Adaptive covariance hybridization for the assimilation of SST observations within a coupled Earth system reanalysis

Sébastien Barthélémy[1], François Counillon[1,2], Yiguo Wang[2]

[1]Geophysical Institute, University of Bergen and Bjerknes Centre for Climate research, Bergen, Norway
[2]Nansen Environmental and Remote Sensing Centre and Bjerknes Centre for Climate Research, Bergen, Norway

Key Points:

- Hybrid covariance handles sampling error and improves the update of deep water masses when assimilating surface observation with an EnKF
- The method is well suited to provide a long coupled reanalysis of the past century
- Hybrid covariance with adaptive hybrid coefficients explicitly estimated in space and time achieved the best performance

Corresponding author: Sébastien Barthélémy, `sebastien.barthelemy@uib.no`

Abstract

Ensemble data assimilation methods, such as the Ensemble Kalman Filter, are well suited for climate reanalysis and predictions because they feature flow-dependent covariance, and the ensemble can be used to estimate the uncertainty of the system. However, because Earth System Models are heavy computationally, the method typically uses a few tens of members. Sampling error in the covariance matrix can introduce biases in the unobserved regions (e.g. in the deep ocean), which may cause a drift in the reanalysis and in the predictions. Here, we assess the potential of the hybrid covariance approach (EnKF-OI) to counteract sampling error and rank deficiencies. The EnKF-OI linearly combines the superior flow-dependent covariance computed from a dynamical ensemble with another covariance matrix that is static but less prone to sampling error. We test the method within the Norwegian Climate Prediction Model (NorCPM), which combines the Norwegian Earth System Model (NorESM) and the Ensemble Kalman Filter (EnKF). We test the performance of the reanalyses in an idealised twin experiment, where we assimilate synthetic sea surface temperature observations monthly for the period 1980-2010. The dynamical ensemble consists of 30 members, and the static ensemble is composed of 315 seasonal members sampled from a long stable pre-industrial run. We compare the performance of the EnKF to 1) an EnKF-OI with a global hybrid coefficient tuned empirically, referred to as standard hybrid and 2) an EnKF-OI with adaptive hybrid coefficients explicitly estimated in space and time. Both hybrid covariance methods cure the bias introduced by the EnKF at intermediate and deep water. The adaptive EnKF-OI performs best overall, and that by doing smaller updates than the standard hybrid version. In the adaptive EnKF-OI, the hybrid coefficient remains nearly constant throughout the reanalysis, with only a weak seasonal variability.

Plain Language Summary

Data assimilation is a statistical method that reduces uncertainty in a model, based on observations. A popular method for climate reanalysis and prediction are ensemble method that relies on statistics from a finite ensemble of model realisations. However, observations are sparse – mostly near the surface – and the sampling error from data assimilation method introduces a degradation in the deep ocean. We use a method that complements this ensemble with an pre-existing database of model states to reduce sampling error. We show that the approach substantially reduces error at the intermediate and deep ocean. The method typically requires the tunning of a parameter, but we show that it can be estimated online, achieving the best performance.

1 Introduction

Data assimilation estimates the state of a model (called the analysis) that approaches the "unknown true state" of the system based on observations, a prior model estimate, and statistical information on their uncertainties. Data assimilation is applied sequentially/recursively to provide a reanalysis, which can also be used to understand the mechanisms of variability and initialise predictions. Data assimilation has been one of the key ingredients in the progress of numerical weather prediction (Bauer et al., 2015) and is now used in a wide range of geosciences applications (Carrassi et al., 2018), including climate prediction. The ensemble Kalman Filter (EnKF, Evensen, 2003), is an advanced data assimilation method that provides flow-dependent covariance – *i.e.,* that can evolve in time and space with a transient state or a regime shift – and the ensemble provides a quantification of the uncertainty of the system. These properties are well suited for climate reanalysis and predictions, and the method is becoming increasingly popular in that field (*e.g.*, Zhang et al., 2009; Counillon et al., 2014; Brune et al., 2015; Karspeck et al., 2018; O'Kane et al., 2019).

Ensemble-based data assimilation methods (e.g., the EnKF) approximate the forecast error covariance matrix using a finite-size ensemble from a Monte Carlo integration step. Rank deficiencies and sampling errors emerge and deteriorate the performance of the system, causing an artificial reduction of the ensemble spread and may even lead to filter divergence. Several ad-hoc methods have been introduced to counteract sampling errors. Localization (Houtekamer & Mitchell, 2001; Evensen, 2003; Ott et al., 2004; Anderson, 2007) limits the spatial extent of the corrections, based on the approximation that the covariance function decays as a function of the distance – and can be seen as a way to effectively reduce the degree of freedom of the system. Inflation (Anderson, 2001; Whitaker & Hamill, 2012; Raanes et al., 2019) counteracts the spread-collapse by artificially inflating the ensemble spread at every assimilation step. The last method is the covariance hybridization method, which is the topic of our paper.

Covariance hybridization (Hamill & Snyder, 2000) combines linearly the flow-dependent covariance computed from a finite Monte-Carlo ensemble with another covariance matrix that is less prone to sampling error. The static matrix can be parameterized (Hamill & Snyder, 2000; Weaver & Courtier, 2001), computed from a long model simulation (Counillon et al., 2009), computed as the average of the background error covariance matrices from a previous data assimilation run (Carrió et al., 2021) or computed from a dynamical ensemble at a lower resolution (Rainwater & Hunt, 2013). The hybrid covariance method achieves better performance than the standalone EnKF, particularly for small ensembles, and performance converges to that of the EnKF for large ensembles (X. Wang et al., 2007; Counillon et al., 2009; Raboudi et al., 2019). The computational cost of the hybrid covariance methods is customizable to the desired cost.

The linear coefficients combining the static and the dynamic covariance are called the "hybridization coefficients", which optimally balances the superior but noisy sample covariance with that of less noisy but static covariance. To achieve optimal performance, it is crucial to tune these coefficients (X. Wang et al., 2007; Counillon et al., 2009; Raboudi et al., 2019; Gharamti et al., 2014). The optimal values of these coefficients depend on the non-stationarity of the dynamical system as well as the data assimilation settings, such as the dynamical ensemble size, localisation and inflation settings. As such, it is expected that the optimal value of the hybridization coefficients should vary in space and time. A first attempt to estimate spatial and time-varying hybridization coefficient has been developed (Gharamti, 2020) with a Bayesian framework (using fixed localisation settings). They found that a spatially heterogeneous hybridization coefficient formulation outperforms a homogeneous formulation. Ménétrier and Auligné (2015) and Ménétrier et al. (2015) formulated the problem of hybridization as a linear filtering problem of the background error covariance matrix to optimize both the localization and the hybridization coefficients simultaneously.

In this work, we aim to investigate the benefit of background error covariance hybridization for climate reanalysis and climate prediction systems, as for example with the CMIP6 Decadal Climate Prediction Project (DCPP, Boer et al., 2016). We use the Norwegian Climate Prediction Model (NorCPM, Counillon et al., 2014, 2016) that provides coupled reanalysis and contributed to CMIP6 DCPP (Bethke et al., 2021). Here, we focus on long coupled reanalysis as NorCPM will produce such a reanalysis from 1850 to the present. Sea surface temperature (SST) dominates the ocean observation network before the emergence of altimetry in the 90s and Argo data in the 2000s. Using only SST, NorCPM can control the upper ocean heat content, and major indices of climate variability in the North Atlantic well (Counillon et al., 2016). Two features of NorCPM, flow-dependent assimilation and assimilation in isopycnal coordinates, were found to be particularly important in that success. However, it also yields an unrealistic update of the intermediate water masses in the North Atlantic Subpolar Gyre (SPG) (typically below 1000 m) (Counillon et al., 2016), which subsequently causes a drift in the multi-year pre-

dictions in the North Atlantic SPG region (Bethke et al., 2018). We aim to address this limitation in our current assimilation system.

We test the performance of NorCPM for coupled reanalysis in the framework of idealised twin experiments with the assimilation of SST. We assess whether 1) hybrid covariance can solve this issue and compare the performance of a 31-year coupled reanalysis produced with the EnKF (currently used in NorCPM) and hybrid covariance methods, 2) compare robustness and optimality of two flavours of hybrid covariance methods: one using an empirically tuned globally constant hybridization coefficient and one where the hybridization coefficients are estimated adaptively in space and time (Ménétrier et al., 2015; Ménétrier & Auligné, 2015; Ménétrier, 2021).

The results show that the hybridization method – both with fixed and adaptive hybridization coefficients – performs equally or better than the EnKF in most places and depths, and substantially improves results in intermediate and deep water masses. The adaptive hybrid performs better than its counterpart with fixed hybridization factors.

This paper is organised as follows. Section 2 presents the Earth System Model (ESM) used in this work, the Norwegian Earth System Model (NorESM). Section 3 presents the deterministic EnKF, and its practical implementation within the NorCPM. Section 4 describes hybridization with a static covariance matrix. Section 5 describes the adaptive covariance hybridization method and its practical implementation within NorCPM for SST assimilation. In section 6, we introduce the experimental design and the evaluation metrics. The numerical results are presented in section 7. Section 8, provides a conclusion, discussion and future perspective to this work.

2 Model system: the Norwegian Earth System Model

NorESM1-ME (Bentsen et al., 2013) is based on version 1.0.4 of the Community Earth System Model (Hurrell et al., 2013). Its atmosphere component is the CAM4-OSLO, the ocean component is the Bergen Layered Ocean Model, BLOM, (Bentsen et al., 2013), the land component is the Community Land Model, CLM4, (Lawrence et al., 2011), the sea ice component is the Los Alamos Sea Ice Model, CICE4, (Bitz et al., 2012), and the coupler is CPL7 (Craig et al., 2012).

The atmosphere and the land components have $1.9^\circ \times 2.5^\circ$ latitude-longitude resolution. The atmosphere component has 26 hybrid sigma-pressure levels ranging from the surface up to 3 hPa. The ocean and the sea ice components have a 1° horizontal resolution in both latitude and longitude with a bipolar grid. BLOM comprises in the vertical a stack of 51 isopycnic layers and two layers for representing the bulk mixed layer. Before 2005, the forcings are the CMIP5 historical forcings (Taylor et al., 2012), while after 2005, they are the representative Concentration Pathway 8.5 forcings (van Vuuren et al., 2011).

3 The deterministic ensemble Kalman filter

Let $\mathbf{X} \in \mathbb{R}^{n \times N}$ an ensemble of N model states $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$, $\mathbf{x} \in \mathbb{R}^n$ the ensemble mean and $\mathbf{A} \in \mathbb{R}^{n \times N}$ the ensemble anomalies. We note n the model state dimension. $\mathbf{x}$ and $\mathbf{A}$ are given by Eq. (1) and (2):

$$\mathbf{x} = \frac{1}{N}\mathbf{X}\mathbf{1}, \tag{1}$$

$$\mathbf{A} = \mathbf{X}\left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right), \tag{2}$$

where $\mathbf{I} \in \mathbb{R}^{N \times N}$ is the identity matrix and $\mathbf{1} \in \mathbb{R}^{N}$ is a vector with all elements equal to 1. In the following equations, the superscripts f and a stand respectively for the forecasted and analysed states of the mean and the anomalies.

We note $\mathbf{x}^{\mathrm{t}}$ the true state of the system. The observations $\mathbf{y}$ are defined by Eq. (3):

$$\mathbf{y} = \mathbf{H}\mathbf{x}^{\mathrm{t}} + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \tag{3}$$

where $\mathbf{H} \in \mathbb{H}^{p \times n}$ is the observation operator, $\mathbf{R} \in \mathbb{R}^{p \times p}$ is the observation error covariance matrix, and p is the number of observations.

In this study, the deterministic EnKF (DEnKF) introduced by Sakov and Oke (2008) is used. The DEnKF is a square-root (deterministic) flavour of the EnKF that solves the analysis without the need for perturbation of the observations. It inflates the error by construction and is intended to perform well in applications where corrections are small (Sakov & Oke, 2008). The scheme has been robustly tested and validated (Sakov et al., 2012; Counillon et al., 2016; Bethke et al., 2021). The DEnKF decomposes into two steps: a forecast step and an analysis step. In the forecast step, each analyzed member $\mathbf{x}_i^{\mathrm{a}}$ is integrated forward in time, which becomes the prior $\mathbf{x}_i^{\mathrm{f}}$ at the following assimilation cycle:

$$\mathbf{x}_i^{\mathrm{f}} = \mathcal{M}(\mathbf{x}_i^{\mathrm{a}}), \qquad i = 1, \ldots, N, \tag{4}$$

where $\mathcal{M}$ is an operator that stands for the model integration.

The analysis step of the DEnKF proceeds in two steps, the update of the ensemble mean, Eq. (5), and the update of the ensemble anomalies, Eq. (6):

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{f}} + \mathbf{K}\left(\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathrm{f}}\right), \tag{5}$$

$$\mathbf{A}^{\mathrm{a}} = \mathbf{A}^{\mathrm{f}} - \frac{1}{2}\mathbf{K}\mathbf{H}\mathbf{A}^{\mathrm{f}}, \tag{6}$$

where:

$$\mathbf{K} = \mathbf{P}^{\mathrm{f}}\mathbf{H}^{T}\left(\mathbf{H}\mathbf{P}^{\mathrm{f}}\mathbf{H}^{T} + \mathbf{R}\right)^{-1}, \tag{7}$$

$$\mathbf{P}^{\mathrm{f}} = \frac{\mathbf{A}^{\mathrm{f}}\left(\mathbf{A}^{\mathrm{f}}\right)^{T}}{N-1}, \tag{8}$$

are respectively the Kalman gain matrix and the background error covariance matrix estimated from the ensemble anomalies.

In the following, the DEnKF will be referred to as EnKF since general conclusions of this work are independent of the flavour of the EnKF analysis scheme used.

Applying an EnKF with a large dimensional system requires few ad-hoc implementations. In order to avoid a too abrupt start of assimilation, the variance of the observation error is multiplied by a factor 8 at the first assimilation cycle and is then reduced by 1 every two months until it reaches 1 over the course of 14 months. We use the rfactor inflation scheme (Sakov et al., 2012), for which the observation error is inflated by 2 when updating the ensemble anomaly in Eq. 6. We also use pre-screening of the observation; *i.e.*, the observation error variance is inflated so that the analysis remains within

2 standard deviations of the forecast error from the ensemble mean of the forecasts (Sakov et al., 2012). We also used the upscaling method (Y. Wang et al., 2016) that handles the update of the water layers thickness (truncated Gaussian) and better preserve mass, heat and salt. For producing long reanalysis (from 1850) with SST assimilation, we use the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST2.1.0.0) available from 1850–present. This type of product is practical because it handles the biases between different data set and provide a grided ensemble SST that can be used to quantify the uncertainty. Still, it is constructed by statistical interpolation/extrapolation from the raw data and the neighbouring observation errors are highly correlated. Our assimilation code assumes the observation error to be independent (*i.e.,* $\mathbf{R}$ is diagonal) and it was therefore decided to only retain the nearest SST observation to update the water column (we speak of "strong localization"). For more details about the implementation of the EnKF within NorCPM, see Counillon et al. (2014) and Counillon et al. (2016).

4 Background error covariance matrix hybridization

The dynamical covariance matrix $\mathbf{P}_d^{\mathrm{f}}$ is estimated from the dynamic ensemble $\mathbf{X}_d$. The size of $\mathbf{X}_d$ is limited to 30 members in the current version of NorCPM. Such a small ensemble size results in spurious covariances (Anderson, 2007; Bishop & Hodyss, 2007) and rank deficiencies (Oke et al., 2007). Background error covariance hybridization was initially introduced by Hamill and Snyder (2000) to combine an EnKF with a 3DVar and bring some flow-dependency in variational data assimilation. Covariance hybridization has been used in sequential ensemble data assimilation by X. Wang et al. (2007), Counillon et al. (2009) and Gharamti et al. (2014) (hereafter referred to as EnKF-OI) as a way to limit the impact of under-sampling and rank deficiency. The background error covariance matrix combines linearly a dynamical covariance matrix $\mathbf{P}_d^{\mathrm{f}}$ with another covariance matrix $\mathbf{P}_s^{\mathrm{f}}$ (where the subscript s stands for static) computed from a climatological ensemble of size N_s, $\mathbf{X}_s$ (where $N_s \gg N_d$). That static ensemble is constructed by gathering model output before running the assimilation experiment. As such, the EnKF-OI does not increase the computational cost of the integration step and has only a limited impact on the computational cost of the analysis step (Counillon et al., 2009). In the special case of the EnKF-OI in NorCPM the CPU-time of the EnKF-OI is 7% larger than that of the standard EnKF; we discuss in Section 8 possible ways to reduce the computational cost of the EnKF-OI in NorCPM. We denote $\mathbf{P}_h^{\mathrm{f}}$ the hybrid covariance matrix:

$$\mathbf{P}_h^{\mathrm{f}} = \alpha_d \mathbf{P}_d^{\mathrm{f}} + \alpha_s \mathbf{P}_s^{\mathrm{f}}, \qquad \alpha_d, \alpha_s \geq 0 \tag{9}$$

Unless explicitly mentioned, the sum of the coefficients α_d and α_s can be different from 1.

The update of the mean and the anomalies with the EnKF-OI writes:

$$\mathbf{x}_d^{\mathrm{a}} = \mathbf{x}_d^{\mathrm{f}} + \mathbf{K}_h \left(\mathbf{y} - \mathbf{H}\mathbf{x}_d^{\mathrm{f}} \right), \tag{10}$$

$$\mathbf{A}_d^{\mathrm{a}} = \mathbf{A}_d^{\mathrm{f}} - \frac{1}{2}\mathbf{K}_h \mathbf{H} \mathbf{A}_d^{\mathrm{f}}, \tag{11}$$

where $\mathbf{K}_h$ is the hybrid Kalman gain,

$$\mathbf{K}_h = \mathbf{P}_h^{\mathrm{f}} \mathbf{H}^T \left(\mathbf{H} \mathbf{P}_h^{\mathrm{f}} \mathbf{H}^T + \mathbf{R} \right)^{-1}. \tag{12}$$

In practice, we do not compute explicitly the hybrid covariance matrix $\mathbf{P}_h$:

$$\mathbf{P}_h = \frac{\mathbf{A}_h \left(\mathbf{A}_h\right)^T}{N_h - 1}, \tag{13}$$

where $\mathbf{A}_h$ stands for the hybrid anomalies (Counillon et al., 2009):

$$\mathbf{A}_h = \sqrt{N_h - 1} \left[\sqrt{\frac{\alpha_d}{N_d - 1}} \mathbf{A}_d, \sqrt{\frac{\alpha_s}{N_s - 1}} \mathbf{A}_s \right], \tag{14}$$

and $[.,.]$ stand for the concatenation of two sets of vectors. Therefore, the EnKF-OI is implemented within NorCPM following Evensen (2003) and Sakov et al. (2010).

5 Adaptive covariance hybridization

5.1 Explicit optimality of the hybridization coefficients

In this section, we build on the work of Ménétrier et al. (2015); Ménétrier and Auligné (2015) and adapt a method proposed in Ménétrier (2021) to determine the optimal hybridization coefficients.

We take advantage of the strong localization setting (see Section 3) to consider only the covariance between a single point at the surface of the ocean, and the multivariate state vector of the water column below so that the covariance matrix in each point reduces to a vector. In the following, we note $\mathbf{P}$ the true covariance vector at a given point (which would be obtained with an infinite ensemble). We use the subscript i to refer to the i^{th} element of the covariance vector (*e.g.*, $\mathbf{P}_i$ for the i^{th} element of the true covariance vector). We define the scalar product of two random vectors X and Y as $\mathbb{E}\left[X \cdot Y\right] = \sum_{i=1}^{p} \mathbb{E}\left[X_i Y_i\right]$. We assume that $\mathbf{P}_d$ and $\mathbf{P}$ are realizations of two independent random processes and that the sampling error of $\mathbf{P}_d$, *i.e.*, $\mathbf{P}_d - \mathbf{P}$, is unbiased, and orthogonal to the truth, (see Ménétrier et al. (2015), Eqs. (20a)-(20b)). Therefore:

$$\mathbb{E}\left[\mathbf{P}_d - \mathbf{P}\right] = 0, \tag{15a}$$

$$\mathbb{E}\left[(\mathbf{P}_d - \mathbf{P}) \cdot \mathbf{P}\right] = 0. \tag{15b}$$

The optimal hybridization coefficients (α_d, α_s) are defined as those minimizing, in a statistical sense, the square of the error between $\mathbf{P}_h$ and $\mathbf{P}$, *i.e.*, (α_d, α_s) minimize the function e:

$$e(\alpha_d, \alpha_s) = \mathbb{E}\left[\left\|\mathbf{P}_h - \mathbf{P}\right\|^2\right] = \mathbb{E}\left[\left\|\alpha_d \mathbf{P}_d + \alpha_s \mathbf{P}_s - \mathbf{P}\right\|^2\right], \tag{16}$$

where $\|.\|$ stands for the L_2-norm of a vector. It can be shown (see Appendix A) that minimizing the function e is equivalent to solving a system of two equations, with unknown α_d and α_s, whose solution is given by:

$$(\alpha_d, \alpha_s) = \left(\frac{n_d}{\Delta}, \frac{n_s}{\Delta}\right), \tag{17}$$

where:

$$\Delta = \left\|\mathbf{P}_s\right\|^2 \mathbb{E}\left[\left\|\mathbf{P}_d\right\|^2\right] - \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right]^2, \tag{18}$$

$$n_d = \left\|\mathbf{P}_s\right\|^2 \mathbb{E}\left[\left\|\mathbf{P}\right\|^2\right] - \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right]^2, \tag{19}$$

$$n_s = \left(\mathbb{E}\left[\left\|\mathbf{P}_d\right\|^2\right] - \mathbb{E}\left[\left\|\mathbf{P}\right\|^2\right]\right) \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right]. \tag{20}$$

5.2 Properties

The properties highlighted by Ménétrier et al. (2015) in the case of the dual optimization of localization and hybridization hold here:

1. **Behavior of the hybridization coefficients:** if the static covariance is multiplied by a factor λ, then α_s is divided by λ, while α_d remains unchanged. As such, it is not necessary to tune the static covariance with a scalar a priori, as done in Evensen (2003), Oke et al. (2008) and Counillon et al. (2009).
2. **Asymptotic behavior:** with an infinite ensemble, $\mathbb{E}\left[\|\mathbf{P}_d\|^2\right] = \mathbb{E}\left[\|\mathbf{P}\|^2\right]$. We can replace $\mathbb{E}\left[\|\mathbf{P}_d\|^2\right]$ by $\mathbb{E}\left[\|\mathbf{P}\|^2\right]$ in Eqs. (19)-(20), and obtain $(\alpha_d, \alpha_s) = (1, 0)$ as expected – there is no need for hybridization.
3. **Benefits of hybridization:** whatever the choice of the static covariance (see Appendix B),
$$e(1,0) \geq e(\alpha_d, \alpha_s), \tag{21}$$
showing the superiority of the hybrid scheme over the standalone EnKF.
4. **Optimality condition:** we can show that $\mathbf{P}_h$ verifies the following optimality condition:
$$\begin{cases} \dfrac{\partial e}{\partial \alpha_d} = 0 \\ \dfrac{\partial e}{\partial \alpha_s} = 0 \end{cases} \Leftrightarrow \mathbb{E}\left[(\mathbf{P}_d - \mathbf{P}_s) \cdot (\mathbf{P}_h - \mathbf{P})\right] = 0. \tag{22}$$
This means from a statistical point of view that $\mathbf{P}_h$ is the orthogonal projection of the true covariance $\mathbf{P}$ onto a sub-space defined as $\mathbf{P}_d + \mu\,(\mathbf{P}_s - \mathbf{P}_d)$ where $\mu \in \mathbb{R}$, see Fig. 1.

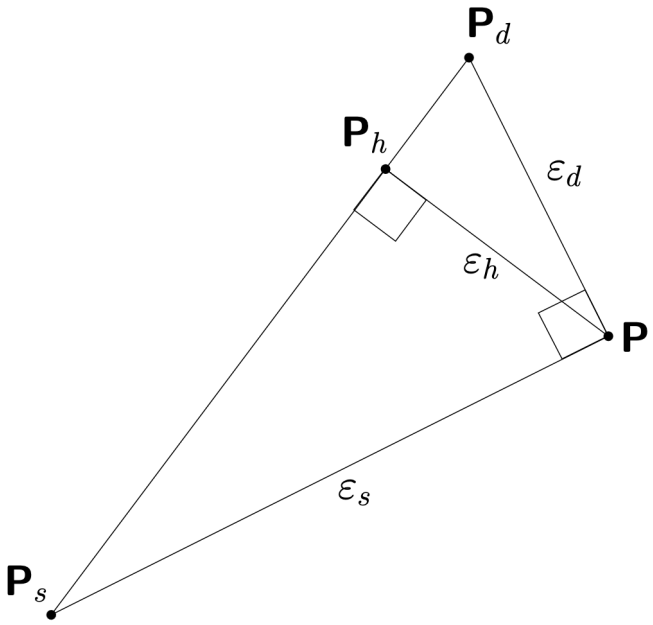


Figure 1. Geometrical representation of the orthogonal projection of the true covariance $\mathbf{P}$ onto the sub-space generated by the dynamic covariance $\mathbf{P}_d$ and the static covariance $\mathbf{P}_s$.

Here are some remarks:

1. if $\alpha_d \neq 1$, $\mathbf{P}_h$ can be interpreted as the linear interpolation between $\mathbf{P}_d$ and $\lambda\mathbf{P}_s$, where:
$$\lambda = \frac{\alpha_s}{1-\alpha_d}, \tag{23}$$
thus, $\lambda \geq 1$ (resp ≤ 1) is equivalent to $\alpha_d+\alpha_s \geq 1$ (resp. ≤ 1). λ acts as an inflation or deflation term for the matrix $\mathbf{P}_s$ and $\mathbf{P}_h$ is the linear interpolation between $\mathbf{P}_d$ and the inflated/deflated $\mathbf{P}_s$.
2. From Eqs (15a)-(15b), it follows that the errors $\mathbf{P}_d-\mathbf{P}$ and $\mathbf{P}_s-\mathbf{P}$ are uncorrelated: $\mathbb{E}\left[(\mathbf{P}_d-\mathbf{P})\cdot(\mathbf{P}_s-\mathbf{P})\right]=0$. As a consequence, the triplet $(\mathbf{P}_d, \mathbf{P}, \mathbf{P}_s)$ forms a triangle rectangle in $\mathbf{P}$, see Fig. 1.
3. The numerator of α_d, n_d, can be interpreted as a measure of the collinearity of the static covariance $\mathbf{P}_s$, and the expectation of the true covariance $\mathbf{P}$. Hence, n_d is equal to 0 if and only if $\mathrm{Var}(\mathbf{P})=0$ and $\mathbf{P}$ and $\mathbf{P}_s$ are collinear (or linearly dependent). Conversely, if $\mathbb{E}\left[\mathbf{P}_d\cdot\mathbf{P}_s\right]=0$, then $\mathbf{P}$ and $\mathbf{P}_s$ are orthogonal (see Appendix C).
4. It follows from Eq. (15b) that (see Appendix B, Eqs. (B1)-(B2)):
$$\mathbb{E}\left[\|\mathbf{P}_d-\mathbf{P}\|^2\right] = \mathbb{E}\left[\|\mathbf{P}_d\|^2\right] - \mathbb{E}\left[\|\mathbf{P}\|^2\right] \geq 0. \tag{24}$$
Therefore, the difference $\mathbb{E}\left[\|\mathbf{P}_d\|^2\right]-\mathbb{E}\left[\|\mathbf{P}\|^2\right]$ can be interpreted as a measure of the optimality (or the non-optimality) of the covariance function computed from the dynamic ensemble $\mathbf{P}_d$: the smaller the difference, the smaller the distance $\|\mathbf{P}_d-\mathbf{P}\|$ in a statistical sense. Conversely, the larger the difference, the larger the distance $\|\mathbf{P}_d-\mathbf{P}\|$.
5. It follows from remarks 3 and 4 that the hybridization coefficients α_d and α_s are the combination of the optimality of the dynamic covariance function $\mathbf{P}_d$ and the collinearity/orthogonality of the static covariance $\mathbf{P}_s$ and the true covariance $\mathbf{P}$.
6. As a consequence of Eq.(24) and Appendix C, $0 \leq \alpha_d \leq 1$. We can not give similar upper and lower boundaries for the values of α_s, as the term $\mathbb{E}\left[\mathbf{P}_d\cdot\mathbf{P}_s\right]$ can be negative and we do not know its lower bound. Numerical simulations showed that this term is almost always positive (not shown). We can just say that α_s is maximal when Δ is minimal and therefore $\mathbb{E}\left[\mathbf{P}_d\cdot\mathbf{P}_s\right]$ is maximal. In that case, α_d is minimal.

5.3 Practical implementation

Quantities in Eq. (19) and Eq. (20) can not be computed directly as they are a function of $\mathbb{E}\left[\|\mathbf{P}\|^2\right]$, $\mathbb{E}\left[\mathbf{P}_d\cdot\mathbf{P}_s\right]$, and $\mathbb{E}\left[\|\mathbf{P}_d\|^2\right]$ that are unknown.

Nonetheless, the sampling theory developed in Ménétrier (2021) allows us to express $\mathbb{E}\left[\mathbf{P}_i^2\right], i=1,\ldots,p$ as a function of the covariance and variance of the dynamic ensemble. Using Eq. (101b) of Ménétrier (2021) one can write:

$$\mathbb{E}\left[\mathbf{P}_i^2\right] = \frac{(N_d-1)^2}{(N_d-2)(N_d+1)}\mathbb{E}\left[\mathbf{P}_{di}^2\right] - \frac{N_d-1}{(N_d-2)(N_d+1)}\mathbb{E}\left[\mathbf{v}_{di}\mathbf{v}_{d1}\right], \tag{25}$$

where $\mathbf{v}_{di}$ is the variance of the dynamic ensemble for the i-th element of the model state. The quantities n_d, n_s, and Δ are then fully expressed as a function of the static covariance and the expectation of the variance and covariance of the dynamic ensemble.

The expectation terms in Eq. (18), (19), (20), and (25) are estimated under a simplifying assumption of "local homogeneity": it is assumed that in an area surrounding the water column, the vertical covariance functions are representative of the covariance function of the water column. The expectation terms are then estimated as the average

of the surrounding vertical covariance functions. Numerical tests have shown that at least 500 covariance functions are necessary to compute reliable statistics. In this work, we consider covariance functions in a radius up to 1000 km around the water column, which usually provides between 500 and 1000 covariance functions to compute the expectation terms.

In order to limit the computational burden of estimating the hybridization coefficients, they are computed on a subgrid of the domain (every 5 grid cells). The hybridization coefficients are then interpolated to the remaining wet points using linear interpolation of the neighbouring wet points.

The hybridization coefficients are estimated based on both temperature and salinity. Doing so yields a lower root mean square error than when computing the hybridization coefficients solely based on the temperature (not shown).

6 Experimental design and evaluation metrics

6.1 Experimental design

The free ensemble run (hereafter referred to as FREE), consists of 30 members run with transient forcing from 1850 to 2014. The true run (hereafter referred to as TRUE) is created by spawning one member (adding noise to surface temperature) on member 1 of FREE in 1960 and running it up to 2010. It was verified in Y. Wang et al. (2022) that TRUE and member 1 of FREE were fully de-synchronised at the start of the experiment in 1980. The synthetic observations of SST are generated by adding white noise to the monthly SST of TRUE. The amplitude of the noise is set equal to the observation uncertainty (in space and time) of HadISST2. As in the real framework for assimilation of SST, we do not use SST data under sea ice.

We produce reanalyses with monthly assimilation of SST observations from January 1980 to December 2010. All experiments start with the same initial dynamic ensemble (taken from FREE in January 1980). The static ensemble is made from the monthly restarts of a 315 years long stable pre-industrial run (monthly varying static ensemble). The experiments are separated into three categories:

- **EnKF:** the standard EnKF used in NorCPM (Counillon et al., 2016).
- **Standard hybrid:** a constant and global hybridization coefficients (see Section 4). The sum of α_d and α_s is 1. We performed 7 reanalyses with $\alpha_d = 0, 0.2, 0.4, 0.6, 0.8, 0.9, 1$. The case where $\alpha_d = 0$ is equivalent to an ensemble of EnOI, and the case where $\alpha_d = 1$ is equivalent to the standard EnKF.
- **Adaptive hybrid:** the hybridization coefficients are estimated at each assimilation cycle, they vary spatially and their sum is not imposed equal to 1 (see Section 5).

6.2 Evaluation metrics

The accuracy of the reanalyses is estimated based on the root mean square error (RMSE). The RMSE is computed as:

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{N} \omega_i \left(\mathbf{x}_i - \mathbf{x}_i^{\mathrm{t}}\right)^2}. \tag{26}$$

In the following, the RMSE is computed either over a time series at a given point (in which case $\omega_i = \frac{1}{N}$), or over the whole domain at a given time (in which case ω_i is the relative size of the grid cell).

In order to easily compare RMSE between the nine different schemes (see Section 6.1), we introduce the Mean Skill Score of one configuration i, MSS_i. It is the relative reduction of RMSE compared to the mean of the RMSE of the nine configurations, Eq. (27):

$$\mathrm{MSS}_i = 1 - \frac{\mathrm{RMSE}_i}{\frac{1}{9}\sum_{j=1}^{9}\mathrm{RMSE}_j}, \tag{27}$$

where RMSE_i is the RMSE of one of the schemes. The MSS is 1 if the scheme is perfect (RMSE is equal to 0), between 0 and 1 if the scheme performs better than the mean of the other schemes and negative otherwise.

Another important metric to evaluate the relative efficiency of different data assimilation schemes is to consider the "degrees of freedom for signal" (DFS, Cardinali et al., 2004; Wahba et al., 1995). It can be interpreted as the number of modes of variability reduced from the ensemble by the assimilation (*i.e.* the assimilation change). The DFS is defined as follows:

$$\mathrm{DFS} = \frac{\partial \mathbf{H}\mathbf{x}_d^{\mathrm{a}}}{\partial \mathbf{y}} = \mathrm{Tr}\left(\mathbf{K}\mathbf{H}\right) \tag{28}$$

The DFS is between 0 (*i.e.,* the observations have no impact on the ensemble), and the total number of degrees of freedom (*i.e.,* observations has collapsed the number of modes of variability into a single one, Xie et al., 2018). The total number of degrees of freedom is the minimum between the ensemble size and the number of observations used for the local assimilation. In NorCPM, in the context of strong localisation (where we retain only the nearest observation, see Section 3), it implies that the DFS is between 0 and 1 (independently of the ensemble size). This allows for an inter-comparison of the DFS even though the schemes have different ensemble sizes.

7 Results

7.1 Stability of the adaptive covariance hybridization

The adaptive covariance hybridization method (see Section 5) estimates adaptive hybridization coefficients both in space and time.

Figure 2 shows the time series of globally averaged α_d, α_s, and $\alpha_d+\alpha_s$ (sea ice-covered points where there are no SST data are masked). After a spin-up period of approximately three years, the averaged values of the hybridization coefficients converge to a global average of 0.7 for α_d, and 0.175 for α_s. This shows that the mean values of α_d and α_s are stable in time and display a limited temporal variability despite an important spatial variability, and so does the sum $\alpha_d+\alpha_s$. Hence, the mean values of α_d and α_s computed in specific basins show similar behaviour and converge within 3 years (not shown). The global averaged value of $\alpha_d + \alpha_s$ is roughly 0.875. Following remark 1 in Section 5.2, it implies that the static ensemble has a larger error variance than the error growth within one assimilation cycle and needs to be reduced - in agreement with Oke et al. (2008), Counillon and Bertino (2009), and Evensen (2003).

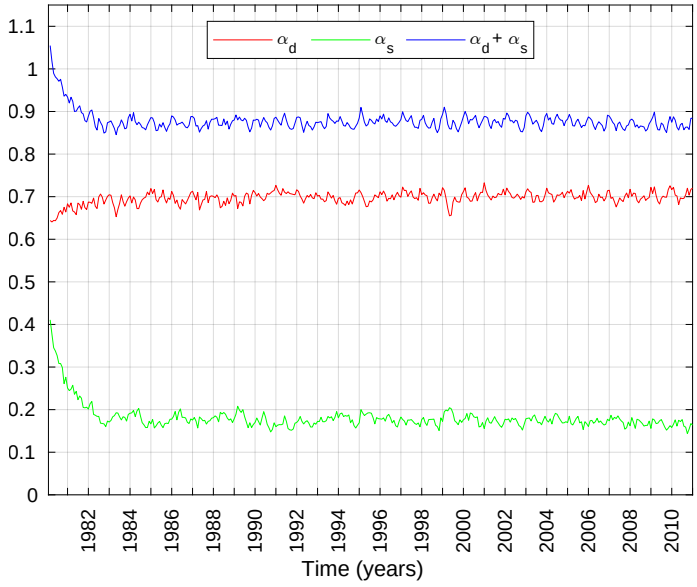


Figure 2. Time series of the global average values of α_d, α_s, and $\alpha_d + \alpha_s$. Ice-covered regions are masked.

Figures 3 shows the pointwise averaged map of α_d and α_s computed over the years 1983 to 2010 for January and July. The values of α_d display an important spatial variability with values ranging approximately from 0.4 up to 1. Regions where α_d is small coincide with places where α_s is large. The spatial variations of the values of α_d and α_s (Section 5.2) can be explained from the perspective of the optimality of the dynamic covariance $\mathbf{P}_d$ (depending on the sampling error in the dynamical ensemble), and the collinearity between the static covariance $\mathbf{P}_s$ and the true covariance $\mathbf{P}$ (meaning that static covariance is sufficient). Larger values of α_d are found in locations where the dynamic is non-stationary and internal variability is large; *e.g.,* in the Northern part of the Atlantic Ocean (Gulf Stream pathway, Subpolar Gyre, near the ice edge), the North Pacific, El Niño–Southern Oscillation and in the Southern Ocean. Conversely, there are relatively low values of α_d in the Indian Ocean where variability is primarily externally forced (Guemas et al., 2013), the decadal fluctuations are less pronounced than in the Atlantic or the Pacific Ocean and where the Pacific Ocean teleconnections dominate the regionally driven variability (Frankcombe et al., 2015). In the tropical Atlantic, the model is performing very poorly and has no skill (Counillon et al., 2021); it is thus not surprising that α_d is also low.

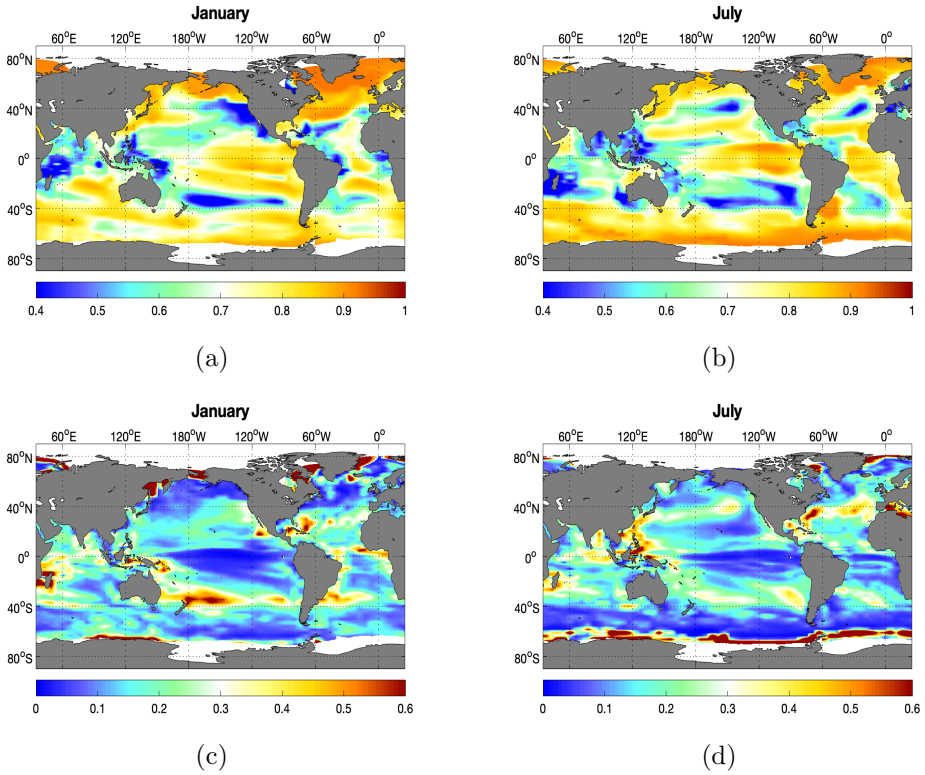
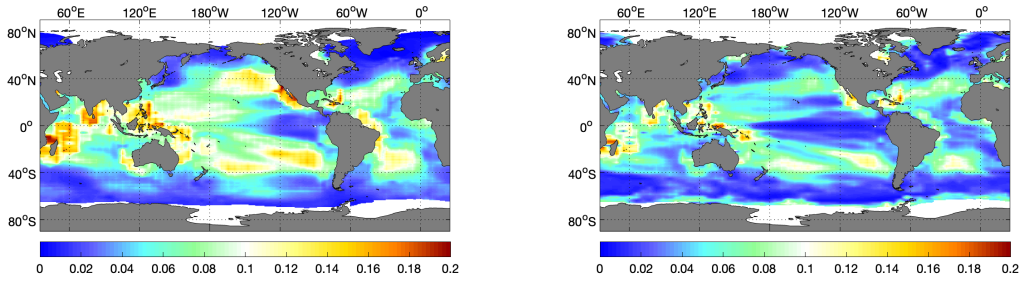


Figure 3. Pointwise averaged estimate of α_d (top row), and of α_s (bottom row) computed over 1983–2010 for the months of January and July.

In Fig. 4, we analyse the interannual de-seasoned standard deviation of the hybrid coefficient beyond year 3 (once it has converged). We can see that the variability is very small, except in a few places, *e.g.* in the Arctic, in Indian Ocean and in the tropical Atlantic and Pacific Gyre. In those places, the performance between the standard hybrid coefficient method is relatively small (not shown).

7.2 Intercomparison of the performance of the EnKF and the hybrid covariance schemes

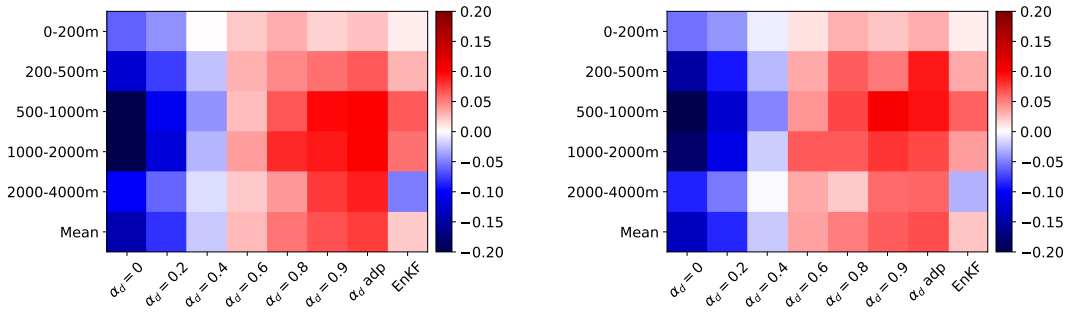
Figure 5 shows the MSS (see Section 6.2) of ocean heat and salt content for the different schemes at different depth ranges (0-200 m, 200-500 m, ..., 2000-4000 m). We include the EnKF ($\alpha_d = 1$ and $\alpha_s = 0$) and the ensemble of EnOI ($\alpha_d = 0$ and $\alpha_s = 1$) as particular cases of the standard hybrid covariance method. A red cell (resp. blue cell) indicates that the scheme provides a reduction (resp. an increase) of RMSE compared to the average performance of all the schemes for a given depth range. For example, the adaptive hybrid and the standard hybrid scheme with $\alpha_d = 0.9$ reduces the RMSE of the temperature at depth 500-1000 m by 10% compared to the average performance, while the standard hybrid with $\alpha_d = 0$ increases the RMSE at the depth 200-500 m by 15%. The results for heat and salt content are very similar. As expected, the EnKF is outperforming the ensemble of EnOI (*i.e.,* $\alpha_d = 0$), showing the superiority of flow-dependent covariance over static covariance. It also shows the importance of tuning the hybrid coefficient as for a large span of standard hybrid coefficient values, the hybrid covariance methods perform poorer than the EnKF. When α_d is larger than 0.8 ($\alpha_d = 0.9$ being optimal), the standard hybrid covariance outperforms the EnKF; notably between 2000 and 4000 m. In the latter, a bias is gradually increasing due to spurious covariance at



(a) (b)

Figure 4. Standard-deviation of the de-seasoned values of α_d (a), and of α_s (b) computed over 1983–2010.

454 depth (Y. Wang et al., 2022; Bethke et al., 2018). The adaptive hybrid covariance method
455 performs best at nearly all depth levels for heat and salt content. In the following, we
456 will therefore present the adaptive hybrid and assess the spatial distribution of the im-
457 provements over the EnKF, but results with the best standard hybrid are nearly com-
458 parable (not shown).



(a) (b)

Figure 5. MSS of all the schemes for temperature (a) and salinity (b) at different depth intervals. $\alpha_d = 0, \alpha_d = 0.2, \ldots, \alpha_d = 0.9$ refer to the hybridization coefficient of the standard hybrid. $\alpha_d = 0$ is the ensemble of EnOI, full static case, and $\alpha_d = 1$ is the EnKF – the default scheme used in NorCPM. "α adp" stands for the adaptive hybridization scheme. The warm colour indicates that the scheme performs better than the average skill of all systems.

The adaptive hybrid and the EnKF achieve similar performance in the top 1000 m (Figure 5), and we focus on performance below this depth range. We compare the reduction of RMSE of the EnKF and the adaptive hybrid compared to that of FREE for two depths range.

Between 1000 and 2000 m (Fig. 6), the EnKF reduces the error overall (warmer value is predominant) compared to FREE. Still, there are few places where it increases the RMSE of temperature, *e.g.*: in the North Pacific, the subtropical Atlantic, and near the Weddell Sea. Results are relatively comparable for salt content (see Fig. 6-b). The overall pattern is similar with the adaptive hybrid. Still, it yields further improvement, as in the North Atlantic subpolar gyre and it mitigates the degradation in the aforementioned regions. The degradation in the Weddell Sea is nearly completely removed.

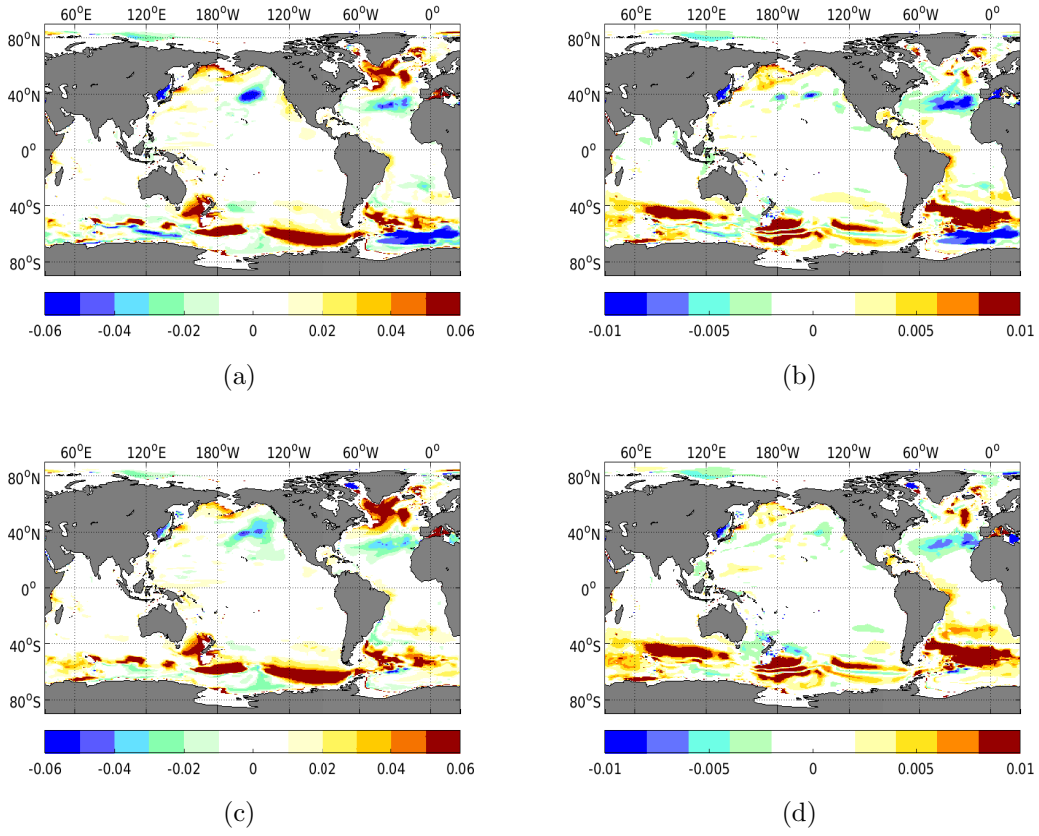


Figure 6. RMSE difference between FREE and the EnKF (a,b), and FREE and the adaptive hybrid (c,d) between 1000 and 2000 m depth for the temperature (left column) and the salinity (right column). Warm colour indicates that assimilation reduces error compared to FREE.

Between 2000 and 4000 m (Fig. 7), the EnKF degrades overall performance compared to FREE. The degradation is larger in the North Pacific, the North Atlantic, and the Southern Ocean for both the temperature and the salinity. The improvements are also limited to the South Atlantic Ocean. The adaptive hybrid corrects or mitigates these biases. Some degradation remains (in the North Atlantic subpolar gyre, the Sea of Japan in particular for salinity), but the assimilation yields an overall improvement over FREE.

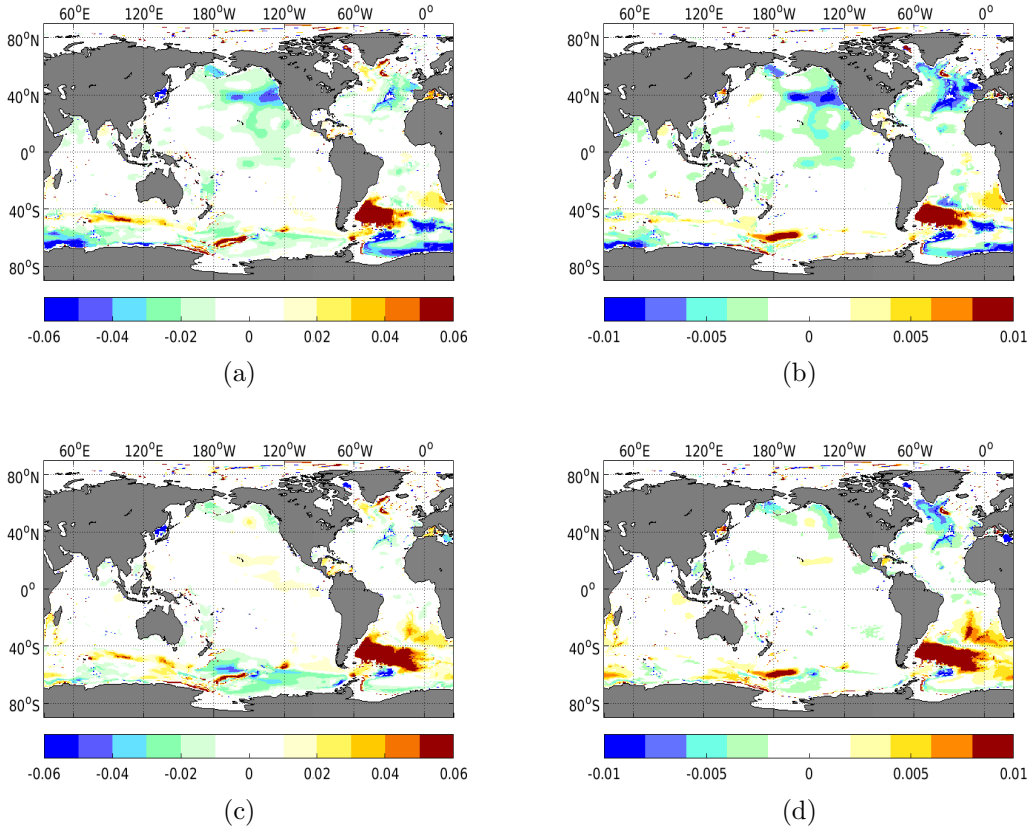
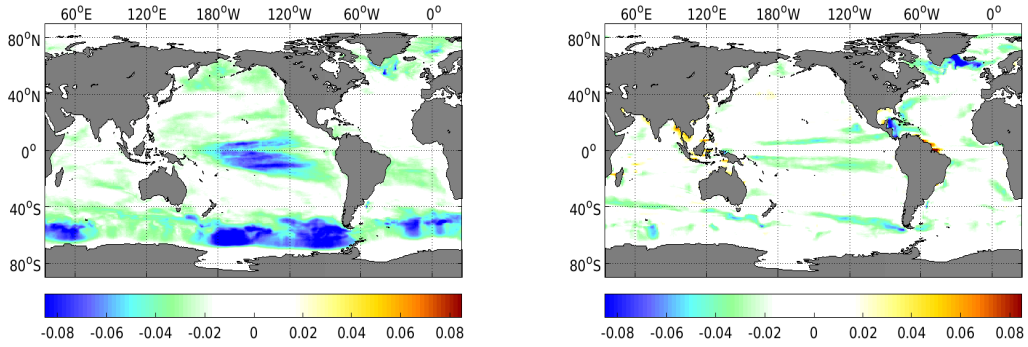


Figure 7. Same as fig. 6 but for 2000–4000 m depth interval.

An ideal assimilation system achieves minimal error while making the smallest change possible during the assimilation. Figure 8 shows the difference of DFS (that quantifies the assimilation change) between the EnKF and the best standard hybrid ($\alpha_d = 0.9$) (panel a) and between the EnKF and the adaptive hybrid (panel b). The standard hybrid has a larger DFS value than the EnKF (negative values), implying that the data assimilation induces more change. This is most notable in the Southern Ocean and the tropical Pacific. In the Southern Ocean, the standard hybrid covariance method performs better than the EnKF, so it can be argued that the larger corrections are beneficial. However, in the tropical Pacific, the ΔRMSE of the two remains quite close, meaning the analysis induces more changes without improving performance. On the contrary, the adaptive hybrid, Fig. 8-(b), has a DFS close to that of the EnKF. There are some slight differences (in the North Pacific, the North Atlantic, and the Southern Ocean), with a maximum in the Irminger Sea, where it strongly outperforms the EnKF (*e.g.,* 1000–2000 m). It implies that the adaptive hybrid induces only change where this yields improved performance.



(a) (b)

Figure 8. Difference of DFS between the EnKF and the standard hybrid covariance with $\alpha_d = 0.9$ (a), and between EnKF and adaptive hybrid covariance (b). The cold colour indicates that the hybrid covariance yields a larger reduction of DFS than with the EnKF.

8 Discussion and conclusion

In this work, we compare two different versions of hybrid covariance data assimilation with the standard EnKF for producing climate reanalysis. We use the Norwegian Climate Prediction Model (NorCPM) and work in an idealised twin experiment framework. The reanalyses are performed with sole assimilation of SST for the period 1980-2010. In the first hybrid coefficient method, the hybridization coefficients are tuned empirically to optimize the performance, while in the second, the hybridization coefficients are estimated adaptively, both in space and time. The two hybrid coefficient methods outperform the standard EnKF and mitigate the degradation it introduces in the intermediate and deep ocean compared to unassimilated simulations. The adaptive performs best and is doing so by making smaller corrections than the standard hybrid. The hybridization coefficients with the adaptive hybrid are converging quickly (less than 3 years) to stable values and only show small seasonal variations.

Other alternatives have been developed in parallel to address the sampling error with the EnKF in NorCPM – namely the isopycnal vertical localization (Y. Wang et al., 2022). The latter limits the assimilation update of temperature and salinity to a fixed isopycnal level and was shown to mitigate the degradation seen in the standard EnKF. Combining the two approaches is straightforward and will be tested in the future. However, the isopycnal vertical localization detailed in Y. Wang et al. (2022) was tuned for an ensemble size of 30 members, while now the ensemble size is much larger (350 members). The vertical tapering will thus need to be revised.

The adaptive hybrid coefficients method is slightly more expensive than the standard hybrid as it requires additional computation related to the estimation of the hybrid coefficient at each assimilation step. Here, the hybridization coefficients are estimated at every 5 grid cells, but we could have estimated them at every 10 grid cells (reducing the cost by 4) with a comparable solution (not shown). Furthermore, as these coefficients converge rapidly to stable estimates (within 3 model years, showing only a weak seasonal variability). They could be stored and directly used instead of being recalculated every time. As such, we do not consider that the additional computational cost would be much larger than the standard hybrid, which also has an additional cost (empirical estimation of the global coefficient).

In this study, the estimation of the hybridization coefficients in the adaptive method is constant in the vertical. Nonetheless, adapting the method to estimate different hybridization coefficients for different vertical levels or variables would be relatively trivial. Furthermore, we tested the method for the particular case of assimilation of SST objective analysis, where we update a single water column with a single observation, (*i.e.,* "strong localization"). For the assimilation of temperature/salinity profile data, the observation error is uncorrelated, and a larger localisation radius is used in NorCPM, the method can be adjusted following Ménétrier (2021).

A critical assumption made with the adaptive hybrid covariance method is that models are unbiased and that for an infinite ensemble size, the ensemble covariance matrix converges with the true covariance matrix. These assumptions fall apart with Earth System Models with considerable biases (Palmer & Stevens, 2019). It remains to be verified how the method would perform in a real framework. Verification of the method in a real framework, both for coupled reanalysis and testing the impact on prediction initialised from it, will be tested in the future.

Appendix A Minimization of the function e

The function e is defined as:

$$e(\alpha_d, \alpha_s) = \mathbb{E}\left[\left\|\alpha_d \mathbf{P}_d + \alpha_s \mathbf{P}_s - \mathbf{P}\right\|^2\right]. \tag{A1}$$

By linearity of the expectation operator and by definition of the L_2-norm $\|.\|$, we can write:

$$\begin{aligned} e(\alpha_d, \alpha_s) &= \alpha_d^2 \mathbb{E}\left[\left\|\mathbf{P}_d\right\|^2\right] + \alpha_s \left\|\mathbf{P}_s\right\|^2 + \mathbb{E}\left[\left\|\mathbf{P}\right\|^2\right] \\ &\quad + 2\alpha_d\alpha_s \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right] - 2\alpha_d \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}\right] - 2\alpha_s \mathbb{E}\left[\mathbf{P} \cdot \mathbf{P}_s\right]. \end{aligned} \tag{A2}$$

It follows from Eq. (15b) that:

$$\mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}\right] = \mathbb{E}\left[\left\|\mathbf{P}\right\|^2\right] \tag{A3}$$

Replacing Eq. (A3) in Eq. (A2), we obtain the following expression of e:

$$e(\alpha_d, \alpha_s) = \alpha_d^2 \mathbb{E}\left[\left\|\mathbf{P}_d\right\|^2\right] + \alpha_s^2 \left\|\mathbf{P}_s\right\|^2 + (1 - 2\alpha_d)\mathbb{E}\left[\left\|\mathbf{P}\right\|^2\right] + 2\alpha_s(\alpha_d - 1)\mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right]. \tag{A4}$$

e being a quadratic function of two variables, α_d and α_s, with positive coefficients associated to α_d^2 and α_s^2, its has a unique minimum where both the partial derivatives with respect to α_d and α_s are null. Hence, minimizing the function e is equivalent to solve the following system of two equations:

$$\begin{cases} \dfrac{\partial e(\alpha_d, \alpha_s)}{\partial \alpha_d} = 0 \\ \dfrac{\partial e(\alpha_d, \alpha_s)}{\partial \alpha_s} = 0 \end{cases} \tag{A5}$$

The partial derivatives of e with respect to α_d and α_s are given by Eq. (A6) and (A7):

$$\frac{\partial e(\alpha_d, \alpha_s)}{\partial \alpha_d} = 2\alpha_d \mathbb{E}\left[\left\|\mathbf{P}_d\right\|^2\right] + 2\alpha_s \mathbb{E}\left[\mathbf{P}_s \cdot \mathbf{P}_d\right] - 2\mathbb{E}\left[\left\|\mathbf{P}\right\|^2\right], \tag{A6}$$

$$\frac{\partial e(\alpha_d, \alpha_s)}{\partial \alpha_s} = 2\alpha_d \mathbb{E}\left[\mathbf{P}_s \cdot \mathbf{P}_d\right] + 2\alpha_s \left\|\mathbf{P}_s\right\|^2 - 2\mathbb{E}\left[\mathbf{P}_s \cdot \mathbf{P}_d\right]. \tag{A7}$$

From which it follows that minimizing the function e is equivalent to solving the system of two equations:

$$\begin{cases} \alpha_d \mathbb{E}\left[\left\|\mathbf{P}_d\right\|^2\right] + \alpha_s \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right] = \mathbb{E}\left[\left\|\mathbf{P}\right\|^2\right] \\ \alpha_d \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right] + \alpha_s \left\|\mathbf{P}_s\right\|^2 = \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right], \end{cases} \tag{A8}$$

Appendix B Benefits of hybridization

By definition of the function e:

$$e(1, 0) = \mathbb{E}\left[\left\|\mathbf{P}_d - \mathbf{P}\right\|^2\right]. \tag{B1}$$

By linearity of the expectation operator and by definition of the L_2-norm, we can show that

$$e(1,0) = \mathbb{E}\left[\|\mathbf{P}_d\|^2\right] + \mathbb{E}\left[\|\mathbf{P}\|^2\right] - 2\mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}\right]. \tag{B2}$$

Replacing Eq. (A3) in Eq. (B2) we get:

$$e(1,0) = \mathbb{E}\left[\|\mathbf{P}_d\|^2\right] - \mathbb{E}\left[\|\mathbf{P}\|^2\right] \tag{B3}$$

From which we can express the difference of the errors between the EnKF and the hybrid scheme:

$$e(1,0) - e(\alpha_d,\alpha_s) = (1-\alpha_d^2)\mathbb{E}\left[\|\mathbf{P}_d\|^2\right] - 2(1-\alpha_d)\mathbb{E}\left[\|\mathbf{P}\|^2\right] - \alpha_s^2\|\mathbf{P}_s\|^2 - 2\alpha_s(\alpha_d-1)\mathbb{E}\left[\mathbf{P}_s \cdot \mathbf{P}_d\right]. \tag{B4}$$

For the sake of simplicity, we note:

$$\begin{aligned} \alpha &= \mathbb{E}\left[\|\mathbf{P}\|^2\right], && \text{(B5)}\\ \beta &= \|\mathbf{P}_s\|^2, && \text{(B6)}\\ \gamma &= \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right], && \text{(B7)}\\ \delta &= \mathbb{E}\left[\|\mathbf{P}_d\|^2\right]. && \text{(B8)} \end{aligned}$$

Eq. (B4) rewrites:

$$e(1,0) - e(\alpha_d,\alpha_s) = (1-\alpha_d^2)\delta - 2(1-\alpha_d)\alpha - \alpha_s^2\beta - 2\alpha_s(\alpha_d-1)\gamma. \tag{B9}$$

Given that $\alpha_d = \frac{n_d}{\Delta}$ and $\alpha_s = \frac{n_s}{\Delta}$:

$$e(1,0) - e(\alpha_d,\alpha_s) = \frac{\left(\Delta^2 - n_d^2\right)\delta - 2\left(\Delta^2 - \Delta n_d\right)\alpha - n_s^2\beta - 2n_s\left(n_d - \Delta\right)\gamma}{\Delta^2}. \tag{B10}$$

Δ^2 being positive, showing that $e(1,0) - e(\alpha_d,\alpha_s) \geq 0$ is equivalent to show that:

$$\left(\Delta^2 - n_d^2\right)\delta - 2\left(\Delta^2 - \Delta n_d\right)\alpha - n_s^2\beta - 2n_s\left(n_d - \Delta\right)\gamma \geq 0. \tag{B11}$$

On the other hand, n_d, n_s, and Δ write:

$$\begin{aligned} n_d &= \alpha\beta - \gamma^2, && \text{(B12)}\\ n_s &= \gamma\delta - \alpha\gamma, && \text{(B13)}\\ \Delta &= \beta\delta - \gamma^2. && \text{(B14)} \end{aligned}$$

Replacing n_d, n_s, and Δ by their expression given by Eqs. (B12), (B13), and (B14) in the left hand side of Eq. (B11), and developing all the terms we can show that Eq. (B11) is verified if and only if:

$$\beta(\beta\delta - \gamma^2)(\delta - \alpha)^2 \geq 0. \tag{B15}$$

$\beta \geq 0$ as a sum of squares and $(\delta - \alpha)^2 \geq 0$ as a square. Showing that $e(1,0) - e(\alpha_d, \alpha_s) \geq 0$ is then equivalent to show that $\beta\delta \geq \gamma^2$ i-e $\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}_d\|^2\right] - \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right]^2 \geq 0$. This inequality holds, see Appendix C, Eq. (C10), and replacing $\mathbf{P}$ by $\mathbf{P}_d$. As a consequence, $e(1,0) \geq e(\alpha_d, \alpha_s)$.

Appendix C Collinearity/orthogonality of $\mathbb{E}[\mathbf{P}]$ and $\mathbf{P}_s$

We write $\mathbf{P}_d$ as $\mathbf{P}$ plus some error ε, and we assume that the static covariance $\mathbf{P}_s$ is independent from that error, i-e: $\mathbb{E}\left[\mathbf{P}_s \cdot \varepsilon\right] = 0$. Therefore:

$$\mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right] = \mathbb{E}\left[\mathbf{P} \cdot \mathbf{P}_s + \mathbf{P}_s \cdot \varepsilon\right] = \mathbb{E}\left[\mathbf{P} \cdot \mathbf{P}_s\right] + \mathbb{E}\left[\mathbf{P}_s \cdot \varepsilon\right] = \mathbb{E}\left[\mathbf{P} \cdot \mathbf{P}_s\right]. \tag{C1}$$

Thus, if $\mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right] = 0$ then $\mathbb{E}\left[\mathbf{P} \cdot \mathbf{P}_s\right] = 0$ and $\mathbf{P}$ and $\mathbf{P}_s$ are orthogonal in a statistical sense.

Following Eq. (C1), we have:

$$\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] - \mathbb{E}\left[\mathbf{P}_d \cdot \mathbf{P}_s\right]^2 = \|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] - \mathbb{E}\left[\mathbf{P} \cdot \mathbf{P}_s\right]. \tag{C2}$$

By definition of $\|\mathbf{P}_s\|^2$ and $\mathbb{E}\left[\|\mathbf{P}\|^2\right]$:

$$\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] = \sum_{i=1}^{p} \mathbf{P}_{si}^2 \mathbb{E}\left[\mathbf{P}_i^2\right] + \sum_{1 \leq i \neq j \leq p} \mathbf{P}_{si}^2 \mathbb{E}\left[\mathbf{P}_j^2\right], \tag{C3}$$

$$\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] = \sum_{i=1}^{p} \mathbf{P}_{si}^2 \left(\mathrm{Var}\left(\mathbf{P}_i\right) + \mathbb{E}\left[\mathbf{P}_i\right]^2\right) + \sum_{1 \leq i \neq j \leq p} \mathbf{P}_{si}^2 \left(\mathrm{Var}\left(\mathbf{P}_j\right) + \mathbb{E}\left[\mathbf{P}_j\right]^2\right), \tag{C4}$$

$$\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] = \sum_{i=1}^{p} \mathbf{P}_{si}^2 \mathbb{E}\left[\mathbf{P}_i\right]^2 + \sum_{1 \leq i \neq j \leq p} \mathbf{P}_{si}^2 \mathbb{E}\left[\mathbf{P}_j\right]^2 + \sum_{i=1}^{p} \mathbf{P}_{si}^2 \mathrm{Var}\left(\mathbf{P}_i\right) + \sum_{1 \leq i \neq j \leq p} \mathbf{P}_{si}^2 \mathrm{Var}\left(\mathbf{P}_j\right). \tag{C5}$$

On the other hand:

$$\mathbb{E}\left[\mathbf{P} \cdot \mathbf{P}_s\right]^2 = \sum_{i=1}^{p} \mathbf{P}_{si}^2 \mathbb{E}\left[\mathbf{P}_i\right]^2 + 2 \sum_{1 \leq i < j \leq p} \mathbf{P}_{si} \mathbb{E}\left[\mathbf{P}_i\right] \mathbf{P}_{sj} \mathbb{E}\left[\mathbf{P}_j\right]. \tag{C6}$$

Therefore:

$$\begin{aligned}
\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] - \mathbb{E}\left[\mathbf{P}\cdot\mathbf{P}_s\right]^2 &= \sum_{i=1}^{p} \mathbf{P}_{si}^2 \mathrm{Var}\left(\mathbf{P}_i\right) + \sum_{1\le i\neq j\le p} \mathbf{P}_{si}^2 \mathrm{Var}\left(\mathbf{P}_j\right) \\
&\quad + \sum_{1\le i\neq j\le p} \mathbf{P}_{si}^2 \mathbb{E}\left[\mathbf{P}_j\right]^2 \\
&\quad -2 \sum_{1\le i<j\le p} \mathbf{P}_{si}\mathbb{E}\left[\mathbf{P}_i\right]\mathbf{P}_{sj}\mathbb{E}\left[\mathbf{P}_j\right], \qquad \text{(C7)} \\
\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] - \mathbb{E}\left[\mathbf{P}\cdot\mathbf{P}_s\right]^2 &= \sum_{i=1}^{p} \mathbf{P}_{si}^2 \mathrm{Var}\left(\mathbf{P}_i\right) + \sum_{1\le i\neq j\le p} \mathbf{P}_{si}^2 \mathrm{Var}\left(\mathbf{P}_j\right) \\
&\quad + \sum_{1\le i<j\le p} \mathbf{P}_{si}^2 \mathbb{E}\left[\mathbf{P}_j\right]^2 + \mathbf{P}_{sj}^2 \mathbb{E}\left[\mathbf{P}_i\right]^2 \\
&\quad -2 \sum_{1\le i<j\le p} \mathbf{P}_{si}\mathbb{E}\left[\mathbf{P}_i\right]\mathbf{P}_{sj}\mathbb{E}\left[\mathbf{P}_j\right], \qquad \text{(C8)} \\
\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] - \mathbb{E}\left[\mathbf{P}\cdot\mathbf{P}_s\right]^2 &= \sum_{i=1}^{p} \mathbf{P}_{si}^2 \mathrm{Var}\left(\mathbf{P}_i\right) + \sum_{1\le i\neq j\le p} \mathbf{P}_{si}^2 \mathrm{Var}\left(\mathbf{P}_j\right) \\
&\quad + \sum_{1\le i<j\le p} \left(\mathbf{P}_{si}\mathbb{E}\left[\mathbf{P}_j\right] - \mathbf{P}_{sj}\mathbb{E}\left[\mathbf{P}_i\right]\right)^2 \qquad \text{(C9)}
\end{aligned}$$

As a consequence,

$$\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] - \mathbb{E}\left[\mathbf{P}\cdot\mathbf{P}_s\right]^2 \geq 0, \qquad \text{(C10)}$$

as the sum of positive terms, and $\|\mathbf{P}_s\|^2 \mathbb{E}\left[\|\mathbf{P}\|^2\right] - \mathbb{E}\left[\mathbf{P}\cdot\mathbf{P}_s\right]^2$ is equal to 0 if and only if $\mathrm{Var}\left(\mathsf{P}_i\right) = 0$ for all $i = 1, \ldots, p$ and $\left(\mathbf{P}_{si}\mathbb{E}\left[\mathbf{P}_j\right] - \mathbf{P}_{sj}\mathbb{E}\left[\mathbf{P}_i\right]\right)^2 = 0$ for all $1 \leq i < j \leq p$. In particular, this condition is equivalent to:

$$\mathbf{P}_{s1}\mathbb{E}\left[\mathbf{P}_j\right] = \mathbf{P}_{sj}\mathbb{E}\left[\mathbf{P}_1\right], \qquad j = 2, \ldots, p, \qquad \text{(C11)}$$

which means that $\mathbf{P}_{\boldsymbol{s}}$ and $\mathbb{E}\left[\mathbf{P}\right]$ are collinear.

Open Research

All the data used to make this study are available from: https://zenodo.org/record/8037714 (DOI: 10.5281/zenodo.8037714)

The figures were made using:

- Matlab version 2021-b available from: https://se.mathworks.com/products/matlab.html
- matplotlib version 3.5.2 available from: https://matplotlib.org/

Acknowledgments

This study was supported by Horizon Europe (#101081555), the Research Council of Norway (Grant No. 301396) the Trond Mohn Foundation under project number: BFS2018TMT01. This work has also received a grant for computer time from the Norwegian Program for supercomputing (NOTUR2, project number NN9039K) and a storage Grant (NORSTORE, NS9039K).

References

Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, *129*(12), 2884–2903. doi: 10.1175/1520-0493(2001)129⟨2884:AEAKFF⟩2.0.CO;2

Anderson, J. L. (2007). Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D: Nonlinear Phenomena*, *230*(1-2), 99–111. doi: 10.1016/j.physd.2006.02.011

Bauer, H.-s., Schwitalla, T., Wulfmeyer, V., Bakhshaii, A., Ehret, U., Neuper, M., & Caumont, O. (2015). Quantitative precipitation estimation based on high-resolution numerical weather prediction and data assimilation with WRF – a performance test. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *0870*(67), 25047. doi: 10.3402/tellusa.v67.25047

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., ... Kristjánsson, J. E. (2013). The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate. *Geoscientific Model Development*, *6*(3), 687–720. doi: 10.5194/gmd-6-687-2013

Bethke, I., Wang, Y., Counillon, F., Keenlyside, N., Kimmritz, M., Fransner, F., ... Eldevik, T. (2021). NorCPM1 and its contribution to CMIP6 DCPP. *Geoscientific Model Development*, *14*(11), 7073–7116. doi: 10.5194/gmd-14-7073-2021

Bethke, I., Wang, Y., Counillon, F., Kimmritz, M., Langehaug, H., Bentsen, M., & Keenlyside, N. (2018). Subtropical north atlantic preconditioning key to skillful subpolar gyre prediction.

Bishop, C. H., & Hodyss, D. (2007). Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *133*, 2029–2044.

Bitz, C. M., Shell, K. M., Gent, P. R., Bailey, D. A., Danabasoglu, G., Armour, K. C., ... Kiehl, J. T. (2012). Climate sensitivity of the community climate system model, version 4. *Journal of Climate*, *25*(9), 3053–3070. doi: 10.1175/JCLI-D-11-00290.1

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., ... others (2016). The decadal climate prediction project (dcpp) contribution to cmip6. *Geoscientific Model Development*, *9*(10), 3751–3777.

Brune, S., Nerger, L., & Baehr, J. (2015). Assimilation of oceanic observations in a global coupled earth system model with the seik filter. *Ocean Modelling*, *96*, 254–264.

Cardinali, C., Pezzulli, S., & Andersson, E. (2004). Influence-matrix diagnostic of a data assimilation system. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *130*(603), 2767–2786.

Carrassi, A., Bocquet, M., Bertino, L., & Evensen, G. (2018). Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, *9*(5), 1–50. doi: 10.1002/wcc.535

Carrió, D. S., Bishop, C. H., & Kotsuki, S. (2021). Empirical determination of the covariance of forecast errors: An empirical justification and reformulation of hybrid covariance models. *Quarterly Journal of the Royal Meteorological Society*, *147*(736), 2033–2052. doi: 10.1002/qj.4008

Counillon, F., & Bertino, L. (2009). Ensemble optimal interpolation: multivariate properties in the gulf of mexico. *Tellus A: Dynamic Meteorology and Oceanography*, *61*(2), 296–308.

Counillon, F., Bethke, I., Keenlyside, N., Bentsen, M., Bertino, I., & Zheng, F. (2014). Seasonal-to-decadal predictions with the ensemble kalman filter and the Norwegian earth System Model: A twin experiment. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *66*(1). doi: 10.3402/tellusa.v66.21074

Counillon, F., Keenlyside, N., Bethke, I., Wang, Y., Billeau, S., Shen, M. L., &

Bentsen, M. (2016). Flow-dependent assimilation of sea surface temperature in isopycnal coordinates with the Norwegian Climate Prediction Model. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *68*(1). doi: 10.3402/tellusa.v68.32437

Counillon, F., Keenlyside, N., Toniazzo, T., Koseki, S., Demissie, T., Bethke, I., & Wang, Y. (2021). Relating model bias and prediction skill in the equatorial atlantic. *Climate Dynamics*, *56*, 2617–2630.

Counillon, F., Sakov, P., & Bertino, L. (2009). Application of a hybrid EnKF-OI to ocean forecasting. *Ocean Science*, *5*(4), 389–401. doi: 10.5194/os-5-389-2009

Craig, A. P., Vertenstein, M., & Jacob, R. (2012). A new flexible coupler for earth system modeling developed for CCSM4 and CESM1. *International Journal of High Performance Computing Applications*, *26*(1), 31–42. doi: 10.1177/1094342011428141

Evensen, G. (2003). The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*(4), 343–367. doi: 10.1007/s10236-003-0036-9

Frankcombe, L. M., McGregor, S., & England, M. H. (2015). Robustness of the modes of indo-pacific sea level variability. *Climate dynamics*, *45*, 1281–1298.

Gharamti, M. E. (2020). Hybrid ensemble-variational filter: A spatially and temporally varying adaptive algorithm to estimate relative weighting. *Monthly Weather Review*, *149*(1), 65–76. doi: 10.1175/MWR-D-20-0101.1

Gharamti, M. E., Valstar, J., & Hoteit, I. (2014). An adaptive hybrid EnKF-OI scheme for efficient state-parameter estimation of reactive contaminant transport models. *Advances in Water Resources*, *71*, 1–15. Retrieved from `http://dx.doi.org/10.1016/j.advwatres.2014.05.001` doi: 10.1016/j.advwatres.2014.05.001

Guemas, V., Corti, S., García-Serrano, J., Doblas-Reyes, F., Balmaseda, M., & Magnusson, L. (2013). The indian ocean: The region of highest skill worldwide in decadal climate prediction. *Journal of Climate*, *26*(3), 726–739.

Hamill, T. M., & Snyder, C. (2000). A Hybrid Ensemble Kalman Filter–3D Variational Analysis Scheme. *Monthly Weather Review*, *128*(8), 2905–2919. Retrieved from `http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282000%29128%3C2905%3AAHEKFV%3E2.0.CO%3B2` doi: 10.1175/1520-0493(2000)128⟨2905:AHEKFV⟩2.0.CO;2

Houtekamer, P. L., & Mitchell, H. L. (2001). A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, *129*, 123–137.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., ... Marshal, S. (2013). The community earth system model: a framework for collaborative research. *Amercian Meteorological Society*(September), 1339–1360. doi: 10.1175/BAMS-D-12-00121.1

Karspeck, A. R., Danabasoglu, G., Anderson, J., Karol, S., Collins, N., Vertenstein, M., ... others (2018). A global coupled ensemble data assimilation system using the community earth system model and the data assimilation research testbed. *Quarterly Journal of the Royal Meteorological Society*, *144*(717), 2404–2430.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., ... Slater, A. G. (2011). Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, *3*, 1–27. doi: 10.1029/2011ms000045

Ménétrier, B. (2021). *Sample covariance filtering* (Tech. Rep.). Retrieved from `https://github.com/benjaminmenetrier/covariance_filtering/blob/8b2fa7366943cff8de40fa858e0a0a153b63eb90/covariance_filtering.pdf`

Ménétrier, B., & Auligné, T. (2015). Optimized Localization and Hybridization to Filter Ensemble-Based Covariances. *Monthly Weather Review*, *143*(10), 3931–

3947. doi: 10.1175/mwr-d-15-0057.1

Ménétrier, B., Montmerle, T., Michel, Y., & Berre, L. (2015). Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part I: Optimality Criteria and Application to Variance Filtering and Covariance Localization. *Monthly Weather Review*, *143*(5), 1622–1643. Retrieved from `http://journals.ametsoc.org/doi/10.1175/MWR-D-14-00157.1` doi: 10.1175/MWR-D-14-00157.1

Oke, P. R., Brassington, G. B., Griffin, D. A., & Schiller, A. (2008). The bluelink ocean data assimilation system (bodas). *Ocean modelling*, *21*(1-2), 46–70.

Oke, P. R., Sakov, P., & Corney, S. P. (2007). Impacts of localisation in the EnKF and EnOI: Experiments with a small model. *Ocean Dynamics*, *57*(1), 32–45. doi: 10.1007/s10236-006-0088-8

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., ... Yorke, J. A. (2004). A local ensemble Kalman filter for atmospheric data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, *56*(5), 415–428. doi: 10.3402/tellusa.v56i5.14462

O'Kane, T. J., Sandery, P. A., Monselesan, D. P., Sakov, P., Chamberlain, M. A., Matear, R. J., ... Stevens, L. (2019). Coupled data assimilation and ensemble initialization with application to multiyear enso prediction. *Journal of Climate*, *32*(4), 997–1024.

Palmer, T., & Stevens, B. (2019). The scientific challenge of understanding and estimating climate change. *Proceedings of the National Academy of Sciences*, *116*(49), 24390–24395.

Raanes, P. N., Bocquet, M., & Carrassi, A. (2019). Adaptive covariance inflation in the ensemble Kalman filter by Gaussian scale mixtures. *Quarterly Journal of the Royal Meteorological Society*, *145*(January 2018), 53–75. doi: 10.1002/qj.3386

Raboudi, N. F., Ait-El-Fquih, B., Dawson, C., & Hoteit, I. (2019). Combining Hybrid and One-Step-Ahead Smoothing for Efficient Short-Range Storm Surge Forecasting with an Ensemble Kalman Filter. *Monthly Weather Review*, 3283–3300. doi: 10.1175/MWR-D-18-0410.1

Rainwater, S., & Hunt, B. (2013). Mixed-Resolution Ensemble Data Assimilation. *Monthly Weather Review*, *141*(9), 3007–3021. doi: 10.1175/mwr-d-12-00234.1

Sakov, P., Counillon, F., Bertino, L., Lister, K. A., Oke, P. R., & Korablev, A. (2012). TOPAZ4: An ocean-sea ice data assimilation system for the North Atlantic and Arctic. *Ocean Science*, *8*(4), 633–656. doi: 10.5194/os-8-633-2012

Sakov, P., Evensen, G., & Bertino, L. (2010). Asynchronous data assimilation with the enkf. *Tellus A: Dynamic Meteorology and Oceanography*, *62*(1), 24–29.

Sakov, P., & Oke, P. R. (2008). A deterministic formulation of the ensemble Kalman filter: An alternative to ensemble square root filters. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *60 A*(2), 361–371. doi: 10.1111/j.1600-0870.2007.00299.x

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. doi: 10.1175/BAMS-D-11-00094.1

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., ... Rose, S. K. (2011). The representative concentration pathways: An overview. *Climatic Change*, *109*(1), 5–31. doi: 10.1007/s10584-011-0148-z

Wahba, G., Johnson, D. R., Gao, F., & Gong, J. (1995). Adaptive tuning of numerical weather prediction models: Randomized gcv in three-and four-dimensional data assimilation. *Monthly Weather Review*, *123*(11), 3358–3370.

Wang, X., Hamill, T. M., Whitaker, J. S., & Bishop, C. H. (2007). A Comparison of Hybrid Ensemble Transform Kalman Filter – Optimum Interpolation and Ensemble Square Root Filter Analysis Schemes. *Monthly Weather Review*, 1055–1076. doi: 10.1175/MWR3307.1

Wang, Y., Counillon, F., Barthélémy, S., & Barth, A. (2022). Benefit of vertical localization for sea surface temperature assimilation in isopycnal coordinate model. *Frontiers in Climate*, *4*.

Wang, Y., Counillon, F., & Bertino, L. (2016). Alleviating the bias induced by the linear analysis update with an isopycnal ocean model. *Quarterly Journal of the Royal Meteorological Society*, *142*(695), 1064–1074.

Weaver, A., & Courtier, P. (2001). Generalized Diffusion. *Quarterly Journal of the Royal Meteorological Society*, *127*, 1815–1846.

Whitaker, J. S., & Hamill, T. M. (2012). Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. *Monthly Weather Review*, *140*(9), 3078–3089. doi: 10.1175/mwr-d-11-00276.1

Xie, J., Counillon, F., & Bertino, L. (2018). Impact of assimilating a merged sea-ice thickness from cryosat-2 and smos in the arctic reanalysis. *The Cryosphere*, *12*(11), 3671–3691.

Zhang, S., Rosati, A., & Harrison, M. (2009). Detection of multidecadal oceanic variability by ocean data assimilation in the context of a "perfect" coupled model. *Journal of Geophysical Research: Oceans*, *114*(C12).