2] The Nansen Environmental and Remote Sensing Center, Thormøhlens gate 47, Bergen 5006, Norway
[3] Bjerknes Centre for Climate Research, University of Bergen, Bergen 5007, Norway

Abstract

Data Assimilation (DA) in a simple coupled system set-up is explored using linear and non-linear coupled toy models capturing the macroscopic properties of the ocean-atmosphere interactions. We especially concentrate on the effects of temporal scale separation between the oceanic and atmospheric sub-components, and its effects on the optimal implementation and possible modifications at the zeroth order to the typical DA procedures employed in weather forecasting and climate prediction.

1 Introduction

Data Assimilation (*abbrev*. DA), broadly speaking, is a conceptual and mathematical framework that aims to combine the information from observational datasets with the predictions from proposed model forms in order to yield a 'true' estimate of the state; the 'true' state in this case refers to the derived state that is a more accurate representation of the system than what is predicted by the model, or implied by the observations, independently. In some sense, DA could be interpreted as a propagation of an initial forecast state predicted by a given model towards higher likelihood by *assimilating* the observations. In context of complex earth systems, DA has been long utilised in geosciences, especially in meteorology and weather prediction [2], climate change [3], and more recently in attempting long-term climate prediction [7]. In this paper, we will concentrate on the application of DA to meteorology – often interpreted as accurate forecasts on short time-scales ($\lesssim 2$ weeks), and climate prediction – on time-scales longer than typical range covered by meteorology or weather prediction ($\gtrsim 2$ weeks). In such earth systems, the proposed models are typically dynamic, high-dimensional and qualitatively error-prone due to the inherent complexity in modelling the system. In particular, the dynamic nature of the models result in a discrete and sequential implementation of DA in time [2]; this essentially differentiates DA from a matched-filtering procedure where the model forms are deterministic and incorporate explicit time-dependence.

Theoretically speaking, the complexity – and accuracy – of DA for a given system is shared by the intrinsic complexity in the model form (e.g. model degrees of freedom, participating sub-systems, etc) as well as the nature of observations (e.g. the spatial and temporal scarcity of data points, accuracy of the observational datasets, etc). In practice, however, limitations on DA due to the nature of observational datasets are largely systematic and extrinsic, while the impact of the model form has a more fundamental and profound impact. In this respect, the case of coupled systems is extremely relevant and interesting since it presents with a realistic challenge where the model form of the system entails two or more sub-components with differing time-scales, e.g. ocean-atmosphere coupled earth system. This has a fundamental impact on the accuracy of DA on each observed sub-component on any intended time-scale and to varying magnitudes [7, 8]. In this paper, we will study the effects of multi-component coupling in the model form on the DA procedure using two toy models capturing the macroscopic properties of ocean-atmosphere interactions – one linear, and one non-linear and chaotic. The intention, in the end, is to develop a zero-order treatment of sequential coupled data assimilation and possible modifications to it in the linear as well as non-linear and chaotic regime.

2 Linear coupled model

The linear toy model $\mathbb{M}_{\text{lin}}$ for the first case study is adopted from Barsugli and Battisti [1], which is a simple one-dimensional, thermally coupled, purely temporal and stochastically forced atmosphere-ocean system of the form

$$\mathsf{m}\frac{d\mathsf{T_O}}{dt} = \mathsf{C_{OO}T_O} + \mathsf{C_{OA}T_A}, \tag{1}$$

$$\frac{d\mathsf{T_A}}{dt} = \mathsf{C_{AO}T_O} + \mathsf{C_{AA}T_A} + \mathsf{F}(t). \tag{2}$$

In matrix form, the coupled system may be written as

$$\boldsymbol{\nabla}\,\mathbf{T} = \mathbf{C}\,\mathbf{T} + \mathbf{F}, \text{ where,} \tag{3}$$

$$\boldsymbol{\nabla} = \begin{bmatrix} \mathsf{m}\dfrac{d}{dt} & \varnothing \\ \varnothing & \dfrac{d}{dt} \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} \mathsf{T_A} \\ \mathsf{T_O} \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} \mathsf{C_{OO}} & \mathsf{C_{OA}} \\ \mathsf{C_{AO}} & \mathsf{C_{AA}} \end{bmatrix}, \text{ and } \quad \mathbf{F} = \begin{bmatrix} \mathsf{F}(t) \\ \varnothing \end{bmatrix}. \tag{4}$$

where $\mathbf{T} = [\mathsf{T_O}, \mathsf{T_A}]$ is the state vector of the temperature anomalies in the ocean and atmosphere respectively, matrix $\mathbf{C}$ encodes the exchange of information between the two sub-components, 'm' encodes the temporal scale separation between the two sub-components, $\boldsymbol{\nabla}$ is the

†avneet.singh@uib.no

translation operator and $\mathsf{F}(t)$ is the stochastic forcing term. We choose the members of $\mathbf{C}$ and the value of scale separation 'm' such that the system is stable and in dynamic thermal equilibrium, e.g. $\mathsf{C_{OO}} = -0.1$, $\mathsf{C_{OA}} = 0.1$, $\mathsf{C_{AO}} = 0.01$ and $\mathsf{C_{AA}} = -0.1$, and ocean is relatively slow evolving ($\mathsf{m} > 1$). The stability of the system is ensured by negative eigenvalues of $\mathbf{C} \leftrightarrow |\mathbf{C}| > 0 \leftrightarrow \mathbf{C}$ is positive definite, which also ensures that the Lyapunov exponents are negative. The presence of the forcing term ensures that the model is dynamically spun up to mimic realistic scenarios and its cross-section σ_{F} is chosen accordingly. In figure 1, we show a sample time series for a single member simulation.

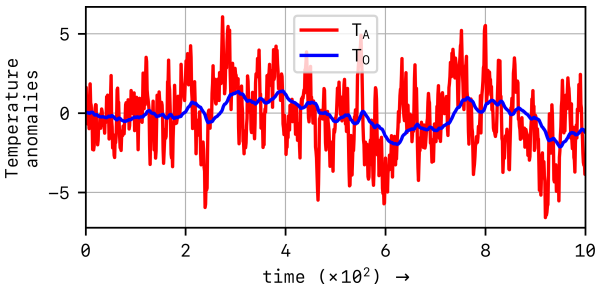

Figure 1: 1-member simulation for $\mathsf{m} = 10$, $\sigma_{\mathsf{F}} = 1.0$, and $\mathsf{T}_{\mathsf{O}}^{t=0} = \mathsf{T}_{\mathsf{A}}^{t=0} = 0.0$.

One may easily calculate the frequency spectrum of the solutions ($t \longrightarrow f$) in the Fourier space analytically in this case (5), to quantify the relative slowness of the ocean with respect to the atmosphere; this is shown in figure 2 – the gap between the Fourier transforms $\hat{\mathsf{T}}_{\mathsf{O}}$ and $\hat{\mathsf{T}}_{\mathsf{A}}$ quantifies the slowness of the ocean relative to the atmosphere.

$$\hat{\mathsf{T}}_{\mathsf{O}}(f) = \sigma_{\mathsf{F}}^2 \frac{\mathsf{C_{OA}}}{(2\pi i f - \mathsf{C_{AA}})(2\pi i \mathsf{m} f - \mathsf{C_{OO}}) - \mathsf{C_{OA}}\mathsf{C_{AO}}}$$
$$\hat{\mathsf{T}}_{\mathsf{A}}(f) = \sigma_{\mathsf{F}}^2 \frac{2\pi i \mathsf{m} f - \mathsf{C_{OO}}}{(2\pi i f - \mathsf{C_{AA}})(2\pi i \mathsf{m} f - \mathsf{C_{OO}}) - \mathsf{C_{OA}}\mathsf{C_{AO}}} \tag{5}$$

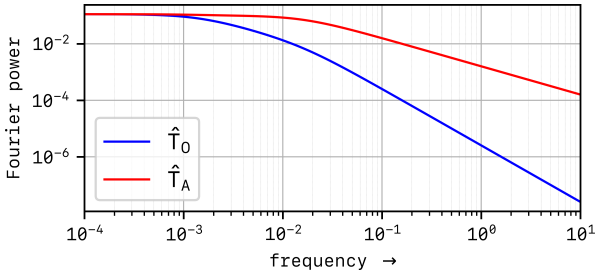

Figure 2: Fourier transforms $\hat{\mathsf{T}}_{\mathsf{O}}$ and $\hat{\mathsf{T}}_{\mathsf{A}}$ for $\mathsf{m} = 10$, $\sigma_{\mathsf{F}} = 1.0$.

Lastly, this temporal scale separation (i.e. slowness of the ocean) leads to a net lagging effect on the ocean which is quantifiable by calculating the auto- and cross-correlation functions (6).

$$\langle \mathsf{T_X}, \mathsf{T_Y} \rangle \equiv \langle \mathsf{T_X}(t), \mathsf{T_Y}(t+\tau) \rangle = \int_{-\infty}^{+\infty} \mathsf{T_X}(t)\, \mathsf{T_Y}(t+\tau)\, \mathrm{d}t. \tag{6}$$

where τ is the *lag time*. In figure 3, we show the auto- and cross-correlation functions plotted for our linear model where the slowness of ocean is again reflected in shallower decline of the oceanic auto-correlation function $\langle \mathsf{T_O}, \mathsf{T_O} \rangle$ relative to the atmospheric auto-correlation function $\langle \mathsf{T_A}, \mathsf{T_A} \rangle$. Notably, the cross-correlation function $\langle \mathsf{T_O}, \mathsf{T_A} \rangle$ peaks at negative value of τ (denoted by τ_{p}) depicting that the ocean *lags* the atmosphere, or that the atmosphere *leads* the ocean. We intend to use this lag to our advantage in the data assimilation procedure. The statistics of $|\tau_{\mathsf{p}}|$ as a function of scale separation 'm' are shown in figure 4; we find that the modes of the distributions (i.e. values of $|\tau_{\mathsf{p}}|$) increase with increasing scale separation (i.e. 'm') as expected. Moreover, while $|\tau_{\mathsf{p}}|$ primarily depends on 'm' and $|\mathbf{C}|$, the higher moments (skewness, kurtosis, etc) depend on $\mathbf{C} - \mathsf{diag}(\mathbf{C})$.

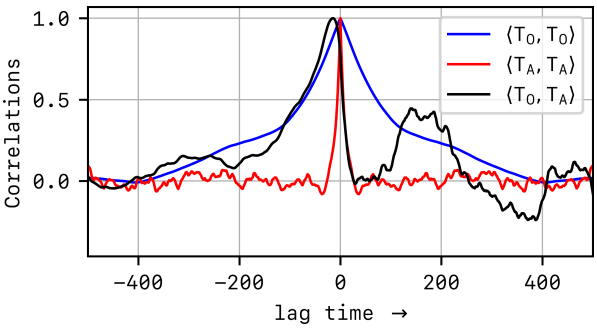

Figure 3: Cross- and auto-correlation functions $\langle \mathsf{T_X}, \mathsf{T_Y} \rangle$ for a 1-member simulation for $\mathsf{m} = 10$, $\sigma_{\mathsf{F}} = 1.0$, and $\mathsf{T}_{\mathsf{O}}^{t=0} = \mathsf{T}_{\mathsf{A}}^{t=0} = 0.0$. Note that the auto-correlation functions naturally peak at $\tau = 0$ while the cross-correlation peaks at $\tau < 0$ depicting the *lagging effect* on the ocean.

2.1 DA on the linear model

In order to study and optimise data assimilation using this coupled linear model $\mathbb{M}_{\mathsf{lin}}$, we utilise the existing framework of sequential *Ensemble Kalman Filter* (*abbrev.* EnKF); it is a well known fact that in the linear case, EnKF is the most optimal choice for data assimilation considering the analytical constraints as it equates to a unique and global minimisation of the standard variational approach [2]. If the model error covariance matrix is denoted by $\mathbf{P}$, the observational error covariance matrix by $\mathbf{R}$, then the Kalman Gain $\mathbf{K}$ at any time t is written as

$$\mathbf{K}_t = \mathbf{P}_t \left[\mathbf{P}_t + \mathbf{R}_t\right]^{-1}. \tag{7}$$

The resulting analysed and forecast state at any time t are further written as

$$\mathbf{T}_t^{\mathsf{ana}} = \mathbf{T}_t^{\mathsf{for}} + \mathbf{K}_t \left[\mathbf{T}_t^{\mathsf{obs}} - \mathbf{T}_t^{\mathsf{for}}\right], \tag{8}$$

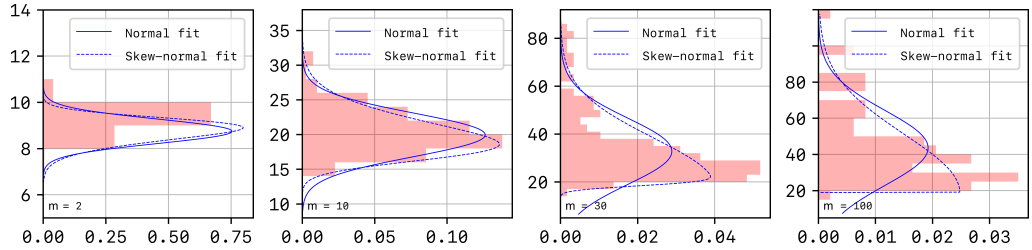

Figure 4: Normalised probability densities (*horizontal axes*) of peak lag times $|\tau_p|$ (*vertical axes*) as a function of scale separation 'm'; $\sigma_F = 1.0$, and $T_O^{t=0} = T_A^{t=0} = 0.0$. The solid lines denote normal distribution fits while the dashed lines denote skew-normal fits. Note the shift in the mode, the skewness and the spread of the distributions as a function of increasing values of $m \in \{2, 10, 30, 100\}$.

where $\mathbf{T}_t^{obs}$ and $\mathbf{T}_t^{for}$ are the observation and forecast vectors. The analysed model error covariance matrix evolves as $\mathbf{P}_t^{ana} = \mathbf{P}_t^{for}[\mathbf{I} - \mathbf{K}_t]$ although this calculation is not required explicitly by the ensemble version of the Kalman filter. The analysed state $\mathbf{T}_t^{ana}$ is finally translated in time from t to $t+1$ by the translation operation, $\mathbf{T}_{t+1}^{for} = \mathbb{M}_{lin} \cdot \mathbf{T}_t^{ana}$, by numerically integrating the coupled stochastic differential equations[1] (1)–(2). In our experiments, we reasonably assume the observation errors in the ocean and the atmosphere to be mutually uncorrelated and constant in time[2], i.e.

$$\mathbf{R}_t = \begin{bmatrix} \sigma_{A,\,obs} & \varnothing \\ \varnothing & \sigma_{O,\,obs} \end{bmatrix}_{\forall t}, \tag{9}$$

while the model error covariance matrix $\mathbf{P}_t$ is explicitly written as

$$\mathbf{P}_t = \begin{bmatrix} \langle \vec{T}_A^{for} \mid \vec{T}_A^{for} \rangle & \langle \vec{T}_A^{for} \mid \vec{T}_O^{for} \rangle \\ \langle \vec{T}_O^{for} \mid \vec{T}_A^{for} \rangle & \langle \vec{T}_O^{for} \mid \vec{T}_O^{for} \rangle \end{bmatrix}_t. \tag{10}$$

The *true* observation vectors, namely $\mathbf{T}_t^{true}$, for the ocean and the atmosphere in our experiments are generated by perturbing a 1-member simulation of the model $\mathbb{M}_{lin}$ with $\sigma_{O,\,obs}$ and $\sigma_{A,\,obs}$ respectively. Lastly, the Kalman Gain matrix $\mathbf{K}_t$ is explicitly written as

$$\mathbf{K}_t = \frac{1}{D_t} \begin{bmatrix} \langle \vec{T}_A^{for} \mid \vec{T}_A^{for} \rangle [\langle \vec{T}_O^{for} \mid \vec{T}_O^{for} \rangle + \sigma_{O,\,obs}^2] - \langle \vec{T}_O^{for} \mid \vec{T}_A^{for} \rangle^2 & \sigma_{O,\,obs}^2 \langle \vec{T}_O^{for} \mid \vec{T}_A^{for} \rangle \\ \sigma_{A,\,obs}^2 \langle \vec{T}_A^{for} \mid \vec{T}_O^{for} \rangle & \langle \vec{T}_O^{for} \mid \vec{T}_O^{for} \rangle [\langle \vec{T}_A^{for} \mid \vec{T}_A^{for} \rangle + \sigma_{A,\,obs}^2] - \langle \vec{T}_A^{for} \mid \vec{T}_O^{for} \rangle^2 \end{bmatrix}_t, \tag{11}$$

where $D_t = [\langle \vec{T}_A^{for} \mid \vec{T}_A^{for} \rangle + \sigma_{A,\,obs}^2][\langle \vec{T}_O^{for} \mid \vec{T}_O^{for} \rangle + \sigma_{O,\,obs}^2] - \langle \vec{T}_O^{for} \mid \vec{T}_A^{for} \rangle^2$, and $\vec{\cdot}$ represents the ensemble vector of length N_e, the time-varying spread of the ensembles is denoted by $\sigma_{A/O,\,for/ana}$, and the $\langle \cdot | \cdot \rangle$ operation represents inner product. Naturally, $\langle \vec{T}_O^{for} \mid \vec{T}_A^{for} \rangle = \langle \vec{T}_A^{for} \mid \vec{T}_O^{for} \rangle$ and that $\langle \vec{T}_X^{for} \mid \vec{T}_Y^{for} \rangle$ is generally equivalent to the discrete formulation of the zero-lag correlation function $\langle T_X, T_Y \rangle_{\tau=0}$ in (6). This completes the required ingredients for EnKF. In table 1, we list the values of the parameters defined above.

We quantify the performance of DA by calculating the RMSE vector defined by

$$\boldsymbol{\sigma}_{RMSE} = \sqrt{\sum_{\forall t} |\mathbf{T}_t^{true} - \mathbf{T}_t^{ana}|^2}. \tag{12}$$

In standard large-scale data assimilation applications, a *strongly* coupled approach (*abbrev.* SCDA) as outlined above is not yet adopted although its application has

Parameter	*Value*	*Parameter*	*Value*
m	$\in [1, 11]$	τ	$\in [1, 41]$
$T_O^{t=0}$	0.0	$T_A^{t=0}$	0.0
$\sigma_{O,\,for}^{t=0}$	1.0	$\sigma_{A,\,for}^{t=0}$	2.0
$\sigma_{O,\,obs}$	1.0	$\sigma_{A,\,obs}$	2.0
σ_F	1.0	N_e	25

Table 1: Ranges and values of the described parameters. Note that the largest value of 'm' explored in our simulations is dictated by the cross-section of the stochastic forcing i.e. σ_F; for $\sigma_F = 1.0$, values of $m \gtrsim 11$ result in significant (and dominant) dissipation in the model $\mathbb{M}_{lin}$ – this has an adverse impact on the testability of EnKF due to the *collapse* of the forecast ensemble.

[1]The expression (3) is of the *Ito* form and its numerical integration is performed using sdeint.itoint algorithm available in Python's SciPy library. More specifically, sdeint.itoint uses the Euler-Maruyama algorithm to integrate the Ito equation.

[2]We use the explicit form of $\mathbf{R}$ in our simulations instead of the ensemble form for simplicity. Note that each ensemble member receives its unique perturbed version of the truth defined by $\mathbf{R}$.

been encouraged in recent times [6]. The current adoption of *weakly* coupled implementation (*abbrev.* WCDA) simply involves setting the off-diagonal terms in $\mathbf{P}_t$ to zero, thereby explicitly ignoring the model error covariances in the analysis step and only allowing the coupled system to share information during the forecast step; this approach is prone to instabilities.

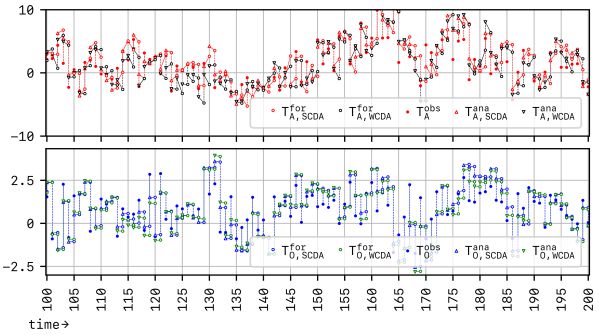

Figure 5: One random member EnKF simulation for m = 6 showing the WCDA and the SCDA implementations. Top panel (*red*) shows the atmospheric component while the bottom panel depicts the ocean (*blue*).

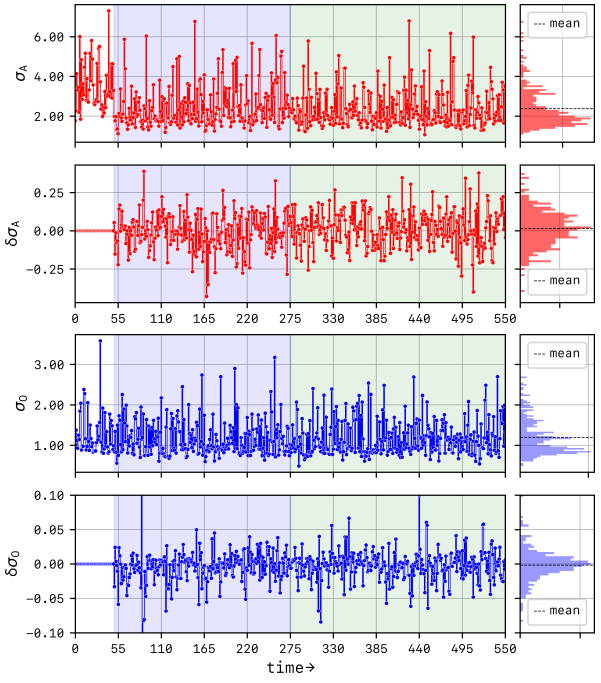

Figure 6: RMSE statistics derived from a 25–member EnKF simulation of the SCDA ($\boldsymbol{\sigma}_{\mathsf{A/O}}$), and the difference between the WCDA and the SCDA implementation ($\delta\boldsymbol{\sigma}_{\mathsf{A/O}}$); m = 6. The model is spun up to stability during the *free run* in time $t \in [0, 50]$ whereas the RMSE means are calculated for time $t \in (275, 550]$. In this particular experiment, SCDA seems to perform marginally better in the slower ocean component than WCDA although this discrepancy is much smaller than the climatological background error due to the stochastic forcing term and should not be over-emphasised.

In order to show the difference between the two, we run WCDA simulations in parallel with SCDA in all our experiments and calculate $\delta\boldsymbol{\sigma}_{\mathsf{RMSE}} = \boldsymbol{\sigma}^{\mathsf{SCDA}}_{\mathsf{RMSE}} - \boldsymbol{\sigma}^{\mathsf{WCDA}}_{\mathsf{RMSE}}$. In figure 5, we show a snapshot of a random ensemble member's simulation while figure 6 shows the corresponding RMSE statistics derived from a 25–member EnKF simulation by fixing m = 6.

Inflation

We note that the spread of the ensemble in our simulation contracts over time since the model error covariance matrix evolves as $\mathbf{P}^{\mathsf{ana}}_t = \mathbf{P}^{\mathsf{for}}_t\,[\mathbf{I} - \mathbf{K}_t]$ besides the usual variation due to the model flow [2]. This leads to convergence issues and degradation in the performance of DA over time; we counter this effect by incorporating a constant multiplicative inflationary factor of the ensemble for the SCDA simulation, denoted by γ, applied to both the oceanic as well as the atmospheric ensemble.

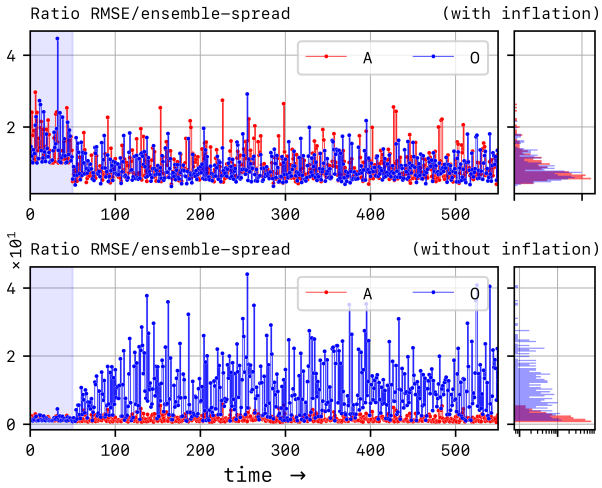

Figure 7: The convergence of SCDA is probed by calculating the ratio between the RMSE and the spread of the ensemble for each sub-component over time t for m = 6. The criteria of convergence requires that this ratio tends to $\lesssim 1$ in the limit $t \to \infty$. We find that inflation is necessary to achieve convergence for both the SCDA and the WCDA (not shown). Note that the shaded regions denote the duration of the free run that has been ignored in the histograms on the right.

The values of the constant multiplicative inflationary factor γ is shown in figure 8. It is noteworthy that the constant multiplicative inflation counters the deflation of ensemble due to the dissipative nature of the model. Thus, it follows that γ must depend on the choice of 'm' in the model. In particular, γ must decrease monotonically with increasing value of 'm'; we ran a tuning simulation to optimise the choice of γ_{m} and its results are shown in figure 8.

In addition, figures 7 and 9 show the detailed properties of the ensemble for the curious reader. Note the impact of constant multiplicative inflation γ on preserving the spread of the ensemble in figure 9 (top attached panels).

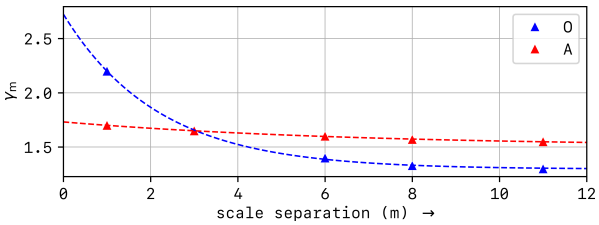

Figure 8: Calculated values of γ_m are shown on the vertical axes for different choices of of scale separation on the horizontal axes. We fit the data points with decaying exponential functional forms, marked by dashed (−−) lines, in order to derive a generalised analytical expression; using this formulae, we calculate the optimal choices of γ_m for any given 'm'. Note that both the WCDA and the SCDA implementations are inflated by the same value of γ_m.

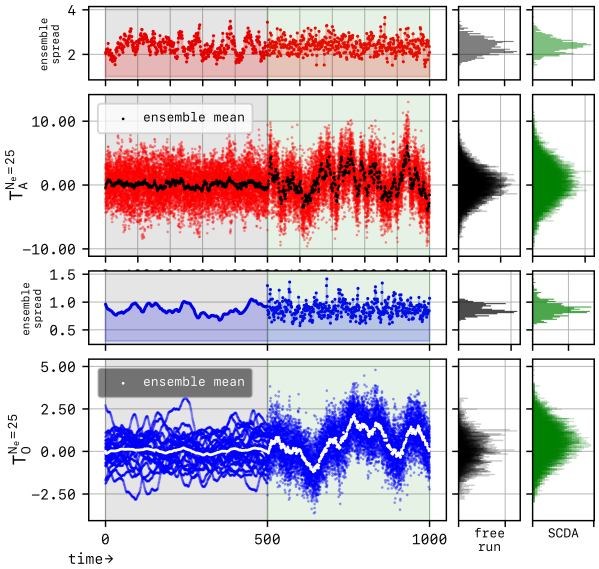

Figure 9: Detailed properties of the ensembles and their spreads in the ocean and the atmosphere for $m = 6$ and $t \in [0, 1000]$; note that the SCDA starts from $t > 500$ (shaded green in the central plot) whereas the free run corresponds to $t \leq 500$ (shaded grey). Moreover, the colors of the histograms correspond to the shaded regions in the central plot from where the statistics are derived from, e.g. grey histogram is derived from the statistics gathered during the free run whereas the green histogram corresponds to statistics from the SCDA part of the simulation.

2.2 Leading averaged cross-covariance

In order to maximise the effects of cross-covariance in coupled data assimilation, Lu *et al* [4, 5] suggested that it is advantageous to use leading forecasts and observations in the fast sub-component, i.e. the atmosphere, while updating the slower sub-component, i.e. the ocean, instead of assimilating the instantaneously available atmospheric observation at any time t. This suggestion simply follows from figure 3 where we found that the cross-correlation function is maximum when atmosphere leads the ocean, i.e. when $\tau < 0$. This essentially means that the information shared between the oceanic and the atmospheric sub-components in the cross-update may be optimised further by choosing $\tau < 0$, thus maximising the signal-to-noise ratio (SNR) in the cross-update. In order to further minimise the effects of noise arising due to stochastic forcing, Lu *et al* [4] used an *average* over some τ_{win} number of leading observations in the atmosphere to update the ocean; they termed this implementation as LACC [4, 5]. Note that no modification was necessary in reverse, i.e. ocean → atmosphere, since the faster sub-component doesn't benefit from any such modification directly.

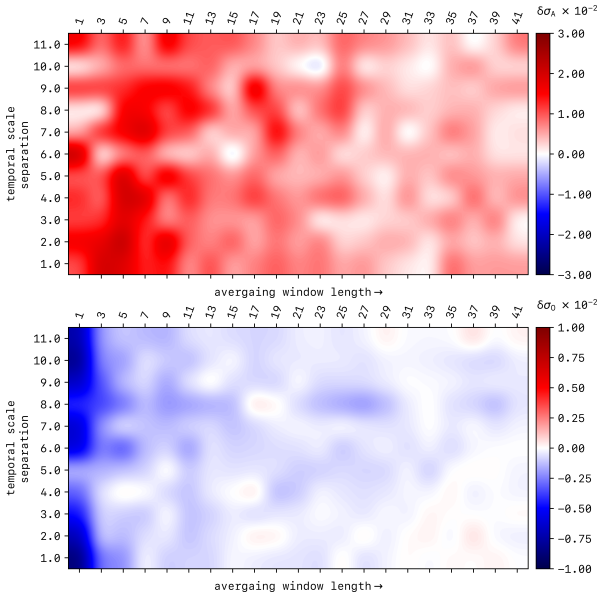

Figure 10: RMSE statistics derived from a 25–member EnKF simulation as a function of 'm' (*vertical axes*) and τ_{win} (*horizontal axes*) showing the difference between the WCDA and the SCDA/LACC implementations captured by $\delta\boldsymbol{\sigma}_{A/O}$; the bottom panel shows the oceanic component while the top panel depicts the atmosphere. Note that the SCDA case depicted in figures 5–6 corresponds to $\tau_{\text{win}} = 1$ in this plot. In other words, $\text{LACC}_{[\tau_{\text{win}}=1]} \equiv \text{SCDA}$.

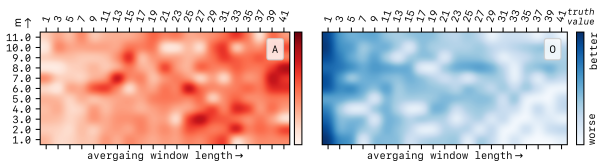

Figure 11: RMSE statistics from figure 10 mapped onto a log-scaled ordinal *truth value* space, i.e. *better* or *worse* than the WCDA. In this context, large positive and small negative values are mapped onto 'worse' while large negative values are mapped onto 'better'; the left panel shows the atmospheric component while the right panel depicts the ocean. The log-scale of the color-map is intended to highlight the local variations.

In this paper, we will attempt to quantify the effects of choices of τ_{win} on $\delta\boldsymbol{\sigma}_{\text{RMSE}}$. It must be noted that the optimal choice of τ_{win} must depend on 'm' since

$|\tau_p|$ depends on 'm'. Consequently, the results of the LACC implementation compared with the WCDA and the SCDA are shown in figure 10 as a function of 'm' and τ_{win}. In principle, we expect that for larger scale separations, correspondingly longer leading averaged windows should maximise the cross-component SNR, thus leading to a larger disparity between the LACC and the WCDA implementations. However, we note that longer window lengths also decrease the frequency of the coupled cross-updates since this update is made only when preceding τ_{win} observations in the atmosphere become available, which contribute toward diminishing returns of LACC as seen in figure 10. Overall, we find that the LACC and the SCDA methods perform better than the standard WCDA implementation in the ocean whereas the atmosphere almost always suffers when the coupled cross-updates are included.

3 Extension of LACC to non-linear case

bla bla bla

References

[1] J J Barsugli and D S Battisti. *Strongly Coupled Data Assimilation Using Leading Averaged Coupled Covariance (LACC). Part II: CGCM Experiments*. ***J. Atmos. Sci*** , 55(–):477–493, 1998.

[2] A Carrassi *et al*. *Data assimilation in the geosciences: An overview of methods, issues, and perspectives*. ***WIREs Climate Change*** , 9(5):e535, 2018.

[3] A Hannart *et al*. *DADA: data assimilation for the detection and attribution of weather and climate-related events*. ***Climatic Change*** , 136(2):155–174, 2016.

[4] F Lu *et al*. *Strongly Coupled Data Assimilation Using Leading Averaged Coupled Covariance (LACC). Part I: Simple Model Study*. ***Mon. Wea. Rev.*** , 143(–):3823–3837, 2015.

[5] F Lu *et al*. *Strongly Coupled Data Assimilation Using Leading Averaged Coupled Covariance (LACC). Part II: CGCM Experiments*. ***Mon. Wea. Rev.*** , 143(–):4645–4659, 2015.

[6] S G Penny and T M Hamill. *Strongly Coupled Data Assimilation for Integrated Earth System Analysis and Prediction*. ***Bull. Amer. Meteor. Soc.*** , 98 (–):ES169–ES172, 2017.

[7] S G Penny *et al*. *Data assimilation in the geosciences: An overview of methods, issues, and perspectives*. ***Journal of Advances in Modeling Earth Systems*** , 11(6):1803–1829, 2019.

[8] M Tondeur *et al*. *On temporal scale separation in coupled data assimilation with the ensemble Kalman filter*. ***submission in progress*** , –(–): –, 2019.