On Temporal Scale Separation in Coupled Data Assimilation with the Ensemble Kalman Filter

Maxime Tondeur[1] · Alberto Carrassi[2,3,4] · Stephane Vannitsem[5] · Marc Bocquet[6]

Abstract

Data assimilation for systems possessing many scales of motions is a substantial method-ological and technological challenge. Systems with these features are found in many areas of computational physics and are becoming common thanks to increased computational power allowing to resolve finer scales and to couple together several sub-components. Coupled data assimilation (CDA) distinctively appears as a main concern in numerical weather and climate prediction with major efforts put forward by meteo services worldwide. The core issue is the scale separation acting as a barrier that hampers the propagation of the infor-mation across model components (e.g. ocean and atmosphere). We provide a brief survey of CDA, and then focus on CDA using the ensemble Kalman filter (EnKF), a widely used Monte Carlo Gaussian method. Our goal is to elucidate the mechanisms behind information propagation across model components. We consider first a coupled system of equations with temporal scale difference, and deduce that: (i) cross components effects are strong from the slow to the fast scale, but, (ii) intra-component effects are much stronger in the fast scale. While observing the slow scale is desirable and benefits the fast, the latter must be observed with high frequency otherwise the error will grow up to affect the slow scale. Numerical experiments are performed using the atmosphere-ocean model, MAOOAM. Six configura-tions are considered, differing for the strength of the atmosphere-ocean coupling and/or the number of model modes. The performance of the EnKF depends on the model configuration, i.e. on its dynamical features. A comprehensive dynamical characterisation of the model configurations is provided by examining the Lyapunov spectrum, Kolmogorov entropy and Kaplan–Yorke attractor dimension. We also compute the covariant Lyapunov vectors and use them to explain how model instabilities act on different model's modes according to the coupling strength. The experiments confirm the importance of observing the fast scale, but show also that, despite its slow temporal scale, frequent observations in the ocean are beneficial. The relation between the ensemble size, N, and the unstable subspace dimension, n_0, has been studied. Results largely ratify what known for uncoupled system: the condition $N \geq n_0$ is necessary for the EnKF to work satisfactorily. Nevertheless the quasi-degeneracy of the Lyapunov spectrum of MAOOAM, with many near-zero exponents, is potentially the cause of the smooth gradual reduction of the analysis error observed for some model con-

figurations, even when $N > n_0$. Future prospects for the EnKF in the context of coupled ocean-atmosphere systems are finally discussed.

Keywords Data assimilation · Ensemble Kalman filter · Multiscale · Lyapunov

1 Introduction

Data assimilation (DA) is the term used to refer to a broad family of conceptual, mathematical and numerical methods performing the combination of model solutions with data from observing devices of a system. A popular terrain of applications of DA, and one that distinguishes DA from more general classes of optimisation or filtering methods, is its widespread use in the context of chaotic dynamics. The primary goal is to get an estimate of the system's state that is more accurate than the ones given by model and data independently (see e.g., [1]). Data assimilation posed its root in the geosciences, particularly meteorology, but its use is becoming widespread across other areas of the geosciences and beyond [11]. Examples include, but are not limited to, the attribution of climate change [25], neuro- (e.g. [32,40]) and life-sciences [30] or traffic management [43].

This work is about DA for systems possessing a wide range of spatial and temporal scales, in particular coupled dynamical systems, in which the typical temporal scales of the system's components are different and generally not overlapping. This situation is common in physical science and it arises when modelling a continuum system in high resolution. Similarly when the system is modelled by coupling together different sub-systems each one spanning its own band of spatio-temporal scales. Notable examples are the climate models that couple together the different components of the Earth system. But it is nowadays present in other domains, including computational biology, neuroscience or economy, wherever the today's enhanced computational power allows to explicitly couple several sub-systems at an almost constantly increasing resolution.

Together with the increase in models' resolution, computational geosciences in the last decade has also seen the increase of prediction lengths beyond the meteorological time horizon of two weeks. Seasonal-to-decadal (s2d) forecasts, a time horizon bearing enormous societal relevance, are possible because predictable signals arise from the interaction between the fast (e.g., the atmosphere) and the slow (e.g., the ocean, the land surface or the cryosphere) varying components of the system [18]. In the range between weather and s2d predictions stands the sub-seasonal to seasonal (s2s) time range, which corresponds to predictions from two weeks to a season [10]. Sub-seasonal to seasonal predictions have motivated the ongoing transition toward the so-called "seamless" weather/climate prediction, in which the same fully coupled climate model is used to predict from minutes to months (see e.g., [9,44]).

Coupled data assimilation (CDA) is needed to enhance the predictive skill of phenomena connected to the air–sea exchange like hurricanes or coastal weather, or in s2d predictions where climate conditions are triggered by coupled processes such as ENSO. The development of efficient CDA methods has been identified as crucial already in the assessment report of the 5th Intergovernmental Panel on Climate Change, and several research institutions, are involved in studying and developing CDA (see, e.g., [49]).

From a DA perspective, the main issue is that the scale separation renders it extremely difficult to carry out the uncertainty quantification necessary to propagate consistently the information from the observations in one component throughout the full system. If the scale separation is not very large, one can still rely upon standard, uncoupled DA that operates on

each component independently, and then use the full coupled model to propagate information between successive observations, an approach known as *weakly CDA* (wCDA). Although in wCDA the effect of the coupling manifests indirectly via the model forward integration, the cross-component physical correlations (if any) are not exploited in the analysis update. Such a procedure is thus prone to produce imbalances, and a fully CDA (usually referred to as *strongly CDA*, sCDA) is required.

We focus on sequential DA methods, such as the ensemble Kalman filter (EnKF, [20]), with the aim of elucidating the nature of the problem. We will develop our discussion in relation to the geosciences, where the field has been pushed forward. However none of the results nor of the conclusions are restricted to that context exclusively.

1.1 Coupled Data Assimilation in the Geosciences: Brief Survey

We provide here a succinct survey of CDA efforts in the geosciences that is functional to our discussion. Recent reviews of CDA can be found in [47–49]. In particular [48] provides a detailed comparison of different DA methods using the same low dimensional coupled model used here (see Sect. 3)

Early attempts include a Kalman filter type approach used to assimilate sparse data in a system with various spatio-temporal scales [26], and the EnKF in a two-scale low dimensional system [3]. A modification of the four-dimensional variational assimilation (4DVar; see e.g. [11] its Sect. 3.2 and references therein) for coupled dynamics is given in [37], although the approach presented a number of practical issues making difficult its application in an operational scenario.

First wCDA reanalyses have been obtained at the USA National Centre for Environmental Prediction [52] and at the Japanese Agency for Marine-Earth Science and Technology [59], with global coupled models using 3DVar and 4DVar respectively. At the European Centre for Medium Range Weather Forecast, wCDA-like is performed with the incremental 4DVar in an innovative way. While all other terms are (i.e. background error covariance and observation operator) are kept uncoupled, the full coupled model is used in the outer loop of the minimisation, resulting in an implicit coupling that manifest within the assimilation window [34]. The method has been used to produce the reanalyses CERA-20C [35] for the entire 20th century, and CERA-SAT [54] that include satellite data. A comparison between the explicit (i.e. complete sCDA) and the implicit coupling in the incremental 4DVar has shown that for long assimilation window the latter produces accurate analysis, but the explicit coupling is preferable for short assimilation windows. The transition from a reanalysis to real time prediction is currently under study [8].

The EnKF in a wCDA setting has been successfully used to assimilate ocean data and initialise s2d predictions with the Norwegian Earth System Model (NorESM, [14]). Weakly CDA using the EnKF (in particular the Ensemble Adjustment one) has been performed in [68] to constrain independently atmosphere and ocean at the Geophysical Fluid Dynamics Laboratory (GFDL).

The authors of [38] proposed a sCDA approach in which the observed ocean-atmosphere correlation asymmetry is exploited explicitly when performing the coupled analysis. The maximum correlation occurred when the atmosphere leads the ocean by about the decorrelation time of the atmosphere. The method is referred to as Leading Averaged Coupled Covariance (LACC) and the cross atmosphere-ocean covariance are constructed by using the leading (i.e. one decorrelation time ahead) forecasts and observations and the current ocean state. Using the local ensemble transform Kalman filter (LETKF, [28]), the authors of [55]

improved over wCDA using only atmospheric observations in a coupled atmosphere–ocean model. Strongly coupled EnKF was implemented in [61] to recover the Atlantic merid-ional overturning circulation (AMOC) with simulated observations in a low-order coupled atmosphere-ocean model and later with averaged data of the atmosphere from a millennial-scale simulation of a comprehensive coupled atmosphere–ocean climate model [62]. One of the first cases of an operational, EnKF-based, sCDA for a coupled ocean and sea ice model, is the Norwegian TOPAZ system [53].

The Maximum Likelihood Ensemble Kalman filter (MLFE, [69]) has also been success-fully used in a number of sCDA applications [70]. These include land-atmosphere coupling [60], aerosol-atmosphere coupling [71], as well as chemistry-atmosphere coupling [45]. Different 4DVar CDA approaches are discussed in [56] using an idealised single-column atmosphere–ocean model, the estimation of the cross error covariances for use in CDA with 4DVar is studied in [57], while strategies to mitigate the sampling error in CDA have been described in [58].

1.2 Outline

An heuristic explanation of the impact of the temporal scale separation on CDA is provide in Sect. 2. The numerical atmosphere-ocean model MAOOAM is introduced in Sect. 3 together with a detailed analysis of its stability properties in connection to the atmosphere-ocean coupling strength. Definitions and significance of the Lyapunov exponents and vectors used for the stability analysis are recalled in the "Appendix". Numerical experiments using an EnKF are given in Sect. 4, followed by the conclusions in Sect. 5.

2 The Nature of the Problem

This section aims at illustrating key dynamical aspects of DA in coupled systems with time scale separation. We will intentionally set ourselves in a very idealised framework thus that the discussion that follows has only a general qualitative scope. With that in mind, our goal is to highlight: (i) which scale is more important to be observed, and, (ii) why it is desirable to allow observations from one component to impact the other.

Let us consider two coupled, deterministic and autonomous, ordinary differential equa-tions (ODE) as a prototype for a multiscale dynamical system

$$\frac{d\mathbf{x}}{dt} = \epsilon \mathbf{f}(\mathbf{x}, \mathbf{z}),$$
$$\frac{d\mathbf{z}}{dt} = \mathbf{g}(\mathbf{x}, \mathbf{z}),$$

(1)

with $\mathbf{x} \in \mathbb{R}^{m_x}$, $\mathbf{z} \in \mathbb{R}^{m_z}$, $\mathbf{f} : \mathbb{R}^{m_x+m_z} \mapsto \mathbb{R}^{m_x}$, $\mathbf{g} : \mathbb{R}^{m_x+m_z} \mapsto \mathbb{R}^{m_z}$. The processes $\mathbf{f}$ and $\mathbf{g}$ are assumed to have the same time-scale, thus that their temporal scale difference is "artificially" fully accounted for by the constant, $\epsilon \ll 1$, making the variable $\mathbf{x}$ slower than $\mathbf{z}$. The time t is adimensionalized with respect to the typical time scale of the fast variables. We furthermore assume that $\mathbf{f}$ and $\mathbf{g}$ have similar magnitude, are both bounded from above as $\mathcal{O}(1)$, and the characteristic spatial scales of $\mathbf{x}$ and $\mathbf{z}$ are similar. We recall the unrealistic characters of the above hypotheses. In particular the latter one is done here in order to simplify the treatment thus focusing on the effect of the timescale difference exclusively: in realistic coupled atmosphere-ocean models, atmosphere and ocean do have different spatial scales.

Within a time interval $t_k - t_{k-1} = \mathcal{O}(1)$ the slow scale changes such as $\mathcal{O}(\mathbf{x}_k) = \mathbf{x}_{k-1} + \mathcal{O}(\epsilon)$, while the fast scale as $\mathcal{O}(\mathbf{z}_k) = \mathbf{z}_{k-1} + \mathcal{O}(1)$. Let us suppose to have observations of both scales, $\mathbf{y}^x \in \mathbb{R}^{d_x}$ and $\mathbf{y}^z \in \mathbb{R}^{d_z}$, for the slow and fast scale, respectively. Data from each scale are collected with different frequencies, proportional to their respective time scale, so that observations are more frequent for the fast than for the slow scale.

In order for the observations of the slow scale system's component to monitor its variability, the observational interval has to be of order $\Delta t^x = \mathcal{O}(\epsilon^{-1})$. The fast scale observational interval has to be shorter than the slow scale one, $\Delta t^z \leq \Delta t^x$, and we stipulate for convenience that $\Delta t^x = K \Delta t^z$, with $K \in \mathbb{N}$, meaning that every Δt^x both scales are simultaneously observed. Note that when $\Delta t^z = \mathcal{O}(1)$ the solution of the slow system, $\mathbf{x}(t)$, can be considered approximately constant in the interval $t \in [t_k, t_k + \Delta t^z]$.

The model defined by Eq. (1) is used to assimilate recursively data $\mathbf{y}^z$ every Δt^z, and data $\mathbf{y} = (\mathbf{y}^x, \mathbf{y}^z)^T$ whenever t_k is a multiple of Δt^x. The linearised error evolution between two subsequent analyses reads

$$\begin{bmatrix} \Delta \mathbf{x}_k^f \\ \Delta \mathbf{z}_k^f \end{bmatrix} \approx \begin{bmatrix} \epsilon \mathbf{F} \\ \mathbf{G} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}_{k-1}^a \\ \Delta \mathbf{z}_{k-1}^a \end{bmatrix} = \begin{bmatrix} \epsilon \mathbf{F_x} & \epsilon \mathbf{F_z} \\ \mathbf{G_x} & \mathbf{G_z} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}_{k-1}^a \\ \Delta \mathbf{z}_{k-1}^a \end{bmatrix} \tag{2}$$

where $\Delta \mathbf{x}_k$ and $\Delta \mathbf{z}_k$ are the errors in the slow and fast variables, respectively, while the superscripts "f" and "a" stand for forecast and analysis. The terms $\mathbf{F_x}$ and $\mathbf{F_z}$ are components of the tangent linear model of $\mathbf{f}$, in particular its linearisation with respect to $\mathbf{x}$ and $\mathbf{z}$; the same applies to $\mathbf{g}$.

We first consider how the error in one component impacts the other. This is regulated by the cross terms in Eq. (2), namely $\mathbf{F_z} \Delta \mathbf{z}_{k-1}^a$ for the *Fast*-to-*Slow* dependence, and $\mathbf{G_x} \Delta \mathbf{x}_{k-1}^a$ for the *Slow*-to-*Fast*. Their amplitude measures the degree of sensitivity of one component to the other, and depend on the type and strength of the coupling. For instance, atmosphere-ocean coupling, is usually described via two distinct, yet dependent, processes. A mechanical transfer of kinetic energy from the atmospheric wind to the ocean surface, that tends to slow down atmospheric wind and to enhance ocean waves, and a thermal coupling in which heat is transferred from the warmer to the colder model component. These processes would be, at the first order, encoded in the terms $\mathbf{F_z}$ and $\mathbf{G_x}$, and their relative dominance reflected in their amplitudes.

Let us suppose that at the arbitrary analysis time, t_{k-1}, the analysis error on both components, $\Delta \mathbf{x}_{k-1}^a$ and $\Delta \mathbf{z}_{k-1}^a$, is of $\mathcal{O}(1)$. Using the linearised Eq. (2) we can describe the first order error dynamics within the assimilation interval Δt^z. Let us insert the error order in Eq. (2) and take the norm of both sides

$$\begin{aligned} \mathcal{O}(\Delta \mathbf{x}_k^f) &\approx \epsilon[\|\mathbf{F_x} + \mathbf{F_z}\|]\mathcal{O}(1) \leq \epsilon[\|\mathbf{F_x}\| + \|\mathbf{F_z}\|]\mathcal{O}(1), \\ \mathcal{O}(\Delta \mathbf{z}_k^f) &\approx [\|\mathbf{G_x} + \mathbf{G_z}\|]\mathcal{O}(1) \leq [\|\mathbf{G_x}\| + \|\mathbf{G_z}\|]\mathcal{O}(1). \end{aligned} \tag{3}$$

Given that the analysis error is of $\mathcal{O}(1)$, one desires (at the best) the forecast error bound to be also $\mathcal{O}(1)$. By substituting $\mathcal{O}(\Delta \mathbf{x}_k^f) = \mathcal{O}(1)$ and $\mathcal{O}(\Delta \mathbf{z}_k^f) = \mathcal{O}(1)$ in Eq. (3), we get the following bounds for the amplitude of the tangent linear model (i.e. the first order model sensitivity) of the slow component

$$\begin{aligned} \mathcal{O}(\|\mathbf{F_x}\|) &\leq \mathcal{O}(\epsilon^{-1}) \quad Slow \mapsto Slow \text{ sensitivity,} \\ \mathcal{O}(\|\mathbf{F_z}\|) &\leq \mathcal{O}(\epsilon^{-1}) \quad Fast \mapsto Slow \text{ sensitivity,} \end{aligned} \tag{4}$$

and for the fast component

$$
\begin{aligned}
\mathcal{O}(\|\mathbf{G_x}\|) \leq \mathcal{O}(1) \quad Slow \mapsto Fast \text{ sensitivity,} \\
\mathcal{O}(\|\mathbf{G_z}\|) \leq \mathcal{O}(1) \quad Fast \mapsto Fast \text{ sensitivity.}
\end{aligned}
\tag{5}
$$

From Eq. (4), we see that the slow scale sensitivities can be as large as $\mathcal{O}(\epsilon^{-1})$. This means that an $\mathcal{O}(1)$ error in any of the scales will not (in general, $\mathcal{O}(\|\mathbf{F_x}\|), \mathcal{O}(\|\mathbf{F_z}\|) > \mathcal{O}(\epsilon^{-1})$) cause a larger order error in the slow scale forecast. In particular, the second inequality in (4) indicates that the $Fast$-to-$Slow$ scale effect is generally little, and the forecast error will not grow over $\mathcal{O}(1)$, within $\Delta t^z = \mathcal{O}(1)$.

The reduced $Fast$-to-$Slow$ effect is also explained by recalling that within the interval Δt^z the slow scale is almost constant and it largely "feels" the fast one via its smoothed averaged signal, with a time variability of the same order the slow scale. This mechanism is often adduced to explain the somehow little $Fast$-to-$Slow$ effect observed in coupled DA experiments with more realistic atmosphere-ocean coupled models. For instance [62] performed coupled DA with the EnKF in a comprehensive coupled atmosphere–ocean climate model showing that atmospheric observations alone, albeit frequent, do not suffice to properly recover the slowly evolving Atlantic meridional overturning circulation (AMOC), and that, in the absence of ocean data, the use of time-averaged atmospheric measurements was able to successfully track the AMOC (see also [49] for a review of recent coupled DA operational efforts). Note however that, in those cases the fast and slow components do not generally have the same amplitude nor the same spatial scale, as we have hypothesised here.

The sensitivity bounds on the fast scale, Eq. (5), are smaller: an $\mathcal{O}(1)$ internal or $Slow$-to-$Fast$ sensitivity is enough to cause an $\mathcal{O}(1)$ forecast error growth. In particular, and as opposed to the $Fast$-to-$Slow$ case, the first of the inequalities (5), indicates the larger impact of the slow scale on the fast one, again in line with the aforementioned works by [49,62].

Nevertheless, it is the second inequalities in (5) that sets the highest challenge: it implies that the fast scale analysis error must be kept within $\mathcal{O}(1)$ otherwise a "locally" large $\|\mathbf{G_z}\|$, beyond $\mathcal{O}(1)$, will lead the forecast error to grow over $\mathcal{O}(1)$. The only way to achieve this is by directly observing the fast scale and, via coupled DA, allowing the slow scale measurements to update the fast scale. Whenever the fast scale is left unobserved, the error in the scale will generally grow over $\mathcal{O}(1)$ within $\Delta t^z = \mathcal{O}(1)$ and, through the cross component sensitivity $\|\mathbf{F_z}\|$, will inevitably impact the slow scale too.

In conclusion, while the ideal situation is to have data on both scales, those on the fast one are particularly important. They are needed to keep error in the fast scale to a low level, thus preventing the growth of error in the slow scale via the crossing term. Slow scale observations are beneficial and desirable too. They are instrumental to keep error in the slow scale to small levels; they are, however, less capable to contain the growth of the fast scale errors. It is finally worth stressing again the very ideal character of the above conclusions and of Sect. 2 at large. The full picture in a real system can be far more complicated. For instance the relative roles of the atmosphere and ocean in real system is observed to be very different in the Tropics and in mid-latitudes [2].

3 A Coupled Atmosphere-Ocean Model: MAOOAM

3.1 Generalities

In our experiments we shall use the coupled atmosphere-ocean numerical model MAOOAM [15], which is an instructive low-order model for coupled dynamics. MAOOAM is composed of a two-layer quasi-geostrophic (QG) atmosphere that is coupled, both thermally and mechanically, to a QG shallow-water ocean layer in the β-plane approximation. The model solves for the vorticity and the temperature in both media, and is written in spectral Fourier modes, whose full number can be adjusted to the desired resolution.

In our applications we set the total number of Fourier modes alternatively to $m = 36$, 52, or 56. Linear and nonlinear terms in the Fourier expansion are projected onto the phase subspace spanned by the selected modes, using an appropriate scalar product. The model and its properties are described in [15,65].

MAOOAM develops baroclinic instability: the solar forcing induces a horizontal North-South temperature gradient in the atmosphere, which in turns maintains the vertical gradient of the wind. This is possible because the atmosphere possesses two vertical layers. The wind gradient then creates a shear force which is responsible for eddies at the interface of the two layers; they are the cause of instability in the model. Concurrently, the ocean transports the heat to counteract the initial gradient of temperature.

The model is numerically integrated with a time-step of approximately 16 min.

3.2 Selected Model Configurations

We consider three model setups with dimension $m = 36$, $m = 52$ and $m = 56$. In the case, $m = 36$ (the most widely used in previous studies with MAOOAM) the modes, i.e. the model's state vector components, are distributed between atmosphere and ocean as follows: the first 10 are associated to the atmospheric barotropic streamfunction, followed by 10 modes for the atmospheric temperature, 8 for the ocean streamfunction, and 8 for the ocean temperature. In the configuration $m = 52$, 16 modes (8 for both streamfunction and temperature respectively) are added to the ocean. Finally, for $m = 56$, 4 additional atmospheric modes (2 for both barotropic streamfunction and temperature) are added to the atmosphere.

For each of these three model's dimensions, we consider two atmosphere-ocean couplings, hereafter referred to as *weak* and *strong*, making a total of six model configurations: $36wk$, $52wk$, $56wk$, $36st$, $52st$ and $56st$. The coupling strength is varied by acting on the friction coefficients and the heat exchange between the two media, as described in Table 1; other key model parameters are included in Table 2.

An illustration of the long term dynamical behaviour of configurations $36wk$ and $36st$ is given in Fig. 1(panels (a) and (b), respectively). Both panels show the trajectory solution of the model for 10^7 days, projected onto the 3-dimensional portion of the phase space spanned by three key modes $(\psi_2^{\text{ocn}}, \theta_2^{\text{ocn}}, \psi_1^{\text{atm}})$, i.e. the second Fourier modes of the ocean streamfunction and temperature, and the first one of the atmospheric streamfunction; the importance of these three modes as representative of the model dynamics in the full phase space has been put forward in [65].

The marked difference between the attractors' shapes (cf. the two panels of Fig. 1) is a manifestation of the different coupling strength. In the weakly coupled configuration, $36wk$ (panel (a)), the attractor has a sort of regular large scale shape (a spheroid) that is densely, albeit discontinuously due to chaos, filled by the trajectory as typical of an ergodic

Table 1 Summary of the six MAOOAM configurations under consideration, with indication of the atmosphere-ocean coupling strength

	Coupling	k (adim)	kp (adim)	λ (W m^2 K)	d (s^{-1})	T_0^{atm} (K)	T_0^{ocn} (K)	m^{atm}	m^{ocn}
36wk	Weak	0.010	0.020	10	6×10^{-8}	289	301	20	16
52wk								20	32
56wk								24	32
36st	Strong	0.0145	0.029	15.06	9×10^{-8}	290.20	299.35	20	16
52st								20	32
56st								14	32

The table reports the values of the key parameters, modulating the coupling's strength: the friction coefficients at the bottom of the atmosphere, k, between internal atmospheric layers, kp, and between atmosphere and ocean, d, as well as the heat exchange between the two media, λ, and the stationary solutions for the $0-$th order atmospheric and ocean temperature, T_0^{atm} and T_0^{ocn} (see [64] for a complete description of the model parameters and their role). The model dimension, for the atmosphere and ocean, m^{atm} and m^{ocn}, respectively, is reported in the last two columns for each of the configurations. The total dimension is $m = m^{\text{atm}} + m^{\text{ocn}}$,

Table 2 List of the remaining MAOOAM parameters having the same values in both coupling configurations

C^{atm} (W/m^2)	310	Radiation input for the atmosphere
C^{ocn} (W/m^2)	310	Net short wave radiation input for the ocean
γ^{atm} (J/m^2 K)	10^7	Specific heat capacity of the atmosphere
γ^{ocn} (J/m^2 K)	6.6×10^8	Specific heat capacity of the ocean
H (m)	165	Depth of the ocean layer
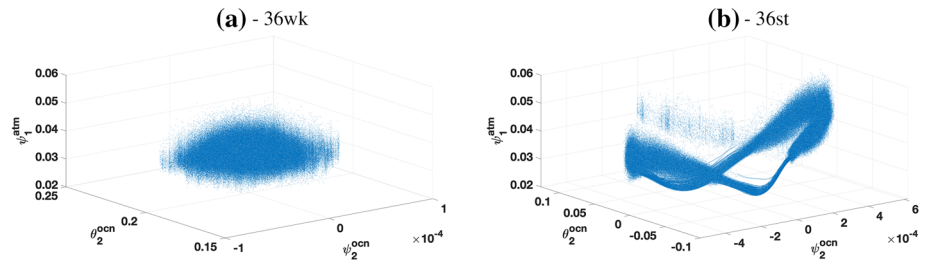
Fig. 1 Illustration of the MAOOAM attractor in the configurations $36wk$ (**a**) and $36st$ (**b**) for the three variables $(\psi_2^{\mathrm{ocn}}, \theta_2^{\mathrm{ocn}}, \psi_1^{\mathrm{atm}})$. The model is integrated forward for 10^7 days

system. The attractor for the configuration $36st$ is still the one of a chaotic dynamics, yet it is organised now around an unstable periodic orbit around which the solution is wandering. This dynamics is accompanied by a succession of recurrences in regions of lower or higher values of the atmospheric streamfunction, ψ_1^{atm}, with low and high variability respectively. We will hereafter refer to them as the "passive" and "active" regimes respectively. This different behaviour appears clear when looking at the time series of ψ_1^{atm} in Fig. 2; the red and green spots in the figure indicate the start of one active and one passive regime, respectively. Note furthermore that ψ_1^{atm} displays a low-frequency variability with a period of about 70 years. This low-frequency variability is characterised by a slow motion along the attractor of the system leading for instance to the succession of peaks and minima in the streamfunction field of Fig. 2.

In order to estimate how the different variables in the model correlate to each other, and globally, how the atmosphere and ocean components are correlated, we compute the model auto-correlation every 10 days and then averaged over 10^5 days, for the configuration $36wk$. Results are shown in Fig. 3 for three cases. Besides the instantaneous values (panel (a)), we also compute the correlation between the ocean and the time-averaged atmosphere with averaging windows of 100 days (panel (b)) and 1000 days (panel (c)).

Not surprisingly, when looking at the instantaneous values of the correlations (a), the self-components (i.e. the atmosphere-atmosphere and ocean-ocean) values are so much greater than the atmosphere-ocean correlation, that the latter values are almost invisible (yet they are not zero). It is interesting to note the well organised band-shape structure of the atmospheric correlation with a second maximum showing the correlation between atmospheric barotropic streamfunction and temperature, as opposed to the unstructured, yet very rich, pattern of the ocean auto-correlation. These are the correlations that would make possible in sCDA to update ocean/atmospheric variables by observing other ocean/atmospheric variables.

A noteworthy feature of Fig. 3 is the substantial increase of the atmosphere-ocean cross correlation when the ocean is correlated with a time-averaged atmosphere (panels (b) and
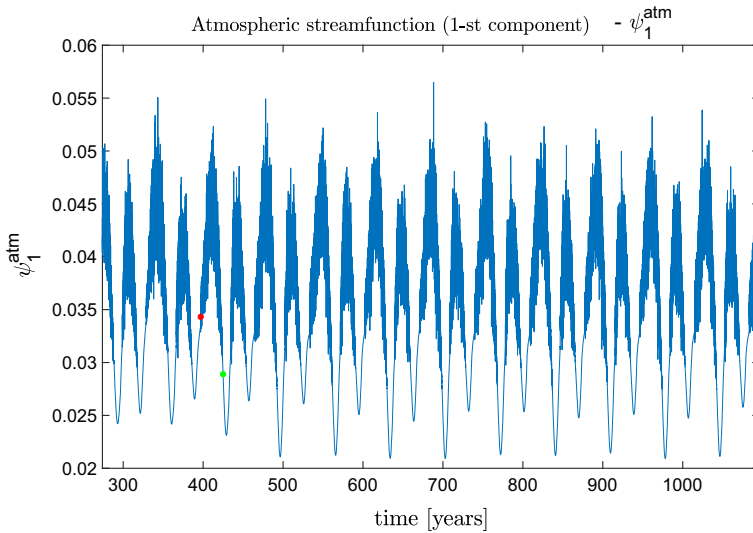
Fig. 2 Time series of the first component of the atmospheric streamfunction, ψ_1^{atm}, for the model configuration 36st. The red and green points indicate the start of active and passive regimes respectively
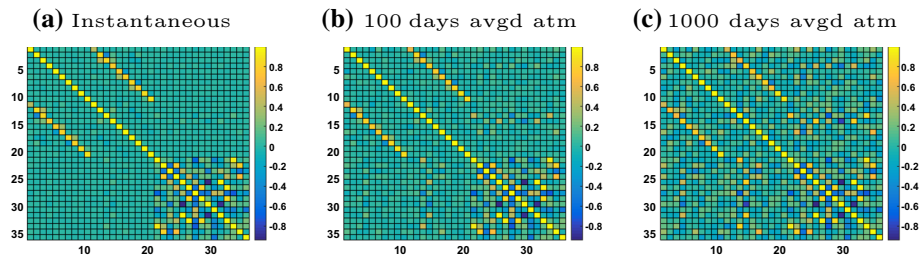

Fig. 3 Correlation matrices in MAOOAM in configuration $36wk$; the axes display the system's state index. The correlation is computed every 10 days, using the instantaneous values (**a**), and the atmospheric time-averaged values over the averaging period of 100 days (**b**) and 1000 days (**c**). Results are averaged over 10^5 days

(c)). This cross correlation increases when the averaging window for the atmosphere is increased from 100 to 1000 days, and decreases further over 1000 days (not shown). This behaviour naturally emerges as a consequence of the time-scale difference between ocean and atmosphere. It has been already put forward in previous studies (see e.g. [61] and references therein), and is what has promoted the use of averaged observations in several early studies on coupled DA [17,29,38].

3.3 Stability Analysis

We characterise the long-term dynamical behaviour of the six model configurations by studying their stability properties. We compute their spectrum of Lyapunov exponents (LEs; see "Appendix") and, based on them, the Kolmogorov entropy (KE; given by the sum of the positive LEs) and the Kaplan–Yorke attractor dimension (KY-dim) (see e.g., [41]). Results are reported in Table 3, while the spectrum's of the LEs for the six model configurations are

Table 3 Summary of the stability analysis results for the six MAOOAM configurations

Model configuration	36wk	52wk	56wk	36st	52st	56st
# Positive $\lambda_i \in [10^{-2},\ 1]$	3	3	4	2	2	1
# Near-neutral$^+$ $\lambda_i \in [10^{-5},\ 10^{-2}]$	3	7	3	2	4	4
# Neutral $\lambda_i \in [-10^{-5},\ 10^{-5}]$	1	1	2	1	1	1
# Near-neutral$^-$ $\lambda_i \in [-10^{-2},\ -10^{-5}]$	13	25	12	11	25	13
# Negative $\lambda_i \in [-1,\ -10^{-2}]$	16	16	34	20	20	37
Kolmogorov entropy	0.498	0.528	0.459	0.139	0.060	0.029
Kaplan–Yorke dimension	25.06	41.03	28.42	20.29	33.35	19.32

The numbers of LEs are counted as distributed in five ranges of values given in the first column. The last two rows report the values of KE and KY-dim respectively

shown in Fig. 4. From Table 3 we see that MAOOAM possesses a large number of almost neutral LEs. To better distinguish real neutral LEs (within numerical accuracy) from very little, albeit non-zero, ones, we split them in five categories: we will consider real neutral exponents those in the interval $\lambda_i \in [-10^{-5},\ 10^{-5}]$. The neighbouring ranges of "near-neutral$^+$" and "near-neutral$^-$" (see Table 3), encompass those exponents that, although not strictly zero, act almost as such.

As anticipated in Sect. 3.1, in the weakly coupled configurations the ocean slowing effect on the atmosphere is less effective, resulting in the model being more chaotic than in the corresponding strongly coupled configurations. By degree of chaos we mean the amplitude and number of the positive LEs. While both factors are merged in the definition of the KE, that in itself already represents a good measure of chaos, KE does not distinguish among the rate of growth along individual Lyapunov modes. All of the strong coupling configurations are characterised by smaller KEs, compared to the weakly ones.

The addition of 20 ocean modes from configuration $36wk/st$ to $52wk/st$ has almost no effect on the positive, nor on the negative, portions of the spectrum, but only on the neutral ones; as a result, the KE is only slightly different between $36wk/st$ and $52wk/st$. However, the KE is slightly larger in the $52wk$ case compared to $36wk$, and slightly smaller in the $52st$ compared to $36st$. This is because in the former case the amplitude of the positive LEs does not change much, while their number is larger for the configuration $52wk$. On the other hand, the larger number of positive LEs in the $52st$ over the $36st$ is counteracted by a reduction in their amplitudes.

In both coupling strength cases, the transition from dimension 36 to 52 leads to almost doubling the number of the almost neutral LEs. These are a manifestation of, and are arisen by, the additional 20 ocean modes. The role of the ocean modes as responsible for the neutral portion of the spectrum was already observed by [66], where a broader analysis of the connection between physical variables and LEs in MAOOAM was presented. Note also that, although the increase of the almost neutral LEs does not change much the overall degree of instabilities (and therefore the intrinsic predictability of the configurations $36wk$ and $52wk$), it changes substantially the KY-dim, that is much larger in the $52wk$ case. In deterministic dynamics, the number of non-negative LEs, n_0, and the KY-dim are known to be directly proportional to the number of ensemble members that an ensemble Kalman filter (EnKF) needs to achieve satisfactory performance [13], with n_0 being the minimum ensemble size required to avoid filter divergence [6]. These findings have recently been explored for coupled dynamics by [50].
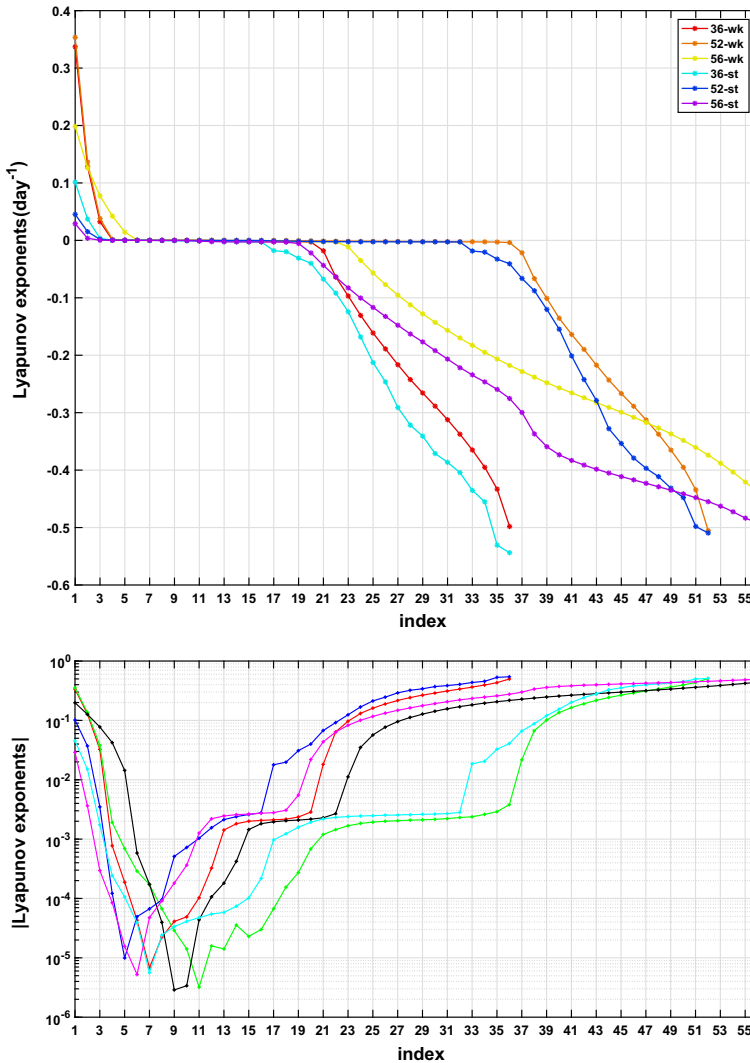
Fig. 4 Spectrum of Lyapunov exponents for the six MAOOAM configurations (top) and absolute values of the Lyapunov exponents (bottom) with log-scale in the *y*-axis

The further dimensional increase from $52wk/st$ to $56wk/st$ causes a surprising, and difficult to interpret, change in the LEs spectrum. In both coupling cases, the number of positive (including small positive) LEs decreases, while that of negative LEs is doubled. Thus, the addition of the 4 atmospheric modes is not increasing the degree of chaos as we might have expected based on the idea that atmosphere brings chaos, while ocean takes it away. The KE and the KY-dim are both smaller in the $56wk/st$ compared to $52wk/st$, implying that, despite the systems' state dimensions, i.e. the full phase-space, is larger, fewer ensemble members may be needed in the $56wk/st$ than in the $52wk/st$ configurations.

We conclude the section by studying the covariant Lyapunov vectors (CLVs; see "Appendix") for the two smallest configurations, $36wk$ and $36st$. Similarly to the analysis
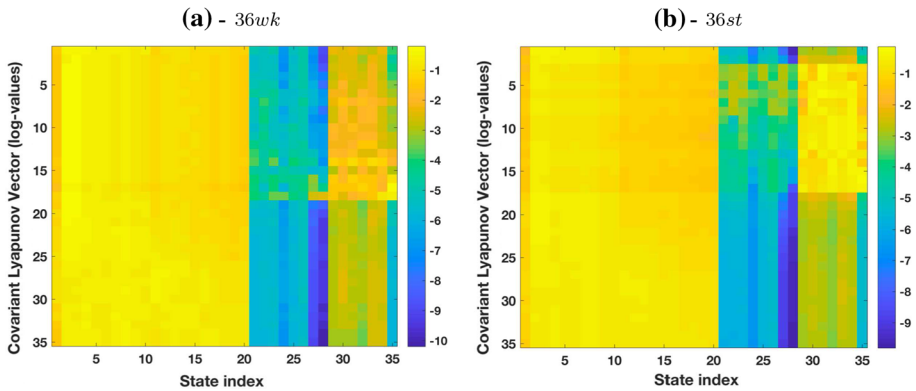
Fig. 5 Time average of the logarithm of the projections of the CLVs (y-axis) onto the model state vector components (index; x-axis) for configurations $36wk$ (**a**) and $36st$ (**b**)

reported in [66], Fig. 5 shows the CLVs amplitude projections (in log-scale) on the individual model's state vector components.

Overall, the CLVs project largely on the atmospheric components (i.e. state index 1 to 20), but the oceanic temperature (state index 30 to 36) also presents significant projections. Some key CLVs associated with exponents close to 0 also display large averaged projections on the oceanic streamflow (state index 21 to 28). This demonstrates the coupling character of the instabilities that span across both atmosphere and ocean. In fact, when the coupling is increased (configuration $36st$, panel (b)) the relative amplitudes of the projections on the ocean components increase commensurately.

4 Coupled Ensemble Kalman Filter with MAOOAM

We present some illustrative numerical experiments using the ensemble Kalman filter (EnKF) with MAOOAM. The specific version of the EnKF adopted here is the *finite-size* EnKF (EnKF-N; [5,7]). The EnKF-N is a deterministic EnKF with high accuracy in low-dimensional systems, and that incorporates the estimation of the inflation meant to counteract sampling errors, that would otherwise have had to be tuned. We do not report here the description of the EnKF-N; readers can find all details in [5,7]. Note also that, to simplify the notation, we will hereafter systematically use the acronym EnKF to refer to the EnKF-N. This choice is also done to stress that the results that follow would be qualitatively the same for any deterministic formulation of the EnKF.

The EnKF is used to perform sCDA and the results that follow refer to this case only. We have also performed experiments using wCDA and the results, not shown, indicate overall lower skills than sCDA.

Experiments are performed with varying ensemble size, N, as well as atmosphere and ocean observational intervals, Δt^{atm} and Δt^{ocn}. Simulated observations are sampled from a trajectory, solution of MAOOAM, that is taken to represent the truth with respect to which we compute the root-mean-square-error (RMSE), as in standard twin experiments. Observational error is simulated by adding Gaussian random noise, and the model-to-data relation reads:

$$\mathbf{y}_k = \mathbf{y}_k^{\text{atm}} = \mathcal{H}^{\text{atm}}(\mathbf{x}_k, \mathbf{z}_k) + \boldsymbol{\epsilon}_k^{\text{atm}} \quad \text{when} \quad \text{mod}\,(t_k, \Delta t^{\text{ocn}}) \neq 0, \tag{6}$$

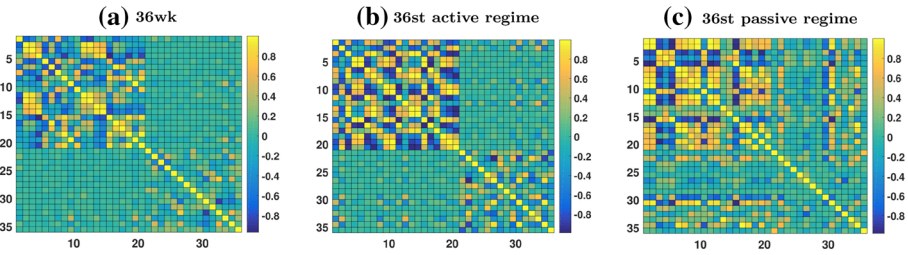**(a)** 36wk   **(b)** 36st active regime   **(c)** 36st passive regime

Fig. 6 EnKF ensemble-based correlation matrices, with $N = 15$ members, at time $t = 1$ year of simulation. The axes display the system's state index

and

$$\mathbf{y}_k = \begin{bmatrix} \mathbf{y}_k^{\text{ocn}} \\ \mathbf{y}_k^{\text{atm}} \end{bmatrix} = \begin{bmatrix} \mathcal{H}^{\text{ocn}}(\mathbf{x}_k, \mathbf{z}_k) + \boldsymbol{\epsilon}_k^{\text{ocn}} \\ \mathcal{H}^{\text{atm}}(\mathbf{x}_k, \mathbf{z}_k) + \boldsymbol{\epsilon}_k^{\text{atm}} \end{bmatrix} \quad \text{when} \quad \text{mod}\,(t_k, \Delta t^{\text{ocn}}) = 0, \qquad (7)$$

with $\mathcal{H}^{\text{ocn}} : \mathbb{R}^{m_{\text{ocn}}+m_{\text{atm}}} \mapsto \mathbb{R}^{d_{\text{ocn}}}$ and $\mathcal{H}^{\text{atm}} : \mathbb{R}^{m_{\text{ocn}}+m_{\text{atm}}} \mapsto \mathbb{R}^{d_{\text{atm}}}$ being the observational operators, $d_{\text{ocn}} \leq m_{\text{ocn}}$ and $d_{\text{atm}} \leq m_{\text{atm}}$, and $\Delta t^{\text{ocn}} = K \Delta t^{\text{atm}}$, $K \in \mathbb{N}$. The observational errors in the two components, $\boldsymbol{\epsilon}_k^{\text{ocn}}$ and $\boldsymbol{\epsilon}_k^{\text{atm}}$, are assumed to be mutually independent and both unbiased and normally distributed with constant error covariances $\mathbf{R}^{\text{ocn}} \in \mathbb{R}^{d_{\text{ocn}} \times d_{\text{ocn}}}$ and $\mathbf{R}^{\text{atm}} \in \mathbb{R}^{d_{\text{atm}} \times d_{\text{atm}}}$, respectively. In the experiments the observational error is assumed to be spatially uncorrelated so that the matrices $\mathbf{R}^{\text{ocn}}$ and $\mathbf{R}^{\text{atm}}$ are diagonal, and the error standard deviation (the square-root of the diagonal entries of the matrices) to be equal to 1% of the MAOOAM component-wise' natural variability, i.e. the long-term time averaged difference between uncorrelated states. MAOOAM variables are directly observed, implying that the observation operators $\mathcal{H}^{\text{ocn}}$ and $\mathcal{H}^{\text{atm}}$ are linear and expressed as matrices of appropriate dimension with only 1 and 0 as entries. The assumed ability to observe directly the model modes is an idealisation. In realistic scenarios one would have, at the best, point-wise measurements of physical quantities that are, usually nonlinearly, related to the model modes. This would introduce "representation error" (see e.g. [31]) and degrade the performance of data assimilation as shown by [48] using MAOOAM.

Figure 6 displays the correlation matrices of the EnKF after 1 year of assimilation. The ensemble size is $N = 15$ members, observations of the full system are assimilated every $\Delta t^{\text{ocn}} = \Delta t^{\text{atm}} = 24$ h. MAOOAM configurations 36wk and 36st are used, and in the latter case the assimilation experiments are initialised alternatively in the active (panel (b)) and passive (panel (c)) regimes (see also Fig. 2). The figure clearly reveals the impact of the coupling on the correlation between atmosphere (top-left 20×20 portion) and the ocean (bottom-right 16×16 portion): when the coupling is weak (panel (a)) the off-diagonal entries are very small, and emerge when the coupling is increased in configuration 36st. As expected, the atmosphere-ocean correlation is much larger in the passive regime; it is in fact the effect of the ocean that dominates in this regime and is reflected in cross-correlation patterns. It is remarkable that an ensemble of as few as 15 members in the EnKF is able to provide physically-sound correlation patterns.

The relation between ensemble size and skill (in terms of the RMSE of the EnKF analysis) is studied in Fig. 7. The figure shows the global RMSE (over the whole model's domain) time averaged over 300 years as a function of the ensemble size N. The RMSE is normalised with the standard deviation of the observational error, so that it has to be lower than 1 for the EnKF to be performing satisfactorily. Observation type and frequency are the same as for Fig. 6: the full system is observed every 24 h.

Fig. 7 EnKF analysis RMSE averaged over 300 years for the six model configurations. The system is fully observed every 24 h. The number of non-negative LEs of each of the model configurations is indicated by the vertical dashed lines
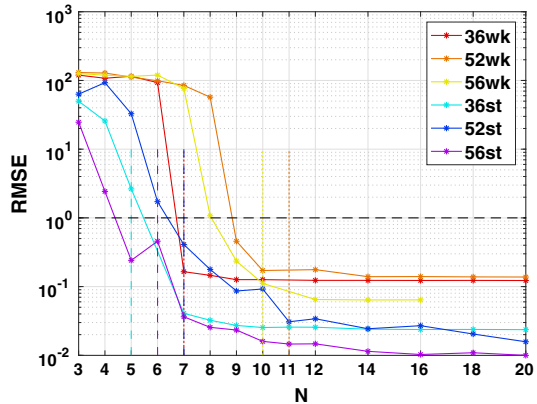
The figure shows clearly how the RMSE of the EnKF analysis decreases below the observational error level, as soon as the number of members exceeds the number of non-negative LEs. This number is indicated by the vertical dashed lines for all of the model configurations (cf. Table 3). Together with [27], this result confirms, and extends to coupled dynamics, what is described in [6] for system having a single scale of motion. This behaviour is due to the fact that, when the system is sufficiently well observed, the error dynamics behaves quasi-linearly and the errors are confined within the unstable subspace of the system. As soon as the ensemble subspace is able to fully align to the unstable subspace, the EnKF effectively reduces the error. Importantly, Fig. 7 implies that, even in coupled systems, when the aforementioned condition on the observations holds, a deterministic EnKF will only need to have a number of ensemble members larger than n_0 to achieve good performance. Nevertheless, as opposed to the behaviour of uncoupled systems, we observe here a gradual error reduction in some cases, e.g. $52st$ and $56wk$, even for $N > n_0$. This behaviour was already observed by [48] and conjectured to be due to the extended spectrum of near-zero LEs in the coupled system. In fact, these quasi-neutral asymptotic LEs have high probability to be instantaneously positive. It is therefore preferable (if not mandatory) to have them accounted for in the EnKF update, so as to counteract the upwell of unfiltered error from asymptotically weakly stable (but often locally unstable) directions, as explained in [24].

Figure 8 shows the time series of the RMSE of the analysis together with the ensemble spread for the configurations $36wk$ and $36st$, for the four variables; $N = 15$ and $\Delta t^{\mathrm{ocn}} = \Delta t^{\mathrm{atm}} = 24$ h. Note that the experiment in the configuration $36st$ lasts for twice the duration of the $36wk$; this choice is done to balance for the slower time scale of the system when the coupling is stronger. The observational error level is also displayed for reference.

As anticipated from Fig. 7, in all cases the RMSE is well below the observational error. The ensemble spread is also consistent with the RMSE, proving the sound functioning of the EnKF. The different temporal scales between atmosphere and ocean, as well as between weak and strong coupling configurations are evident. The figure also highlights the switch between the active and passive regimes in the $36st$ configuration. It is remarkable how well the EnKF is able to adjust to them and properly estimate the state.

The effect of changing the observational intervals is studied in Fig. 9 which shows the RMSE of the EnKF analysis as a function of Δt^{ocn} and Δt^{atm}, for the configuration $36wk$. Results (not shown) for the other configurations and coupling are qualitatively equivalent. Experiments last 1 year, the ensemble size is $N = 15$ members and errors are averaged both
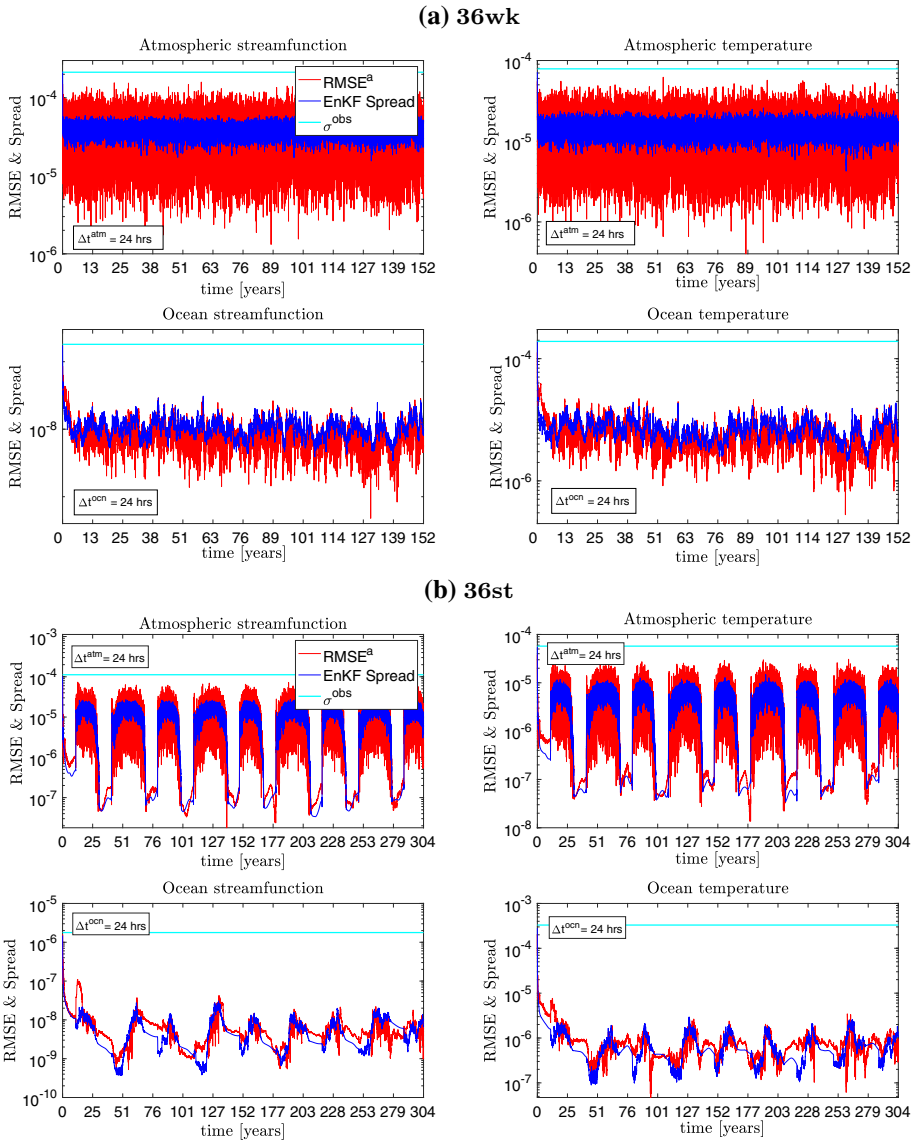
Fig. 8 Time series of the RMSE of the EnKF analysis and the ensemble spread for the configurations $36wk$ and $36st$, for the four class of model variables, atmosphere and ocean temperature and streamfunction

in time (over this 1 year) and on a sample of 10 initial conditions. In Fig. 9a, atmosphere and ocean are both and simultaneously observed, but in Fig. 9b and c, only the atmosphere or the ocean is observed, respectively.

When the atmosphere and the ocean are both observed (panels (a)), the RMSE shows a monotonic growth trend when Δt^{ocn} and Δt^{atm} are increased, although the RMSE of the ocean streamfunction, and to a lesser extent the ocean temperature, seem quite insensitive in the interval $1h \leq \Delta t \leq 3d$. Note that the RMSEs of all four variables stay below the
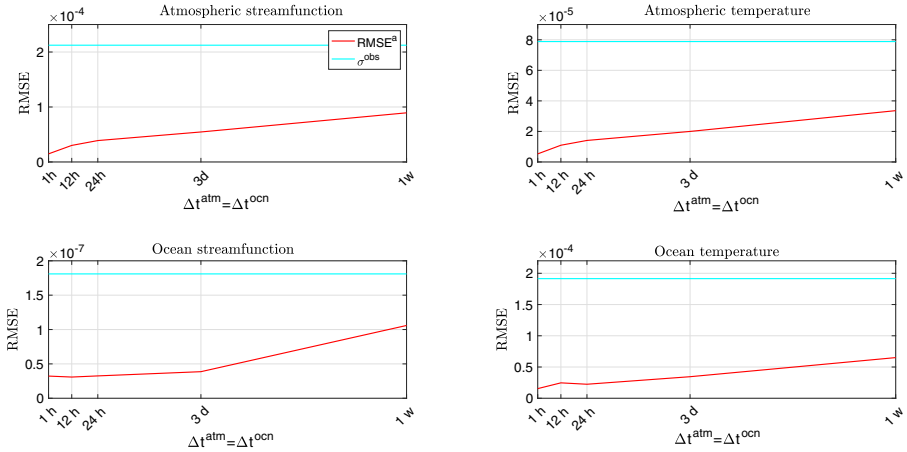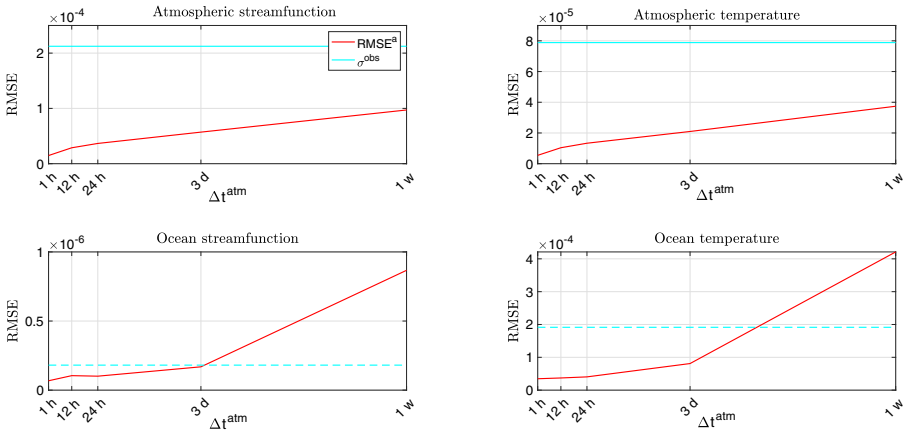
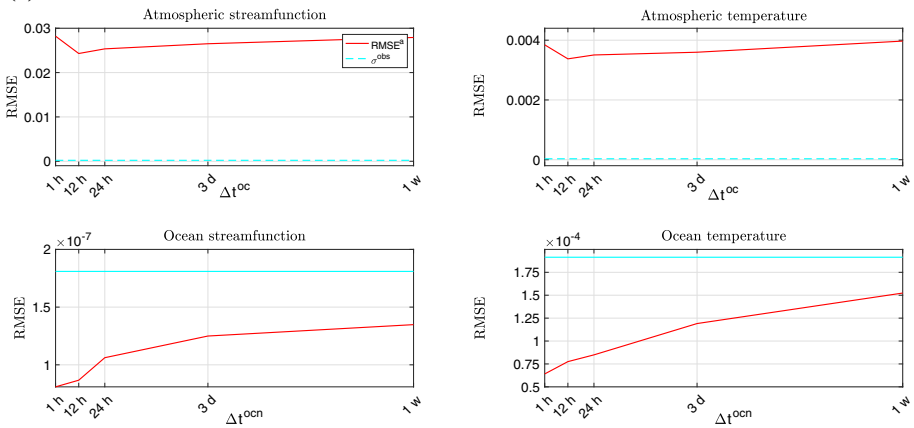**Fig. 9** RMSE of the EnKF analysis as a function of **a** the atmospheric and ocean data interval, $\Delta t^{\mathrm{atm}} = \Delta t^{\mathrm{ocn}}$; **b** atmospheric data interval, Δt^{atm}, **c** ocean data interval, Δt^{ocn}. In case **b** only the atmospheric data are assimilated whereas in case **c** only the ocean data are assimilated. Model configuration $36wk$ and $N = 15$

observational error level for all the considered observational intervals. The situation changes slightly when only atmospheric data are used (panels (b)). Remarkably the error level in the atmosphere appears insensitive to the removal of the ocean observations. While this is obviously not the case for the ocean RMSE, which in fact increases when only atmospheric data are available, it is interesting to observe that it only increases very little. In practice, when $1h \leq \Delta t^{\mathrm{atm}} \leq 3d$, the ocean RMSE stays below the observational level even in the absence of ocean data; all information is brought and propagated from the atmosphere. The importance of atmospheric data is further highlighted by the results in panels (c) in which atmospheric observations are removed and the EnKF only assimilates ocean data. We see how the atmospheric RMSE is now above the observational level consistently in all variables. Ocean RMSE, while slightly lower than when only atmospheric data were available (panels (b)), it is not as low as when both ocean and atmosphere were observed (panels (a)). The importance of observing the fast scale, as conjectured in Sect. 2, is thus corroborated by these numerical findings. It is also relevant to observe that, despite its slow time scale, the ocean analysis improves even when observations are assimilated as frequently as $1h$ and $12h$, in line with what found in [48].

Figure 10 is similar to Fig. 9 except that now the observation interval is kept fixed in one component, either atmosphere or ocean, while varied in the other. In the experiments of panels (a) the ocean data are assimilated every week, $\Delta t^{\mathrm{ocn}} = 1\mathrm{w}$, while the frequency of atmospheric data is changed. Conversely, panels (b) show experiments where the atmospheric data are assimilated every half-day, $\Delta t^{\mathrm{atm}} = 12\,\mathrm{h}$, while the frequency of ocean data is changed. Similarly to Fig. 9, results of Fig. 10 also confirms the importance of providing a sufficient control of the fast scale (the atmosphere) error growth, by keeping the observation frequency high enough. In fact, the comparison of panels (a) and (b) reveal how in the latter case, when ocean data frequency is changed but the atmosphere is observed every $\Delta t^{\mathrm{atm}} = 12\,\mathrm{h}$ the analysis RMSE is maintained consistently below the observational error. This contrasts with the behaviour shown in panels (a) where, although the ocean is observed every $\Delta t^{\mathrm{ocn}} = 1\mathrm{w}$, the analysis RMSE in all variables grows with the increase of the atmospheric data frequency, eventually reaching a level higher than the observational error.

5 Conclusion

The term "coupled data assimilation" (CDA) has been increasingly used in recent years to refer to the application of data assimilation (DA; e.g. [11]) in dynamical systems possessing many, and separated, scales of motion that are coupled together in their dynamical equations [49]. Systems of this sort are common in many areas of sciences, but CDA has emerged distinctively in climate science where Earth system numerical models couple together models of the atmosphere, land, ocean and cryosphere. Classical DA is prone not to work efficiently because the scale separation acts as a barrier hindering the transmission of the information content across model components (e.g. ocean and atmosphere). Understanding origins and causes limiting classical DA is important and may help guiding adaptations and novel solutions.

We have provided an introduction to CDA in Sect. 1, together with a survey of the current status of the research in the field. By using dynamical arguments, in Sect. 2 we traced back the core issue and illustrated in which way, to a first order, information flows from the fast to the slow scale or vice-versa. Furthermore we conjectured how observations of both scales have to be temporally distributed in order to best reduce the state estimation error. We deduced that:
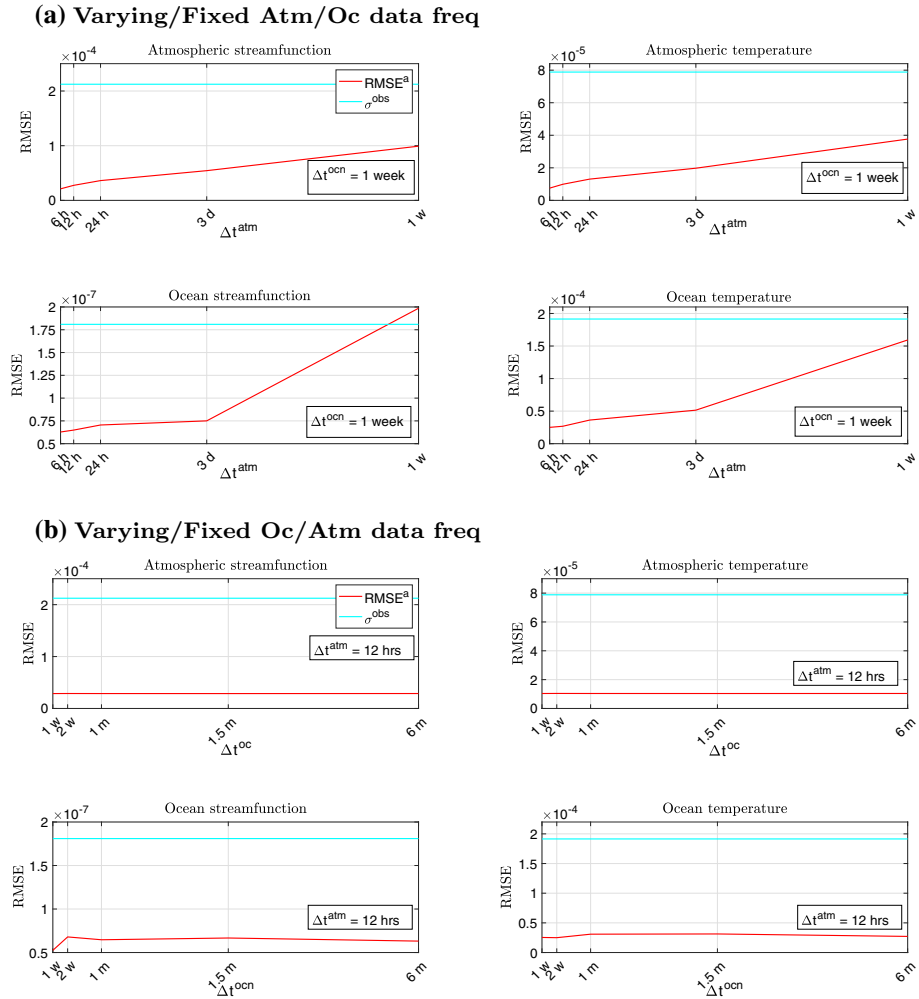
(a) Varying/Fixed Atm/Oc data freq

(b) Varying/Fixed Oc/Atm data freq

Fig. 10 RMSE of the EnKF analysis as a function of **a** the atmospheric data interval with fixed ocean data interval, $\Delta t^{ocn} = 1w$; **b** the ocean data interval with fixed atmospheric data interval, $\Delta t^{atm} = 12h$. The model configuration is $36wk$, with an ensemble size set to $N = 15$

(i) cross components effects are generally stronger in the direction from the slow to the fast scale, so that observations of the slow scale may benefit to the fast, but, (ii) intra-component effects are much stronger in the fast scale. The fast scale must be controlled by frequently enough observations to prevent the error to grow up and affect the slow scale.

The above is in overall agreement with previous works that, while having shown benefit in both directions, have also indicated atmospheric data to be more effective in constraining the ocean than the opposite (see e.g. [48] and references therein). This includes studies (see e.g. [56] and references therein) where uncoupled but forced models of the atmosphere or the ocean are considered. In cases when the ocean is forced with pre-computed atmospheric surface fluxes, error in the latter are responsible for biases in the ocean, revealing an atmosphere-to-ocean impact; see also [48] for an extensive discussion on the transition from uncoupled-forced models to weakly- and strongly-coupled DA.

Our conjectures have been confirmed in numerical experiments performed with the modular arbitrary-order coupled atmosphere-ocean model, MAOOAM [15], in which a state-of-the-art ensemble Kalman filter, the EnKF-N [7] has been implemented. MAOOAM has been used in six different configurations, having different sizes and atmosphere-ocean coupling strengths. The attractors in the weak and strong coupling cases appear very different, with the strong one showing two distinct regimes and a low frequency variability with a period of about 16 years. We have characterised the model stability properties via the spectrum of Lyapunov exponents, the Kolmogorov entropy, the Kaplan–Yorke attractor dimension and the covariant Lyapunov vectors (CLVs). In particular, the averaged projections of the CLVs onto the state vector reveal how the different model instabilities are driven by the atmosphere and/or the ocean.

The experiments with the EnKF-N have confirmed the behaviour anticipated in Sect. 2: atmosphere has to be observed frequently enough (i.e. about every 12 h) in order to achieve a global analysis (including the ocean) with low error (i.e. below the observational error). Moreover, experiments largely prove that, likewise uncoupled dynamics [6], deterministic EnKFs (to which category the EnKF-N belongs) require an ensemble at least as large as the number of non-negative Lyapunov exponents (assuming localisation is not used), to get satisfactory results in coupled systems too. However, as opposed to uncoupled systems, whenever the model displays a degeneracy-like in the Lyapunov spectrum (with many near-zero exponents) the analysis error still gradually decreases for $N > n_0$; in uncoupled systems the analysis error reduction almost fully ceases when $N > n_0$. This behaviour, originally observed in [48], is arguably due to the presence of multiple near-neutral asymptotic directions, with a high chance to be locally unstable.

Although asymptotically neutral or weakly-stable, these directions may display high variance in the local error growth rate, thus be often intermittently unstable. As rigorously proved by [23,24] this situation is known to drive the error upwell from the unfiltered to the filtered subspace, eventually leading to divergence. It seems thus paramount that the EnKF ensemble subspace encompasses at least all of these near-neutral directions, preferably also the asymptotically weakly stable.. Furthermore, given that these directions are generated by the coupling itself (see Fig. 6 and [66]) their impact on the performance of the EnKF in coupled systems is expected to be ubiquitous. Along these lines, using an idealised multi-scale system made up of coupled copies of the Lorenz 3-variables model, the authors of [50] demonstrate that weakly stable directions are needed in situations of strong nonlinear dynamics and intermittent error growth. Similarly, the authors of [39] have suggested that the variability in the number of unstable modes associated to unstable periodic orbits in a simple Earth system model, can explain the observed very different predictability of individual atmospheric blocking events, and have argued that DA must thus cautiously incorporate stable modes too.

One of our current research endeavours is the study of suitable reduced-rank formulations of the EnKF that take into account the unstable modes in coupled systems, in analogy to the assimilation in the unstable subspace (AUS; [42]) so far applied to uncoupled systems. Similarly, the map between the instability rank and the state vector drawn by the CLVs may help designing monitoring strategies in which observing devices are deployed in the areas of large CLVs (see e.g. [12] for a similar strategy based on breeding vectors of the DA cycle). Part of the questions related to extending AUS to coupled dynamics have been undertaken in the recent work [50], although many still remain to be addressed using more realistic models with the aforementioned Lyapunov degeneracy. This can be done using MAOOAM given that its number of positive and neutral exponents can be very large [16], as they manifest the coupling mechanisms. Do we still need such a large amount of ensemble members, or is there a limit beyond which a further increase is not necessary anymore and, if so, under

which circumstances? These questions are worth addressing to properly set up the EnKF in multi-scale dynamics.

Appendix: Lyapunov Exponents and Covariant Lyapunov Vectors

The initial state of a system is never known exactly since the process of measurement and data assimilation is always subjected to finite precision. To clarify the implications of the presence of such an error we consider an initial state displaced slightly from $\mathbf{x}(t_0) = \mathbf{x}_0$ by an initial error $\delta\mathbf{x}_0$. This perturbed initial state generates a new trajectory in phase space and we define the instantaneous error vector as the vector joining the points of the reference trajectory and the perturbed one at a given time, $\delta\mathbf{x}_t$. Provided that this perturbation is sufficiently small, its dynamics can be described by the linearised equation,

$$\frac{\mathrm{d}\delta\mathbf{x}}{\mathrm{d}t} = \frac{\partial\mathbf{f}}{\partial\mathbf{x}_{|\mathbf{x}_t}}\delta\mathbf{x} \tag{8}$$

and the formal solution can be written as,

$$\delta\mathbf{x}_t = \mathbf{M}_{t:t_0}(\mathbf{x}_0)\delta\mathbf{x}_0 \tag{9}$$

where the matrix $\mathbf{M}$, referred as the resolvent matrix, plays an important role in error growth dynamics as revealed when writing the Euclidean norm of the error,

$$\begin{aligned}
E_t &= \|\delta\mathbf{x}_t\|^2 = \delta\mathbf{x}_t^{\mathrm{T}}\delta\mathbf{x}_t \\
&= \delta\mathbf{x}_0^{\mathrm{T}}\mathbf{M}_{t:t_0}^{\mathrm{T}}(\mathbf{x}_0)\mathbf{M}_{t:t_0}(\mathbf{x}_0)\delta\mathbf{x}_0
\end{aligned} \tag{10}$$

The growth of E_t is conditioned by the eigenvalues of the matrix $\mathbf{M}^{\mathrm{T}}\mathbf{M}$, where $(.)^{\mathrm{T}}$ indicates transposition (and complex conjugation if necessary). In ergodic theory of chaotic systems, the double limit of infinitely small initial errors and infinitely long times, is usually considered (e.g. [19]). In these limits the divergence of initially closed states is determined by the logarithm of the eigenvalues of the matrix $[\mathbf{M}^{\mathrm{T}}\mathbf{M}]^{2(t-t_0)}$ that are referred to as the Lyapunov exponents (LEs). The full set of LEs of a system is called the Lyapunov spectrum which are usually represented in decreasing order.

Associated with each of these exponents a natural direction of (in)stability can be defined which is a local property on the attractor of the system, these are known as the covariant

Lyapunov vectors (CLVs). These CLVs were first introduced in [51] and later discussed in [36,63]. They form a norm-independent and covariant basis of the tangent linear space, providing a splitting between the unstable manifold. This splitting describes the unstable perturbations leading to the divergence of the trajectories, the neutral manifold, typically associated with the direction of the flow, and the stable manifold, which corresponds to the contracting directions. Several algorithms for computing these vectors are available [21,22, 33,67].

The CLVs are defined through a suitable geometric construction involving both the orthogonal forward and backward Lyapunov vectors, whose computation can be seen as a byproduct of the usual Benettin et al. algorithm [4] used for estimating of LEs, see also the nice reviews on these topics in [19,21,33]. The size of the perturbations oriented according to the CLVs grows or decays with an approximate exponential law, where the average of the fluctuating rates of growth or decay correspond one-to-one to the LEs. As opposed to the CLVs, the forward and backward Lyapunov vectors (except for the first) are not covariant, so that it is hard to interpret them physically [46]. Therefore, CLVs allow for associating a time-dependent field to each LE, thus providing a connection between observed rates of growth and decay of perturbations and the corresponding physical modes of the system.

A recent analysis of the statistical and dynamical properties of these vectors in the context of MAOOAM has been performed in [66]. A remarkable result is the splitting of the tangent space in two categories of CLVs, a first set mostly associated with the dynamics within the atmosphere, the most unstable and stable directions, and a second set of vectors for which the corresponding Lyapunov exponents are close to 0. The latter set forms a quasi-neutral subspace of truly coupled ocean-atmosphere modes.

References

1. Asch, M., Bocquet, M., Nodet, M.: Data Assimilation: Methods, Algorithms, and Applications. Fundamentals of Algorithms. SIAM, Philadelphia (2016)
2. Bach, E., Motesharrei, S., Kalnay, E., Ruiz-Barradas, A.: Local atmosphere-ocean predictability: dynamical origins, lead times, and seasonality. J. Clim. **32**(21), 7507–7519 (2019)
3. Ballabrera-Poy, J., Kalnay, E., Yang, S.C.: Data assimilation in a system with two scales–combining two initialization techniques. Tellus A **61**(4), 539–549 (2009)
4. Benettin, G., Galgani, L., Giorgilli, A., Strelcyn, J.M.: Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. Part 1: theory. Meccanica **15**(1), 9–20 (1980)
5. Bocquet, M.: Ensemble Kalman filtering without the intrinsic need for inflation. Nonlinear Process. Geophys. **18**, 735–750 (2011). https://doi.org/10.5194/npg-18-735-2011
6. Bocquet, M., Carrassi, A.: Four-dimensional ensemble variational data assimilation and the unstable subspace. Tellus A **69**(1), 1304504 (2017)
7. Bocquet, M., Raanes, P.N., Hannart, A.: Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation. Nonlinear Process. Geophys. **22**, 645–662 (2015). https://doi.org/10.5194/npg-22-645-2015
8. Browne, P.A., de Rosnay, P., Zuo, H., Bennett, A., Dawson, A.: Weakly coupled ocean-atmosphere data assimilation in the ECMWF NWP system. Remote Sens. **11**(3), 234 (2019)
9. Brunet, G., Jones, S., Ruti, P.M., et al.: Seamless Prediction of the Earth System: From Minutes to Months. World Meteorological Organization, Geneva (2015)
10. Brunet, G., Shapiro, M., Hoskins, B., Moncrieff, M., Dole, R., Kiladis, G.N., Kirtman, B., Lorenc, A., Mills, B., Morss, R., et al.: Collaboration of the weather and climate communities to advance subseasonal-to-seasonal prediction. Bull. Am. Meteorol. Soc. **91**(10), 1397–1406 (2010)
11. Carrassi, A., Bocquet, M., Bertino, L., Evensen, G.: Data assimilation in the geosciences: an overview of methods, issues, and perspectives. Wiley Interdiscip. Rev. Clim. Change **9**(5), e535 (2018)
12. Carrassi, A., Trevisan, A., Uboldi, F.: Adaptive observations and assimilation in the unstable subspace by breeding on the data-assimilation system. Tellus A **59**, 101–113 (2007)

13. Carrassi, A., Vannitsem, S., Zupanski, D., Zupanski, M.: The maximum likelihood ensemble filter performances in chaotic systems. Tellus A **61**, 587–600 (2009)
14. Counillon, F., Bethke, I., Keenlyside, N., Bentsen, M., Bertino, L., Zheng, F.: Seasonal-to-decadal predictions with the ensemble Kalman filter and the Norwegian earth system model: a twin experiment. Tellus A **66**(1), 21074 (2014)
15. De Cruz, L., Demaeyer, J., Vannitsem, S.: The modular arbitrary-order ocean-atmosphere model: MAOOAM v1.0. Geosci. Model. Dev. **9**(8), 2793–2808 (2016). https://doi.org/10.5194/gmd-9-2793-2016
16. De Cruz, L., Schubert, S., Demaeyer, J., Lucarini, V., Vannitsem, S.: Exploring the lyapunov instability properties of high-dimensional atmospheric and climate models. Nonlinear Process. Geophys. **25**(2), 387–412 (2018). https://doi.org/10.5194/npg-25-387-2018
17. Dirren, S., Hakim, G.J.: Toward the assimilation of time-averaged observations. Geophys. Res. Lett. **32**(4), L04804 (2005)
18. Doblas-Reyes, F.J., García-Serrano, J., Lienert, F., Biescas, A.P., Rodrigues, L.R.: Seasonal climate predictability and forecasting: status and prospects. Wiley Interdiscip. Rev. Clim. Change **4**(4), 245–268 (2013)
19. Eckmann, J.P., Ruelle, D.: Ergodic theory of chaos and strange attractors. In: The Theory of Chaotic Attractors, pp. 273–312. Springer (1985)
20. Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, 2nd edn. Springer, Berlin (2009)
21. Froyland, G., Hüls, T., Morriss, G.P., Watson, T.M.: Computing covariant lyapunov vectors, oseledets vectors, and dichotomy projectors: a comparative numerical study. Physica D **247**(1), 18–39 (2013)
22. Ginelli, F., Poggi, P., Turchi, A., Chaté, H., Livi, R., Politi, A.: Characterizing dynamics with covariant lyapunov vectors. Phys. Rev. Lett. **99**(13), 130601 (2007)
23. Grudzien, C., Carrassi, A., Bocquet, M.: Asymptotic forecast uncertainty and the unstable subspace in the presence of additive model error. SIAM/ASA J. Uncertainty Quantif. **6**(4), 1335–1363 (2018)
24. Grudzien, C., Carrassi, A., Bocquet, M.: Chaotic dynamics and the role of covariance inflation for reduced rank kalman filters with model error. Nonlinear Process. Geophys. Disc. **2018**, 1–25 (2018). https://doi.org/10.5194/npg-2018-4
25. Hannart, A., Carrassi, A., Bocquet, M., Ghil, M., Naveau, P., Pulido, M., Ruiz, J., Tandeo, P.: DADA: data assimilation for the detection and attribution of weather and climate-related events. Clim. Change **136**(2), 155–174 (2016)
26. Harlim, J., Majda, A.J.: Filtering turbulent sparsely observed geophysical flows. Mon. Weather Rev. **138**(4), 1050–1083 (2010)
27. Haussaire, J.M., Bocquet, M.: A low-order coupled chemistry meteorology model for testing online and offline data assimilation schemes: L95-GRS (v1.0). Geosci. Model Dev. **9**, 393–412 (2016). https://doi.org/10.5194/gmd-9-393-2016
28. Hunt, B., Kostelich, E.J., Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. Physica D **230**, 112–126 (2007)
29. Huntley, H.S., Hakim, G.J.: Assimilation of time-averaged observations in a quasi-geostrophic atmospheric jet model. Clim. Dyn. **35**(6), 995–1009 (2010)
30. Hutt, A., Stannat, W., Potthast, R.: Data Assimilation and Control: Theory and Applications in Life Sciences. Frontiers Media SA, Lausanne (2019)
31. Janjić, T., Bormann, N., Bocquet, M., Carton, J., Cohn, S., Dance, S., Losa, S., Nichols, N., Potthast, R., Waller, J., et al.: On the representation error in data assimilation. Q. J. R. Meteorol. Soc. **144**(713), 1257–1278 (2018)
32. Kadakia, N., Armstrong, E., Breen, D., Morone, U., Daou, A., Margoliash, D., Abarbanel, H.D.: Nonlinear statistical data assimilation for hvc_RA neurons in the avian song system. Biol. Cybern. **110**(6), 417–434 (2016)
33. Kuptsov, P.V., Parlitz, U.: Theory and computation of covariant Lyapunov vectors. J. Nonlinear Sci. **22**, 727–762 (2012)
34. Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K., Janssen, P.: A coupled data assimilation system for climate reanalysis. Q. J. R. Meteorol. Soc. **142**(694), 65–78 (2016)
35. Laloyaux, P., de Boisseson, E., Balmaseda, M., Bidlot, J.R., Broennimann, S., Buizza, R., Dalhgren, P., Dee, D., Haimberger, L., Hersbach, H., et al.: Cera-20c: A coupled reanalysis of the twentieth century. J. Adv. Model. Earth Syst. **10**(5), 1172–1195 (2018)
36. Legras, B., Vautard, R.: A guide to Liapunov vectors. In: Proceedings 1995 ECMWF Seminar on Predictability, vol. 1, pp. 143–156. Citeseer (1996)
37. Lorenc, A.C., Payne, T.: 4D-Var and the butterfly effect: statistical four-dimensional data assimilation for a wide range of scales. Q. J. R. Meteorol. Soc. **133**(624), 607–614 (2007)
38. Lu, F., Liu, Z., Zhang, S., Liu, Y.: Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part I: simple model study. Mon. Weather Rev. **143**(9), 3823–3837 (2015)

39. Lucarini, V., Gritsun, A.: A new mathematical framework for atmospheric blocking events. Clim. Dyn. **54**(1–2), 575–598 (2020)

40. Moye, M.J., Diekman, C.O.: Data assimilation methods for neuronal state and parameter estimation. J. Math. Neurosci. **8**(1), 11 (2018)

41. Ott, E.: Chaos in Dynamical Systems. Cambridge University Press, Cambridge (2002)

42. Palatella, L., Carrassi, A., Trevisan, A.: Lyapunov vectors and assimilation in the unstable subspace: theory and applications. J. Phys. A **46**, 254020 (2013)

43. Palatella, L., Trevisan, A., Rambaldi, S.: Nonlinear stability of traffic models and the use of lyapunov vectors for estimating the traffic state. Phys. Rev. E **88**(2), 022901 (2013)

44. Palmer, T., Doblas-Reyes, F., Weisheimer, A., Rodwell, M.: Toward seamless prediction: calibration of climate change projections using seasonal forecasts. Bull. Am. Meteorol. Soc. **89**(4), 459–470 (2008)

45. Park, S.K., Lim, S., Zupanski, M.: Structure of forecast error covariance in coupled atmosphere-chemistry data assimilation. Geosci. Model Dev. **8**(5), 1315–1320 (2015)

46. Pazó, D., Rodríguez, M.A., López, J.M.: Spatio-temporal evolution of perturbations in ensembles initialized by bred, lyapunov and singular vectors. Tellus A **62**(1), 10–23 (2010)

47. Penny, S.G., Akella, S., Buehner, M., Chevallier, M., Counillon, F., Draper, C., Frolov, S., Fujii, Y., Karspeck, A., Kumar, A., Laloyaux, P., Mahfouf, J.F., Matthew, M., Peña, M., de Rosnay, P., Subramanian, A., Tardiff, R., Wang, Y., Wu, X.: Coupled data assimilation for integrated earth system analysis and prediction: goals, challenges, and recommendations. In: Techincal Report WWRP 2017-3, vol. 3, pp. 1–59. World Meteorological Organization, Geneva (2017)

48. Penny, S.G., Bach, E., Bhargava, K., Chang, C.C., Da, C., Sun, L., Yoshida, T.: Strongly coupled data assimilation in multiscale media: experiments using a quasi-geostrophic coupled model. J. Adv. Model. Earth Syst. **11**(6), 1803–1829 (2019). https://doi.org/10.1029/2019MS001652

49. Penny, S.G., Hamill, T.M.: Coupled data assimilation for integrated earth system analysis and prediction. Bull. Am. Meteorol. Soc. **97**(7), ES169–ES172 (2017)

50. Quinn, C., O'Kane, T.J., Kitsios, V.: Application of a local attractor dimension to reduced space strongly coupled data assimilation for chaotic multiscale systems. Nonlinear Process. Geophys. **27**(1), 51–74 (2020). https://doi.org/10.5194/npg-27-51-2020

51. Ruelle, D.: Ergodic theory of differentiable dynamical systems. Publications Mathématiques de l'Institut des Hautes Études Scientifiques **50**(1), 27–58 (1979)

52. Saha, S., Moorthi, S., Pan, H.L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., et al.: The NCEP climate forecast system reanalysis. Bull. Am. Meteorol. Soc. **91**(8), 1015–1057 (2010)

53. Sakov, P., Counillon, F., Bertino, L., Lisæter, K., Oke, P., Korablev, A.: TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic. Ocean Sci. **8**(4), 633 (2012)

54. Schepers, D., de Boisséson, E., Eresmaa, R., Lupu, C., de Rosnay, P.: Cera-sat: A coupled satellite-era reanalysis. ECMWF Newslett. **155**, 32–37 (2018)

55. Sluka, T.C., Penny, S.G., Kalnay, E., Miyoshi, T.: Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. Geophys. Res. Lett. **43**(2), 752–759 (2016)

56. Smith, P.J., Fowler, A.M., Lawless, A.S.: Exploring strategies for coupled 4D-Var data assimilation using an idealised atmosphere-ocean model. Tellus A **67**(1), 27025 (2015)

57. Smith, P.J., Lawless, A.S., Nichols, N.K.: Estimating forecast error covariances for strongly coupled atmosphere-ocean 4D-var data assimilation. Mon. Weather Rev. **145**(10), 4011–4035 (2017)

58. Smith, P.J., Lawless, A.S., Nichols, N.K.: Treating sample covariances for use in strongly coupled atmosphere-ocean data assimilation. Geophys. Res. Lett. **45**(1), 445–454 (2018)

59. Sugiura, N., Awaji, T., Masuda, S., Mochizuki, T., Toyoda, T., Miyama, T., Igarashi, H., Ishikawa, Y.: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. J. Geophys. Res. Oceans **113**, C10 (2008)

60. Suzuki, K., Zupanski, M., Zupanski, D.: A case study involving single observation experiments performed over snowy siberia using a coupled atmosphere-land modelling system. Atmos. Sci. Lett. **18**(3), 106–111 (2017)

61. Tardif, R., Hakim, G.J., Snyder, C.: Coupled atmosphere-ocean data assimilation experiments with a low-order climate model. Clim. Dyn. **43**(5–6), 1631–1643 (2014)

62. Tardif, R., Hakim, G.J., Snyder, C.: Coupled atmosphere-ocean data assimilation experiments with a low-order model and CMIP5 model data. Clim. Dyn. **45**(5–6), 1415–1427 (2015)

63. Trevisan, A., Pancotti, F.: Periodic orbits, lyapunov vectors, and singular vectors in the lorenz system. J. Atmos. Sci. **55**(3), 390–398 (1998)

64. Vannitsem, S.: The role of the ocean mixed layer on the development of the north atlantic oscillation: a dynamical system's perspective. Geophys. Res. Lett. **42**(20), 8615–8623 (2015)

65. Vannitsem, S., Demaeyer, J., Cruz, L.D., Ghil, M.: Low-frequency variability and heat transport in a low-order nonlinear coupled ocean-atmosphere model. Physica D **309**, 71–85 (2015). https://doi.org/10.1016/j.physd.2015.07.006
66. Vannitsem, S., Lucarini, V.: Statistical and dynamical properties of covariant lyapunov vectors in a coupled atmosphere-ocean model-multiscale effects, geometric degeneracy, and error dynamics. J. Phys. A **49**(22), 224001 (2016)
67. Wolfe, C.L., Samelson, R.M.: An efficient method for recovering lyapunov vectors from singular vectors. Tellus A **59**(3), 355–366 (2007)
68. Zhang, S., Harrison, M., Rosati, A., Wittenberg, A.: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. Mon. Weather Rev. **135**(10), 3541–3564 (2007)
69. Zupanski, M.: Maximum likelihood ensemble filter: theoretical aspects. Mon. Weather Rev. **133**, 1710–1726 (2005)
70. Zupanski, M.: Data assimilation for coupled modeling systems. In: Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, vol. 3, pp. 55–70. Springer (2017)
71. Zupanski, M., Kliewer, A., Wu, T.C., Apodaca, K., Bian, Q., Atwood, S., Wang, Y., Wang, J., Miller, S.D.: Impact of atmospheric and aerosol optical depth observations on aerosol initial conditions in a strongly-coupled data assimilation system. Atmos. Chem. Phys. Discuss. **2019**, 1–25 (2019). https://doi.org/10.5194/acp-2019-2

Affiliations

Maxime Tondeur[1] · Alberto Carrassi[2,3,4] · Stephane Vannitsem[5] · Marc Bocquet[6]

✉ Alberto Carrassi
    n.a.carrassi@reading.ac.uk

[1] Paris, France

[2] Department of Meteorology, University of Reading and National Centre for Earth Observations, Reading, UK

[3] Mathematical Institute, Utrecht University, Utrecht, Holland

[4] Nansen Environmental and Remote Sensing Center, Bergen, Norway

[5] Royal Meteorological Institute of Belgium, Brussels, Belgium

[6] CEREA, Joint Laboratory École des Ponts ParisTech and EDF R&D, Université Paris-Est, Champs-sur-Marne, France