Article: Correcting circulation biases in a lower-resolution global general circulation model with data assimilation.

Martin Canter, Alexander Barth, and Jean-Marie Beckers

GHER, University of Liège, Belgium

December 8, 2016

1 Abstract

In this study, we aim at developing a new method of bias correction using data assimilation. This method is based on the stochastic forcing of a model to correct bias by directly including an additional source term into the model equations. This method is first presented and tested with a twin experiment on the fully controlled Lorenz '96 model. It is then applied to the lower-resolution global circulation NEMO-LIM2 model, with both a twin experiment and a realistic case experiment. Sea surface height observations are used to estimate a forcing aimed at correcting the poorly located currents. Validation is then performed through the use of other variables such as sea surface temperature and salinity. Results show that the method is able to consistently correct a part of the model bias for the twin experiment, and shows the encountered difficulties for the realistic experiment. The bias correction term is presented and is consistent with the limitations of the global circulation model causing bias on the oceanic currents.

2 Introduction

Bias is commonly defined as a systematic error with a non-zero mean. Whether it originates from the model itself, from the observations, or from the assimilation scheme, the effects of bias can significantly deteriorate the solution of the model. In numerical modelling, a current limitation arises from the finite computational power available, which, in ocean models, results in limited resolution. This causes poorly resolved vertical mixing and poor specification of atmospheric fluxes to be a leading term for bias (Gerbig et al., 2008). Our limited knowledge of the system also leads to the imperfect specification of boundary conditions, and a poor representation of subgrid physical processes (Baek et al., 2009). Those differences between the numerical model solution and the dynamics of the real ocean induce systematic errors in the numerical forecasts. When used for prediction or long-term simulations with a limited number of available observations, those systematic errors cause the model to exhibit significant differences in climatologies when compared to the reality. In some circumstances, they can even be comparable or larger than the non-systematic error of the solution of the model. While the random part of the model error has been reduced thanks to several advances in numerical modelling, it has become increasingly necessary to address the systematic model error (Keppenne et al., 2005). Bias in climate modelling can be so large that only variations and anomalies are studied, rather than the absolute results of the model (Zunz et al., 2013).

To reduce the error of the model, observations can be taken into account to correct the model state by using data assimilation. However, a critical assumption for data assimilation analysis schemes is that the mean of the background error is zero. This hypothesis is by definition violated in the presence of bias. Data assimilation schemes that are designed to use non-biased observations to correct random errors with zero mean in a model background estimate are called bias-blind. In the presence of bias, those analysis schemes are suboptimal and can generate spurious corrections and undesired trends in the analysis (Dee, 2005). Most data assimilation schemes are designed to handle small, random errors and make small adjustments to the background fields which are consistent with the spatial structure of random errors (Dee, 2005). Bias-aware data assimilation schemes are designed to simultaneously estimate the model state variables and parameters that are set to represent systematic errors in the system. However, assumptions need to be made about the error covariance of the bias and its attribution to a particular source. It also needs to be represented and expressed in a set of well-defined parameters.

Model-bias estimation was first introduced by Friedland (1969), and more deeply described by Jazwinski (1970); Gelb (1974). Friedland suggested a scheme in which the model state vector should be augmented with a decoupled bias component that can be isolated from the other state vector variables. This allows the estimation of the bias prior to the estimation of the model.

The most known and referred to algorithm for online bias estimation and correction in sequential data assimilation was introduced in Dee and Da Silva (1998). Bias is estimated during the assimilation by adding an extra and separated assimilation step. It was successfully applied in Dee and Todling (2000) to the global assimilation of humidity observations in the Goddard Earth Observing System. A simplified version of this algorithm using a single assimilation step (where Dee and Da Silva (1998) needed two) was applied by Radakovich et al. (2001) to land-surface temperature assimilation, and by Bell et al. (2004) for the online estimation of subsurface temperature bias in tropical oceans. It was also used for model bias estimation by Baek et al. (2006), and observation-bias correction in Fertig et al. (2009). Other examples are Carton et al. (2000); Keppenne et al. (2005); Chepurin et al. (2005); Nerger and Gregg (2008).

Bias-correction approaches can be classified as follows (Keppenne et al., 2005; Chepurin et al., 2005). In offline methods, bias is estimated from the model mean and the climatology, using a preliminary model run. Offline methods are simple to implement and have a small computational cost. In online methods, the bias is updated during the data assimilation step, resulting in an analysed bias.

However, most methods of bias correction need a reference dataset which is defined as bias free, from which a bias estimation can be provided. In practice, it can be difficult to find such a dataset. The bias also needs to be characterised in terms of some well-defined set of parameters. While this is obvious for bias estimation, it is a critical condition when attempting bias correction. The attribution of a bias to an incorrect (unbiased) source will force the assimilation to be consistent with the now biased source. In some cases, the bias correction would even deteriorate the assimilation procedure, and perform worse than a classic, bias-blind assimilation (Dee, 2004).

The effect of bias on the model climatology can not be neglected. The necessity of removing, or at least, reducing the effects of bias on the model has driven to the development of methods allowing to force the model towards a non-biased climatology. Addressing systematic model errors, such as oceanographic biases, is even more tricky, since a representation of the bias itself, or the generation mechanism, is needed. The bias in the background field can be directly modelled by assuming some kind of time behaviour such as persistence (Dee, 2005; Chepurin et al., 2005). As background errors are observable, it is relatively straightforward to formulate a consistent bias-estimation scheme. Suppressing the bias generation during the integration of the model rather than correcting it afterwards would however be preferable.

For example in Derber and Rosati (1989), a variational continuous assimilation technique is applied under the form of a modification of the adjoint technique. A correction term then is added to the equations. The technique aimed at optimally fitting the data throughout the assimilation period, rather than relaxing the solution towards the values at observation times. It has been applied to radiative transfer model in Derber and Wu (1998).

Another example was discussed by Radakovich et al. (2004), where the model is so heavily affected by bias that a classic bias-aware assimilation scheme (Dee and Da Silva, 1998) is insufficient. The bias-correction term is only applied during the assimilation, but due to the model characteristics, the model solution quickly slips back to its biased state and dissipates the correction term. In that study, an adapted bias correction term was applied during the model run which was proportional to the initial term and the time separating two analysis steps.

In the present work, the problem of model-bias correction is tackled by developing a new method, which combines stochastic forcing and data assimilation. Data assimilation is used here to estimate, create and define analysed stochastic forcing terms from which a deterministic forcing term (estimated by the the ensemble mean) is used to reduce the model bias.

Most of the previously developed and existing methods correct bias in the model results and leave its source uncorrected. Some studies have however tackled the bias-correction problem directly, such as in Leeuwenburgh (2008), where an estimation and correction of a surface wind-stress bias was performed through the modification of the bias scheme of Dee and Da Silva (1998) with an Ensemble Kalman filter modification.

The objective of this paper is to correct the effects of the bias by applying a stochastic forcing into the model equation, where the bias is supposed to be generated. An Ensemble Transform Kalman Filter (ETKF) is used to find an optimal forcing term which is directly injected into the modified model equations. The aim is to provide a continuous bias correction by forcing the model towards a non-biased climatology.

The forcing term introduced here does not yet exist in the model equations and the method is only partly similar to a classic parameter estimation problem (Annan et al., 2005; Massonnet et al., 2014). Indeed, we do not aim at optimising an already existing forcing term as in Broquet et al. (2011) (where a weak constraint ocean 4DVAR scheme is used to correct ocean surface forcing), but rather add a new term which itself is optimised. Moreover, as the forcing term optimisation covers the whole time period, our method differs by the fact that it can be considered as an ensemble smoother.

This paper is divided into the following sections: In section 3, the method principle is presented and detailed. In section 4, the Lorenz '96 model is studied with a particular point of view related to the model mean and its global behaviour. In section 5, this novel approach is then tested and implemented with a classic twin experiment on the Lorenz '96 model (Lorenz, 1996; Lorenz and Emanuel, 1998). The efficiency and results of this method are presented. In section 6, this new method is then applied and tested on the realistic sea-ice NEMO-LIM2 ocean model. Again, it is first tested with a twin experiment to control the behaviour of the model. It is afterwards tested with real observations from the mean dynamic topography (MDT) of the CNES (centre national d'études spatiales) (Rio et al., 2011). Section 7 closes this work with a discussion of results and possible extensions of this work.

3 Method

This work aims at developing a new method of bias correction for numerical modelling using data assimilation. While most previously developed and existing methods correct bias in the model results, our objective is to estimate a deterministic bias-correction forcing term from a set of model runs with a stochastic forcing applied to the model equation.

Consider the following nonlinear stochastic discrete-time dynamical system

$$\mathbf{x}^{(m)} = \mathcal{M}_{(m)}\left(\mathbf{x}^{(m-1)}\right), \tag{1}$$

where $m = 1, ..., m_{max}$ is the time index, $\mathbf{x}^{(m)}$ the n dimensional model state and $\mathcal{M}_{(m)}$ the forward model operator. The real dynamical system is described as follow, were we assume the additive model error presented in Evensen (2007)

$$\mathbf{x}^{t(m)} = \mathcal{M}^{t}_{(m)}\left(\mathbf{x}^{t(m-1)}\right) + \boldsymbol{\beta}^{(m)}. \tag{2}$$

Here, $\mathbf{x}^{t(m)}$ is the n dimensional true state, $\mathcal{M}^{t}_{(m)}$ the true model forward operator, and $\boldsymbol{\beta}^{(m)}$ the stochastic error. This model error can be split into two parts, namely a random part whose average is zero: $< \widetilde{\boldsymbol{\beta}}^{(m)} > = 0$, and a systematic error, or bias: $\mathbf{b}$ (Dee, 2005). One can write that

$$\boldsymbol{\beta}^{(m)} = \widetilde{\boldsymbol{\beta}}^{(m)} + \mathbf{b}. \tag{3}$$

Note that we consider the bias to be constant in time. If necessary, this assumption can be relaxed to handle time-varying bias such as seasonal biases. Although finding an adequate correction would prove more difficult and computationally more costly, the principle of the method would remain identical. It is not, however, the objective of this paper and we assume the bias to be constant.

We aim here at handling the bias using an ensemble smoother. To do so, an ensemble of N model trajectories is defined following van Leeuwen (2001); Hunt et al. (2004), with $i = 1, \ldots, N$, as

$$\mathbf{x}_i = \begin{bmatrix} \mathbf{x}_i^{(1)} \\ \mathbf{x}_i^{(2)} \\ \vdots \\ \mathbf{x}_i^{(m_{\max})} \end{bmatrix}. \tag{4}$$

A clear difference is made here between the bias to be corrected $\mathbf{b}$, and the estimator of the bias-correction term $\widehat{\mathbf{b}}_i$, which can be seen as a parameter to be estimated (Barth et al., 2010; Sakov et al., 2010). The state vector is augmented with an estimator of the bias correction term $\widehat{\mathbf{b}}_i$ and one obtains

$$\mathbf{x'}_i = \begin{bmatrix} \mathbf{x}_i^{(1)} \\ \mathbf{x}_i^{(2)} \\ \vdots \\ \mathbf{x}_i^{(m_{\max})} \\ \widehat{\mathbf{b}}_i \end{bmatrix}. \tag{5}$$

One can then write the update of the state vector after the analysis with the Ensemble Transform Kalman filter (Bishop et al., 2001) as

$$\mathbf{x'}^a = \mathbf{x'}^f + \mathbf{K'}\left(\mathbf{y}^o - \mathbf{H'}\mathbf{x'}^f\right), \tag{6}$$

where

$$\mathbf{x'}^a = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x'}^a_i \quad , \quad \mathbf{x'}^f = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x'}^f_i, \tag{7}$$

$$\mathbf{K'} = \mathbf{P'}^f\mathbf{H'}^T\left(\mathbf{H'}\mathbf{P'}^f\mathbf{H'}^T + \mathbf{R}\right)^{-1}. \tag{8}$$

Here, $\mathbf{y}^o$ is the mean state of the observations. Hereafter, the absence of ensemble index i in the equation will refer to the use of the ensemble mean. The observation operator $\mathbf{H'}$ applied to the trajectory $\mathbf{x'}$ also includes a time average and an extraction operator $\mathbf{H}$ of the observed part of the model state

$$\mathbf{H'}\mathbf{x'} = \sum_{m=1}^{m_{max}}\mathbf{H}\mathbf{x}^{(m)} = \mathbf{H}\overline{\mathbf{x}}, \tag{9}$$

$$\overline{\mathbf{x}} = \frac{1}{m_{max}}\sum_{m=1}^{m_{max}}\mathbf{x}^{(m)}, \tag{10}$$

where $\overline{\mathbf{x}}$ is the time average of the model state vector. Since we are only interested in the climatology of the model and the estimator of the bias correction term, the complete model trajectory is not needed. The average state of the model is sufficient, and it is computationally much more interesting to only deal with the latter: to do so, one uses a state vector consisting only of the model mean state and the estimator of the bias correction term

$$\mathbf{x}'' = \begin{bmatrix} \overline{\mathbf{x}} \\ \widehat{\mathbf{b}} \end{bmatrix}, \tag{11}$$

and an observation operator defined as

$$\mathbf{H}''\mathbf{x}'' = \mathbf{H}\overline{\mathbf{x}}. \tag{12}$$

One can show that the analysis using the average model state (Eq. (13)) provides the same analysed bias-estimator correction term $\widehat{\mathbf{b}^a}$ as when the full trajectory is included in the estimation vector (Eq. (6)), which is written as

$$\mathbf{x}''^a = \mathbf{x}''^f + \mathbf{K}'' \left(\mathbf{y}^o - \mathbf{H}''\mathbf{x}''^f\right). \tag{13}$$

The mathematical demonstration of this property is given in the appendix. In practice, the assimilation of observations of the climatology of the model $\overline{\mathbf{x}}$ allows the update and optimisation of the bias-estimator correction $\widehat{\mathbf{b}^a}$ through the Kalman filter/smoother equations. The model is then rerun with the optimal bias-correction term, providing us with a bias-corrected model trajectory $\mathbf{x}^{r(m)}$, expressed as

$$\mathbf{x}^{r(m)} = \mathcal{M}_{(m)} \left(\mathbf{x}^{r(m-1)}\right) - \widehat{\mathbf{b}^a}. \tag{14}$$

The interest of this method is that when the model is rerun, it provides a new model trajectory $\mathbf{x}^{r(m)}$. This new trajectory, hence its average $\overline{\mathbf{x}^r}$, is different from the analysis $\mathbf{x}''^a$. Indeed, the former results from a new run fully governed by the corrected equations of the model (Eq. (14)), whereas the latter results directly from the analysis (Eq. (13)). If the model was completely linear, the analysis provided by the ETKF scheme would be equal to the model bias-corrected run.

To summarise, it is common for bias-correction schemes to estimate the bias during the model run (be it online or offline) using a dynamic model for the bias. This is different from the present approach optimizing the bias-correction term. Also, since the bias estimation with the analysis uses all available information, one can consider this method as a smoother which provides us with a bias-correction term $\widehat{\mathbf{b}^a}$ aimed at modifying the model. This can be used to run a corrected model, either in forecast or reanalysis mode. A schematic view of the method is shown on Fig. 1.

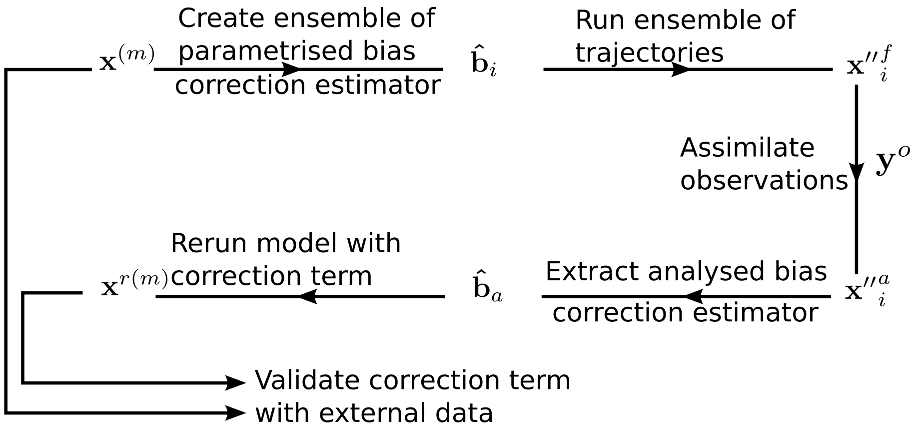


Figure 1: Schematic of the method.

In the next sections of this paper, the reference run (also called true run) corresponds to $\mathbf{x}^t$ (Eq. (2)), and the free run to $\mathbf{x}^m$ (Eq. (1)). The ensemble before analysis, created with an ensemble of guessed estimators $\widehat{\mathbf{b}}_i$, is noted $\mathbf{x}''^{f}_{i}$. The analysed ensemble, after assimilating the observations $\mathbf{y}^o$, is noted $\mathbf{x}''^{a}_{i}$ (Eq. (13)). Finally, the corrected run or rerun will correspond to $\mathbf{x}^{r(m)}$ (Eq. (14)) with the bias correction $\widehat{\mathbf{b}^a}$ provided by the analysis (Eq. (13)).

4 Lorenz '96 Model

We first test our approach on a fully controlled mathematical model. In 1963, Edward Lorenz developed a simplified mathematical model aimed at reproducing atmospheric convection. It is notable for having chaotic solutions for certain parameter values and initial conditions (Lorenz, 1963). Originally, it consisted of a system of three differential equations. In 1996, it was updated in its 40-variables form, known as the Lorenz '96 model (Lorenz, 1996; Lorenz and Emanuel, 1998). It models a circular closed boundaries system with advection and diffusion properties. The system is described by

$$\frac{dX_k}{dt} = -X_{k-2}X_{k-1} + X_{k-1}X_{k+1} - X_k + \mathbf{F}_k, \tag{15}$$

where we slightly modify the original version by taking a spatially changing forcing parameter $\mathbf{F}_k$ instead of a constant one for all the variables.

This model has been widely used to test and improve data-assimilation methods, ensemble filters or parameter estimation (Li et al., 2009; Anderson, 2009; van Leeuwen, 2010). Indeed, developing new methodologies relies on multiple specific procedures which need to be tested. This preparation work is better done beforehand on a very small model which, even if it does not stand comparison with the complexity of realistic models, still enables us to address the multiple issues we will be facing later on. Also, even if the Lorenz '96 model is not particularly complex, it still shows similarities with the ocean, in particular, its chaotic behaviour makes forecasting a real issue.

We will use this model in a different way than previous studies. The latter focused generally on the value of each variable during the model run. Since our aim is not to correct the specific value of the variables, but rather correct the bias that affects those variables, we will look instead at the mean value of those variables over a period of time. This choice is motivated by the fact that, in some sense, bias is defined as a systematic error over a period of time.

Therefore, we first look at the general behaviour of the model when launched with a set of different initial conditions and different $\mathbf{F}_k$ values. It is interesting to note that, even though the model does show a chaotic behaviour which highly depends on the initial conditions and the $\mathbf{F}_k$ values, the model mean tends to stabilise itself after a certain amount of time. Lorenz and Emanuel (1998) already noted that if $\overline{\mathbf{F}} < 4$, the waves can extract energy fast enough to offset the effect of the external forcing. When $\overline{\mathbf{F}} > 4$, the model becomes completely chaotic over time and shows spatially irregular patterns. Even more, when $\overline{\mathbf{F}} > 15$, the model becomes totally unstable and diverges.

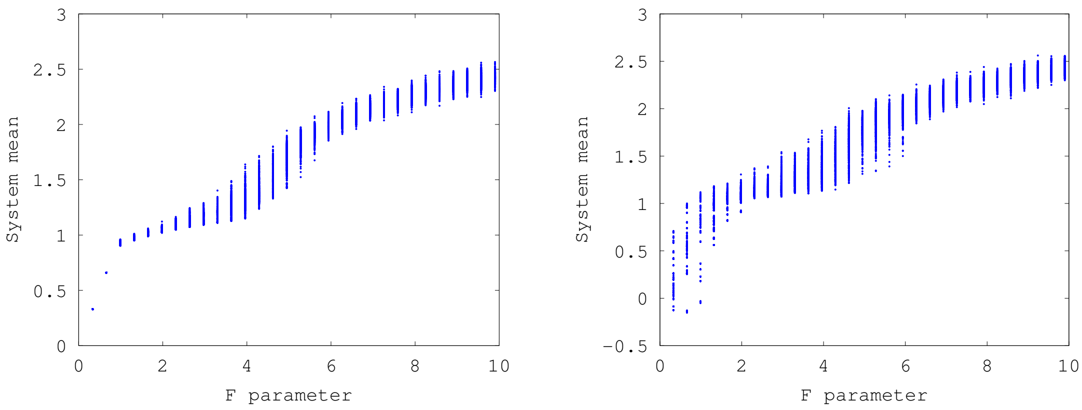
We look at the mean value of the model variables over a certain period of time. We note that there is a significant relationship between the variables' mean over time and the forcing parameter $\mathbf{F}_k$. Parameters are set to $k = 1, .., 40$ (index covering space), and a time step of 0.05, which corresponds to about 6 hours in the atmosphere (Lorenz and Emanuel, 1998). 30 evenly distributed values are chosen for $0 < \mathbf{F}_k < 10$. The model is then run with 450 different initial conditions for each $\mathbf{F}_k$, over 1000 time steps. The 200 first time steps are sufficient for the model to stabilise itself. The mean of the model variables is taken for the last 800 time steps and averaged over the 40 variables to obtain the model mean state.

Two cases are studied: in the first, the $\mathbf{F}_k$ are constant relatively to k for all the variables: $\mathbf{F}_k = \overline{\mathbf{F}}$ (Fig. 2a). In the second, we add a random, spatially-correlated noise on the forcing parameter in order to obtain a different $\mathbf{F}_k$ for each k (Fig. 2b). That new forcing parameter is described by

$$\mathbf{F}_k = \overline{\mathbf{F}} + \mathbf{S_P z}_k, \tag{16}$$

$$P_{i,j} = 0.3e^{\frac{-(i-j)^2}{15}}. \tag{17}$$

Here, $\mathbf{S_P}$ is the Cholesky decomposition of the covariance matrix $\mathbf{P}$ ($\mathbf{P} = \mathbf{S_P}\mathbf{S_P}^T$), and $\mathbf{z}_k$ is a random vector of 40 variables with a normal distribution $\mathbf{z}_k \sim \mathcal{N}(0, \mathbf{I})$.



(a) (b)

Figure 2: Lorenz '96 model mean state as a function of a constant forcing parameter $\overline{\mathbf{F}}$ (Fig. 2a), and as a function of the average of the spatially variable forcing parameter $\mathbf{F}_k$ as defined by Eq. (16) (Fig. 2b). The X-axis represents the 30 different $0 < \overline{\mathbf{F}} < 10$ tested. For Fig. 2b, only the mean part corresponding to $\overline{\mathbf{F}}$ is plotted for more readability. The Y-axis represents the model mean state for the 450 initial conditions as a function of $\overline{\mathbf{F}}$

We can clearly see from Fig. 2a and 2b that there is a monotonic relationship between the system mean and the forcing parameter, whether the latter is constant or not. This encourages the working hypothesis that even a fully non-linear system in each of its variable can be expected to show a simple global behaviour, as long as the system does not include a regime shift. This also confirms that even though the model state at a specific point in time depends on the initial conditions, the time average of the model over the last 800 time steps only has a minimal dependence on the initial conditions. This is important since our aim is not to predict the exact value of the system at a given point in time. We only aim at correcting the model forcing parameter and the bias it causes on the model mean state.

5 Lorenz '96 Model twin experiment

We test our method with a Lorenz '96 model twin experiment. As shown before, the forcing parameter $\mathbf{F}_k$ can be considered to be directly linked to the model mean over a period of time. First, a random, but spatially correlated $\mathbf{F}_k^t$ parameter is created following Eq. (16), with a mean $\overline{\mathbf{F}^t} = 4$. The model is then run once over $m_{max} = 1000$ time steps, with $l_{max} = 15$ different initial conditions. It is then averaged over the initial conditions and over time while ignoring the first 200 time steps to avoid the initial conditions to strongly influence the model mean. This provides the reference (or true) solution $\mathbf{X}_k^t$, obtained from the full model trajectory ${\mathbf{X}'}_{k,l,m}^t$ as follow:

$$\mathbf{X}_k^t = \frac{1}{l_{max}} \sum_{l=1}^{l_{max}} \frac{1}{m_{max}} \sum_{m=200}^{m_{max}} {\mathbf{X}'}_{k,l,m}^t. \tag{18}$$

We follow the exact same procedure to generate an ensemble of $i_{max} = 100$ different $\mathbf{F}_{k,i}^f$. Each one is also run over 1000 time steps, with 15 initial conditions, and averaged without the first 200 time steps, producing an ensemble of model solutions noted $\mathbf{X}_{k,i}^f$.

In the context of a classic twin experiment, we want to assimilate observations $\mathbf{y}_k^o$ from the reference run mean $\mathbf{X}_k^t$. In order to reproduce the behaviour and difficulties of a realistic experiment, noise is added to the reference run mean $\mathbf{X}_k^t$ and observations are created following

$$\mathbf{y}_k^o = \mathbf{X}_k^t + \beta s_{\mathbf{X}_k^t} \mathbf{z}_k. \tag{19}$$

Here $\mathbf{z}_{(k)} \sim \mathcal{N}(0, \mathbf{I})$ is a random vector, $s_{\mathbf{X}_k^t}$ is the standard deviation of $\mathbf{X}_k^t$, and $\beta = 0.1$. An Ensemble Transform Kalman Filter (ETKF) analysis scheme is then used (Bishop et al., 2001; Hunt et al., 2007), where $\mathbf{x}^f$ is the model forecast with error covariance $\mathbf{P}^f$, $\mathbf{K}$ the Kalman gain, $\mathbf{y}^o$ the observations with error covariance $\mathbf{R}$. The best linear unbiased estimator (BLUE) is then given by $\mathbf{x}^a$. The scheme is described by

$$\begin{aligned} \mathbf{x}^a &= \mathbf{x}^f + \mathbf{K}\left(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f\right), && (20) \\ \mathbf{K} &= \mathbf{P}^f\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}\right)^{-1}, && (21) \\ \mathbf{P}^a &= \mathbf{P}^f - \mathbf{K}\mathbf{H}\mathbf{P}^f, && (22) \end{aligned}$$

where $\mathbf{H}$ is the observation operator extracting the observed part of the state vector, and $\mathbf{P}^a$ is the error covariance of the model analysis $\mathbf{x}^a$. We can rewrite and express $\mathbf{P}^a = \mathbf{S}^a{\mathbf{S}^a}^T$ in terms of square-root matrices, which is possible with the following eigenvalue decomposition

$$(\mathbf{H}\mathbf{S}^f)^T\mathbf{R}^{-1}\mathbf{H}\mathbf{S}^f = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T. \tag{23}$$

This helps to avoid forming $\mathbf{P}^a$ explicitly, thus removing the need to handle very large matrices in real applications. Hence, $\mathbf{S}^a$ is given by

$$\mathbf{S}^a = \mathbf{S}^f \mathbf{U}(\mathbf{I} + \boldsymbol{\Lambda})^{-1/2}\mathbf{U}^T, \tag{24}$$

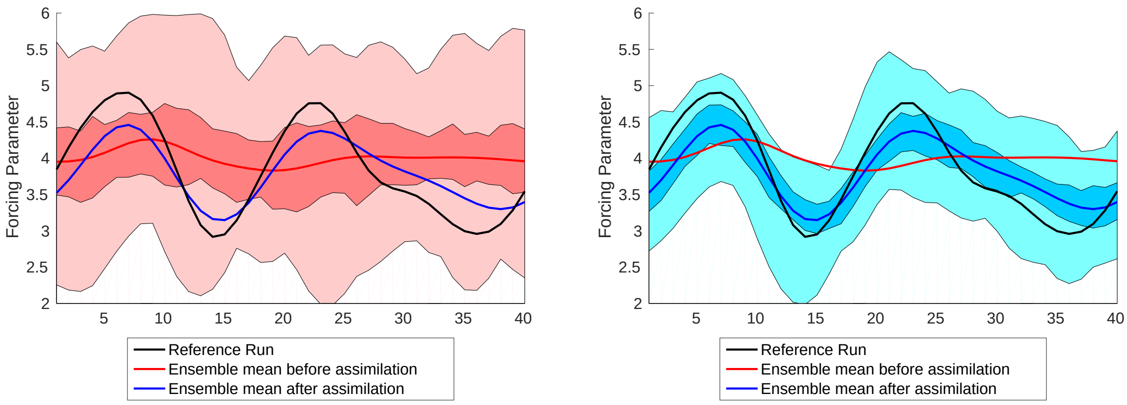
where $\boldsymbol{\Lambda}$ is diagonal and $\mathbf{U}\mathbf{U}^T = \mathbf{I}$. We then compute the Kalman gain and the model analysis with

$$\mathbf{K} = \mathbf{S}^f \mathbf{U}(\mathbf{I} + \boldsymbol{\Lambda})^{-1}\mathbf{U}^T (\mathbf{H}\mathbf{S}^f)^T \mathbf{R}^{-1}, \tag{25}$$

$$\mathbf{x}^{a(k)} = \mathbf{x}^a + \sqrt{N-1}\mathbf{S}^{a(k)}. \tag{26}$$

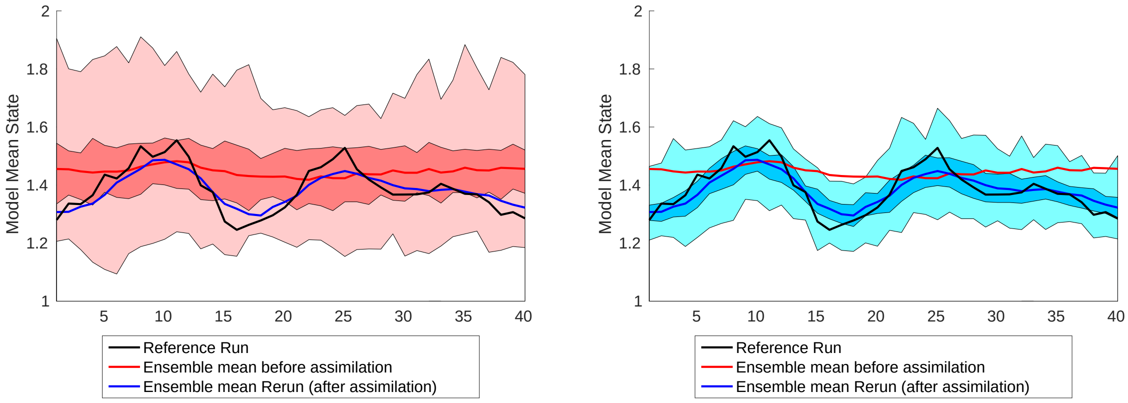
Note that no inflation factor is used for this experiment.

Using this ETKF scheme, we extend our state vector, which consists of the ensemble model mean $\mathbf{X}^f_{k,i}$, with the ensemble $\mathbf{F}^f_{k,i}$ (Eq. 11). After the analysis step, we obtain a new and updated vector of forcing parameter: $\mathbf{F}^a_{k,i}$. We then rerun the model with this updated forcings, and expect the ensemble model mean reruns $\mathbf{X}^a_{k,i}$ to improve and come closer to the reference run. The results of this procedure are shown in Fig. 3a, 3b, 4a and 4b.



(a) (b)

Figure 3: Lorenz '96 model $\mathbf{F}_k$ value (Y-axis) for each $k = 1, .., 40$ (X-axis). The reference run is shown in black: $\mathbf{F}^t_k$. The ensemble mean before assimilation, representing 100 members, is shown in red: $\mathbf{F}^f_k$. The ensemble mean after assimilation is presented in blue: $\mathbf{F}^a_k$. The light and darker areas represent then 25% and 50% percentile of the corresponding colored ensemble before assimilation (a) and after assimilation (b).



(a) (b)

Figure 4: Lorenz '96 model $\mathbf{X}_k$ model mean state (Y-axis) for each $k = 1, .., 40$ (X-axis). The reference run is shown in black: $\mathbf{X}_k^t$. The ensemble mean before assimilation, representing 100 members, is shown in red: $\mathbf{X}_k^f$. The ensemble mean after assimilation is presented in blue: $\mathbf{X}_k^a$. The light and darker red areas represent then 25% and 50% percentile of the corresponding colored ensemble before assimilation (a) and after assimilation (b).

In this experiment, the whole ensemble with assimilated forcings is used for the final run. Fig. 3a and 3b show the forcing ensemble enveloppe before ($\mathbf{F}_k^f$) and after ($\mathbf{F}_k^a$) assimilation respectively. Figures 4a and 4b show the model mean before ($\mathbf{X}_k^f$) and after ($\mathbf{X}_k^a$) assimilation respectively.

The assimilation of observations on the model mean $\mathbf{X}_k^t$ allowed the correction of the bias on $\mathbf{F}_k^f$ (Fig. 3b). The root mean square error (RMSE) on $\mathbf{F}_k^f$ before assimilation was 0.653. After the assimilation, it has been reduced to 0.323 for $\mathbf{F}_k^a$, and it is already able to reproduce the global shape of the reference run. We also need to look at the model mean (Fig. 4b). The RMSE on the ensemble mean $\mathbf{X}_k^f$ is 0.099. However, we can clearly see that the model rerun with the assimilated $\mathbf{F}_k^a$ gives much better results. The RMSE on $\mathbf{X}_k^a$ is only 0.037, and reproduces much better the shape of the observations. Thus, not only does the assimilation show an improvement on the forcing parameter of the model, but its mean climatology is also improved by effectively correcting the source of its bias.

6 NEMO-LIM2

The primitive equations model used in this study is NEMO (Nucleus for European Modelling of the Ocean, Madec (2008)), coupled to the LIM2 (Louvain-la-Neuve Sea Ice Model) sea ice model (Fichefet and Maqueda, 1997; Timmermann et al., 2005; Bouillon et al., 2009). The global ORCA2 implementation is used, which is based on an orthogonal grid with a horizontal resolution of the order of 2° and 31 z-levels (Mathiot et al., 2011; Massonnet et al., 2013). The hydrodynamic model is configured to filter free-surface gravity waves by including a damping term. The leap-frog scheme uses a time step of 1.6 hours for dynamics and tracers. The model is forced using air temperature and wind from the NCEP/NCAR reanalysis (Kalnay et al., 1996). Relative humidity, cloud cover, and precipitation are based on a monthly climatological mean. The sea surface salinity is relaxed towards climatology with a freshwater flux of -27.7 mm/day times the salinity difference in PSU.

Because of its low resolution of 2°, the NEMO-LIM2 model is subject to strong bias due to poorly located currents in the ocean. This leads to a poorly represented heat transport around the globe and causes bias on other variables in the model, such as on the sea surface height and temperature. As announced in section 3, we assume that these bias are constant in time but may have a spatial structure.

We aim here at estimating a forcing term which will correct the oceanic currents of the model. This forcing will be, in practice, a constant acceleration term directly injected into the momentum equations of the ocean-dynamics part of the model. These added constant forces on water masses will create currents correcting the model bias also for other variables. Although the term "forcing" usually refers to external forcings such as atmospheric wind stress, the forcing term here refers thus to an additional source term in the momentum equations. It does not have an external origin, but rather aims at correcting the model error such as those arising from poorly represented physical processes.

However, since the NEMO-LIM2 model is a realistic model, specific constraints need to be imposed to the forcing term in order to maintain a physical and realistic model behaviour. To create a constrained random forcing term, we use DIVA-ND, which is a Data-Interpolating Variational Analysis in N dimensions (Barth et al., 2009, 2014). This tool will allow to generate a random, spatially correlated streamfunction $\Psi(x, y)$. Meridional and zonal forcing fields for the currents can then be derived from $\Psi(x, y)$. However, this could produce currents which are perpendicular to the coasts. In order to avoid such physically impossible currents, an additional constraint is applied when generating the random field Ψ. We subject the generated streamfunction to the strong constraint $\nabla\Psi \bullet \mathbf{t} = 0$ where $\mathbf{t}$ is the vector tangent to the coast.

DIVA-ND defines a cost function $J(\Psi)$, which is expressed as

$$J(\Psi) = \int_\Omega L^4(\nabla^2\Psi)^2 + 2L^2(\nabla\Psi)^2 + \Psi dx, \tag{27}$$

where $\Psi = \Psi(x, y)$ is the random field and Ω the domain on which it is built. This cost function penalises abrupt variations over a given length-scale L, and decouples disconnected areas based on topography. The Hessian matrix of this discretized cost function is used to create random fields taking the periodicity in the model domain into account, with

$$J(\mathbf{x}_\Psi) = \mathbf{x}_\Psi^T \mathbf{P}_\Psi^{-1} \mathbf{x}_\Psi, \tag{28}$$

$$\mathbf{x}_\Psi^{(n)} = \mathbf{P}_\Psi^{1/2} \mathbf{z}_{(n)}, \tag{29}$$

$$\mathbf{z}_{(n)} \sim \mathcal{N}(0, \mathbf{I}). \tag{30}$$

Here, $\mathbf{x}_\Psi$ is the discretized random field on the model grid, $\mathbf{P}_\Psi^{-1}$ the Hessian matrix, and $\mathbf{z}_{(n)}$ a random vector with a normal distribution $\mathcal{N}(0, \mathbf{I})$. More extensive information can be found in Barth et al. (2009).

We observed that additional filtering is needed on the obtained field Ψ in order to remove very small scale signals when calculating the first derivatives of Ψ. This filtering improves the stability of the NEMO-LIM2 model when it is forced. Since we also want to create currents only in the upper layers of the ocean, but avoid modifying the global circulation in depths, the forcing is extended vertically as follow

$$\Psi(x, y, z) = \frac{\Psi(x, y)}{1 + \exp(\frac{z - T(x,y)}{L})}, \tag{31}$$

where $T(x, y)$ is defined as the yearly average ocean mixed-layer thickness. The resulting field is used as a streamfunction from which zonal and meridional divergence-free forces are derived as

$$F_u(x, y, z) = -\frac{\partial \Psi(x, y, z)'}{\partial y}, \tag{32}$$

$$F_v(x, y, z) = \frac{\partial \Psi(x, y, z)'}{\partial x}. \tag{33}$$

We can directly add this stochastic forcing terms into the momentum equations of NEMO-LIM2, where $F_u(x, y, z)$ and $F_v(x, y, z)$ are zonal and meridional components respectively. One then has

$$\frac{du}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + fv + \frac{1}{\rho}\frac{\partial \tau_x}{\partial z} + F_u, \tag{34}$$

$$\frac{dv}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial y} - fu + \frac{1}{\rho}\frac{\partial \tau_y}{\partial z} + F_v. \quad (35)$$

Eq. (34) and (35) provide a set of bias-corrected ocean-dynamics equations governing the NEMO-LIM2 model by applying a forcing term on the ocean currents while the model is running. The forcing term is a physically coherent correction that will remove some part of the bias of the model. It has been calibrated such that the variability of the sea surface height (SSH) caused by the forcing is about 28 cm, which can be compared to the root mean square error between the NEMO-LIM2 model and the CNES mean dynamic topography of 20 cm (Rio et al., 2011).

6.1 Twin Experiment

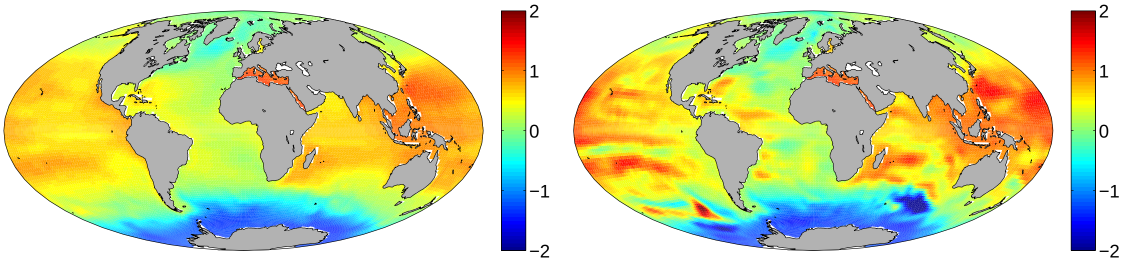
The next step to test the efficiency of our method is to apply it to the realistic ocean model NEMO-LIM2. We proceeded with a twin experiment, using a similar procedure to the one presented in the Lorenz '96 section.

First, a random forcing is generated, with a correlation length of 5000 km. It is afterwards referred to as the truth or the reference forcing. The correlation length is chosen in order to be sufficiently large enough compared to the ORCA2 grid size (about 200 km at the equator). Longer correlation length (up to 10000 km) however did not give a large enough variability in the ensemble. This reference forcing is then used with the NEMO-LIM2 model over one year.

Direct measurements of currents are too sparse. However, climatologies of the sea surface height (SSH) are available, which are inherently related to the currents in the oceans. For the realistic case (see next section), we will thus be using real SSH fields which represents time averages. Due to the geoid problem, SSH altimetry data is represented as anomalies without any information about the mean state. If one would average SSH altimetry data, one would simply obtain zero (or a quantity close to zero). The mean dynamic topography is thus derived by other means, such as drifter and gravimetric measurements. Hence, the observations already represent an average. We will thus create our observations for the twin experiment by taking the mean SSH of the reference run over one year. When we average the model SSH, the reduction in observational error due to this time averaging is already taken into account, since every ensemble member is averaged in time, causing short time-scale variability to be filtered out.

We then create an ensemble of 100 random forcings and run each of them separately. This provides us with an ensemble of yearly mean SSH. We use the Ocean Assimilation Kit (OAK) for the analysis step (Barth et al., 2015). A local assimilation scheme is used with an assimilation length equal to the correlation length of the perturbations (5000 km). The mean SSH from the reference run (Fig. 5b) are taken as the observations. Our state vector (Eq. 11) consists of our ensemble of mean SSH (Fig. 5a), and is extended with its corresponding forcings

$$\mathbf{x}'' = \begin{bmatrix} \overline{SSH} \\ \widehat{F}_u \\ \widehat{F}_v \end{bmatrix}. \tag{36}$$



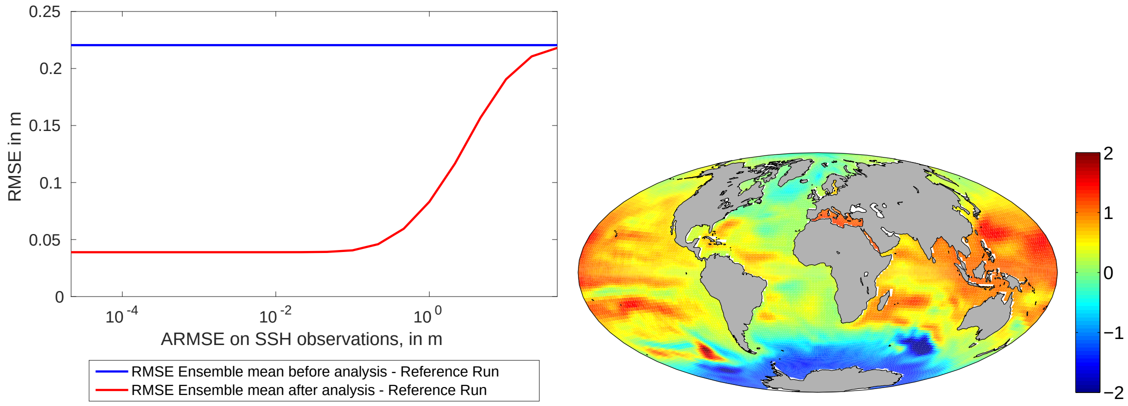
(a) (b)

Figure 5: (a) yearly mean sea surface height (SSH) of the ensemble mean runs (in m). The correlation length of the perturbation is 5000 km. (b) yearly mean sea surface height (SSH) of the twin experiment true run (in m).

Similarly to the Lorenz '96 case, we aim at finding the true forcing from the reference run. Noise is added to the observations, with a value representing 10% of the local SSH variability of the ensemble, in order to have strong noise signal in high variability areas, and low noise in low variability area. We expect here that the assimilation will provide us with a satisfying analysis if the relationship between F_u,F_v and the SSH can be captured by a linear covariance. Additionally, the observations used for the assimilation could contain redundancy. This is expressed by a redundancy factor $\alpha = \sqrt{r}$. It can be shown (Barth et al., 2007) that the error variance must be multiplied by the number of redundant observations r : $\mathbf{R} = r\mu\mathbf{I}$, where μ is the error variance, and $\mathbf{I}$ the identity matrix. $\alpha RMSE$ is thus the square root of the diagonal of $\mathbf{R}$. Hereafter, we refer to $\alpha RMSE$ as the adjusted $RMSE$ ($ARMSE$). Also, all the model errors are not taken into account, which justifies the increase of the $ARMSE$.

The choice of the value of the error variance is critical. Indeed, in the case of an underestimated error variance, the analysis deteriorates unobserved variables. However, if overestimated, the information contained in the observations is not sufficiently transferred into the model.

We perform therefore assimilation with $ARMSE$ values between $10^{-5}m < ARMSE < 10^{2}m$, in order to test the sensitivity and efficiency of the assimilation scheme (Fig. 6a). Indeed, a too small $ARMSE$ on the observations would overconstrain the analysis, and a too large $ARMSE$ would not allow the assimilation scheme to apply a sufficiently large correction. From Fig. 6a, we see that $ARMSE = 4.6$ cm (x-axis) gives the lowest RMSE on the SSH (y-axis) for the assimilation. The corresponding analysed ensemble mean of yearly mean SSH is shown in Fig. 6b. When compared to Fig. 5a, we see that the analysis is satisfactory and is able to retrieve the pattern of the reference run.



(a) (b)

Figure 6: (a) RMSE on SSH from Ensemble Mean before and after analysis, with True Run. (b) Sea surface height of the ensemble mean after assimilation (in m).

However, this is only the first step of our procedure. What we are really interested in is not the direct analysis of the ensemble SSH, but rather the analysis of the zonal and meridional forcings with which we augmented the state vector. Since we considered not to have any information about the true forcing, the initial background estimate (or prior guess) of the forcing is zero. The analysis of the zonal and meridional currents are shown respectively in Fig. 7a and Fig. 7c, and must be compared to the true forcing in Fig. 7b and Fig. 7d. We note that the analysed forcings are convincingly reproducing the structure of the true forcings that we aimed to find.

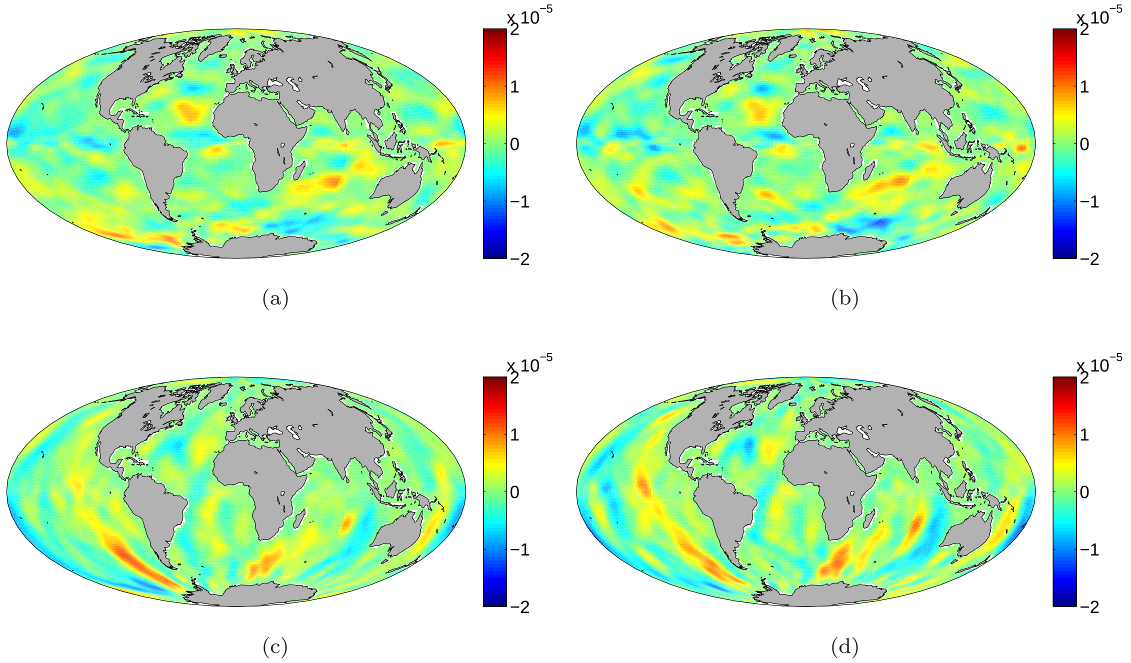
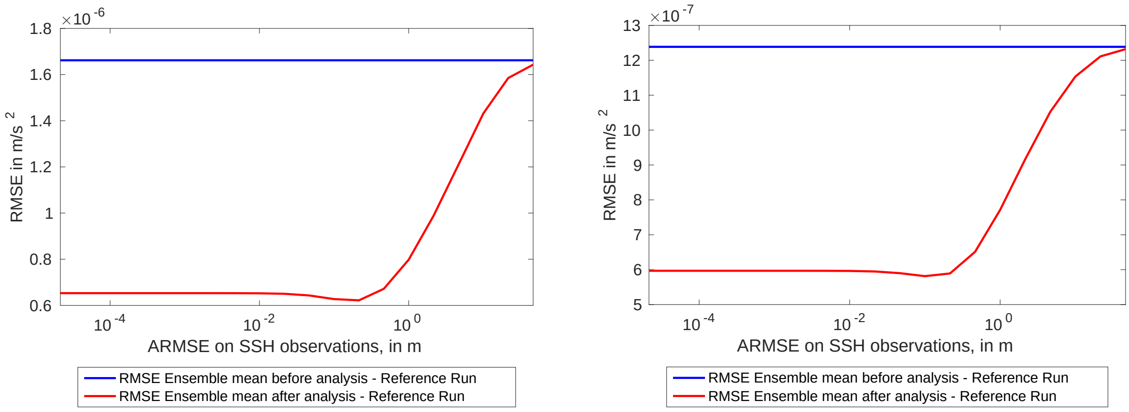


Figure 7: (a) Zonal Forcing ensemble mean after analysis (in ms^{-2}). (b) Zonal Forcing from the true run (in ms^{-2}). (c) Meridional Forcing ensemble mean after analysis (in ms^{-2}). (d) Meridional Forcing from the true run (in ms^{-2}).

Using our twin experiment, and the perfect knowledge that we have on the reference run, we can also look at the RMSE between the analysed forcings and the reference run. This is shown in Fig. 8a and Fig. 8b for the zonal and meridional forcings respectively, with different $ARMSE$ on the SSH observations. We can see that our previous choice of $ARMSE = 4.6$ cm on the observations indeed gives us nearly the best possible results. Since this choice was made solely based on the efficiency of the SSH analysis, we are confident in the relationship between the forcings and the yearly mean SSH of the model.



(a) (b)

Figure 8: (a) RMSE on Zonal Forcing from Ensemble mean before and after Analysis, with True Run. (b) RMSE on Meridional Forcing from Ensemble mean before and after Analysis, with True Run.

One can also compare the total analysed forcing by combining the zonal and meridional components into a vector form in Fig. 9 with the geostrophic currents derived from the SSH bias between the twin experiment reference run and the free model run in Fig. 10. Because of the non-geostrophic balance near the equator, where the horizontal Coriolis force tends to zero, a 5° region around the equator has been removed for this comparison. One can see on Fig. 10 that the geostrophic current derived from the SSH bias is not directly linked to the reference forcing from Fig. 9. This stems from the fact that the forcing affects the model globally, whereas the geostr

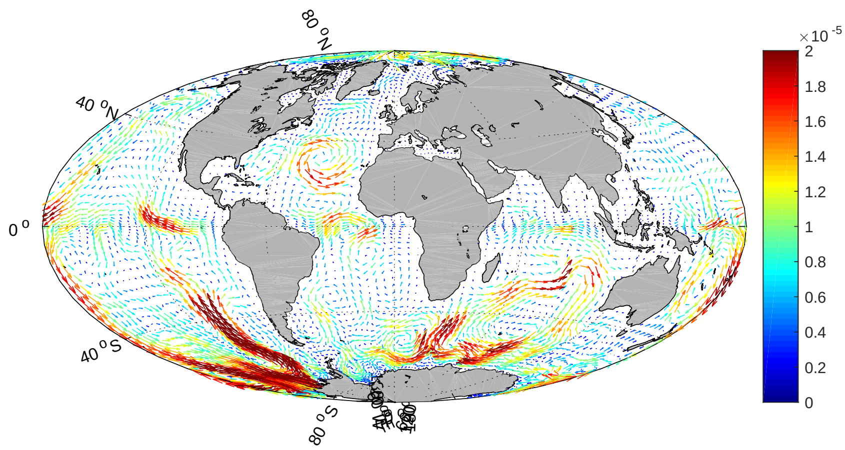


Figure 9: Total forcing ensemble mean after analysis (in ms^{-2}).

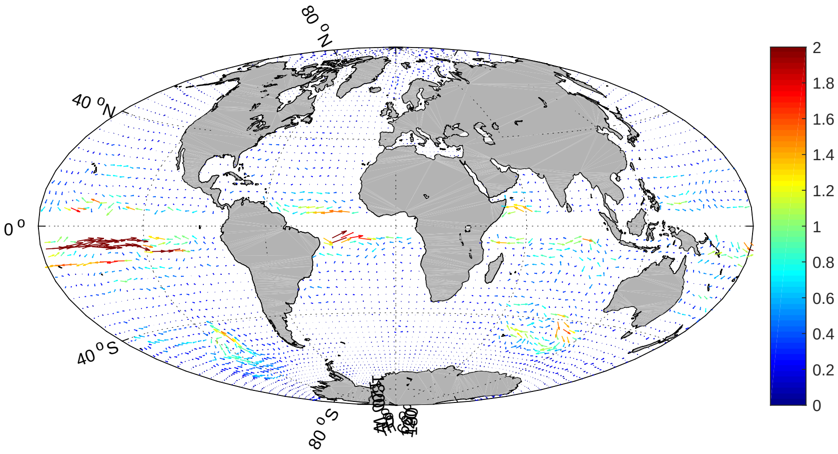
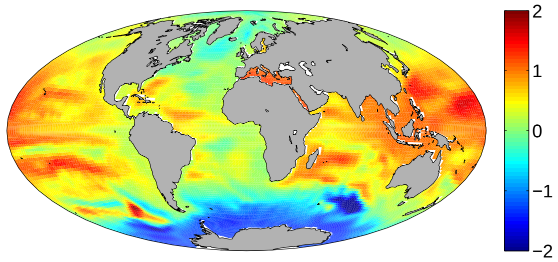
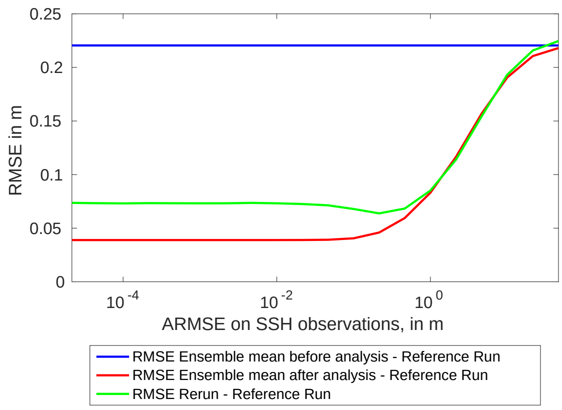


Figure 10: Geostrophic current derived from the SSH bias between the twin experiment reference run and the model free run (in ms^{-1}).

The last step to take is to rerun the model. The forcing from the reference run is considered as the source of the bias acting on the model, and the analysed forcings from the assimilation as the bias correction term to apply to the model. The model is rerun a single time with the analysed ensemble mean forcing, which corresponds to the analysed bias estimator $\widehat{\mathbf{b}}$ from Eq. (14). Without this correction, the model free run without any forcing would be biased. The result of the model rerun with bias correction is shown in Fig. 11a, and can be compared with the true run, displayed in Fig. 5b. Like for the Lorenz '96 case, Fig. 11a is not the result of the assimilation of observations from the true run. It is the rerun of the model with the analysed forcing, obtained from the augmented state vector used during the assimilation procedure. The rerun with bias correction is able to reproduce patterns in the SSH that are particular to the reference run, produced by the true forcing. The last validation of the bias correction term forcing the model is shown in Fig. 11b, where the RMSE on the SSH between the rerun of the model and the true run is compared to the initial ensemble mean and the analysis. One can note that a significant part of the model bias has been removed.



(a)



(b)

Figure 11: (a) Sea surface height (SSH) of the rerun with analysed forcing (in m). (b) RMSE on SSH from Ensemble Mean before and after analysis, and Rerun, with True Run

Further validation of this procedure is done by the comparison of the model forced rerun with the reference run on independent variables. Sea surface temperature (SST) and salinity (SSS) are chosen for their relationship to the currents in the ocean through specific mixing and redistribution of salinity and heat in the ocean. The bias on the currents that this method aims to correct has a direct effect on the SST and SSS. The yearly average SST is shown in Fig. 12a for the ensemble mean, in Fig. 12d for the reference run, and in Fig. 12b for the model rerun with analysed forcing. Fig. 13a, Fig. 13d and Fig. 13b show the SSS for the same runs respectively.

It is clear that typical structures on the SST and SSS fields from the reference run are reproduced by the rerun, and are completely absent on the ensemble mean. One can also note from Fig. 12c and Fig. 13c that the RMSE on the SST and SSS shows a similar behaviour to the RMSE on SSH from Fig. 11b. However, whereas there is a systematic improvement on the SSH reruns with analysed forcings, the analysed forcings appear to be deteriorating the SST and SSS for a specific set of parameters, in particular when the $ARMSE$ on the SSH is large.

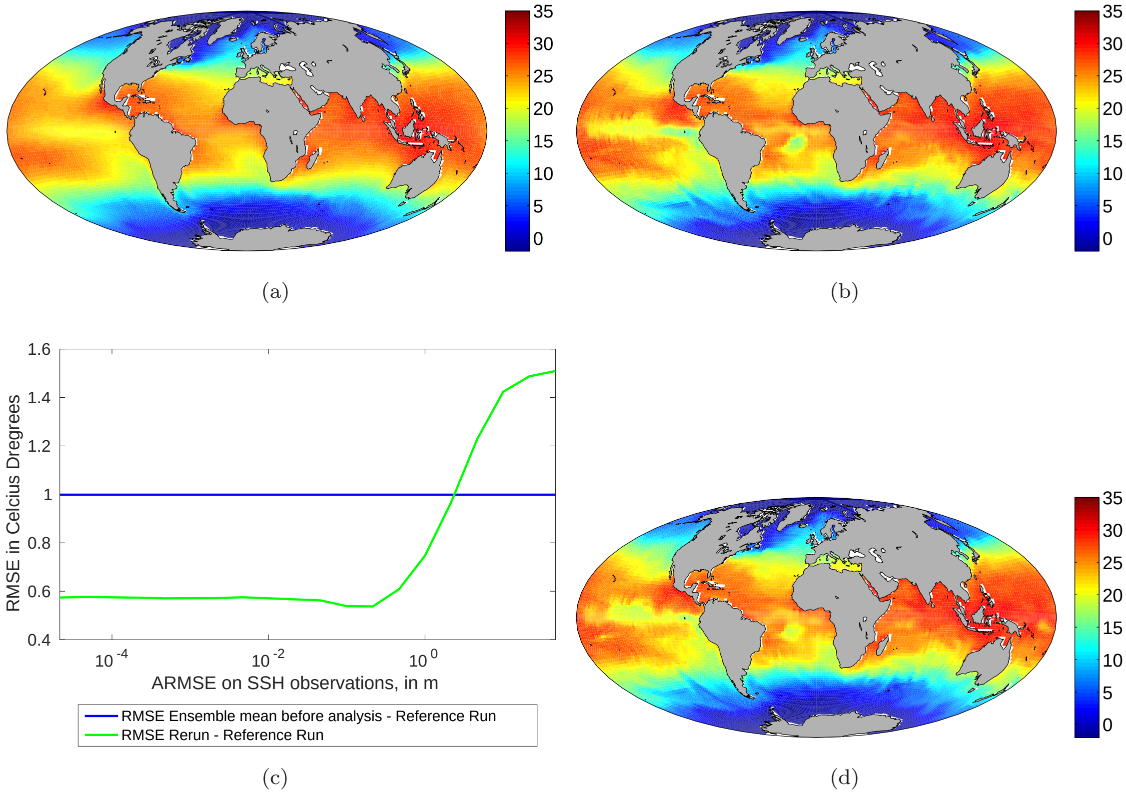


Figure 12: (a) Yearly mean sea surface temperature (SST) of the ensemble mean (in degrees Celsius). (b) Sea surface temperature (SST) of the rerun with analysed forcing (in degrees Celsius). (c) RMSE on SST from Ensemble Mean after analysis, and Rerun, with True Run. (d) Yearly mean sea surface temperature (SST) of the twin experiment true run (in degrees Celsius).

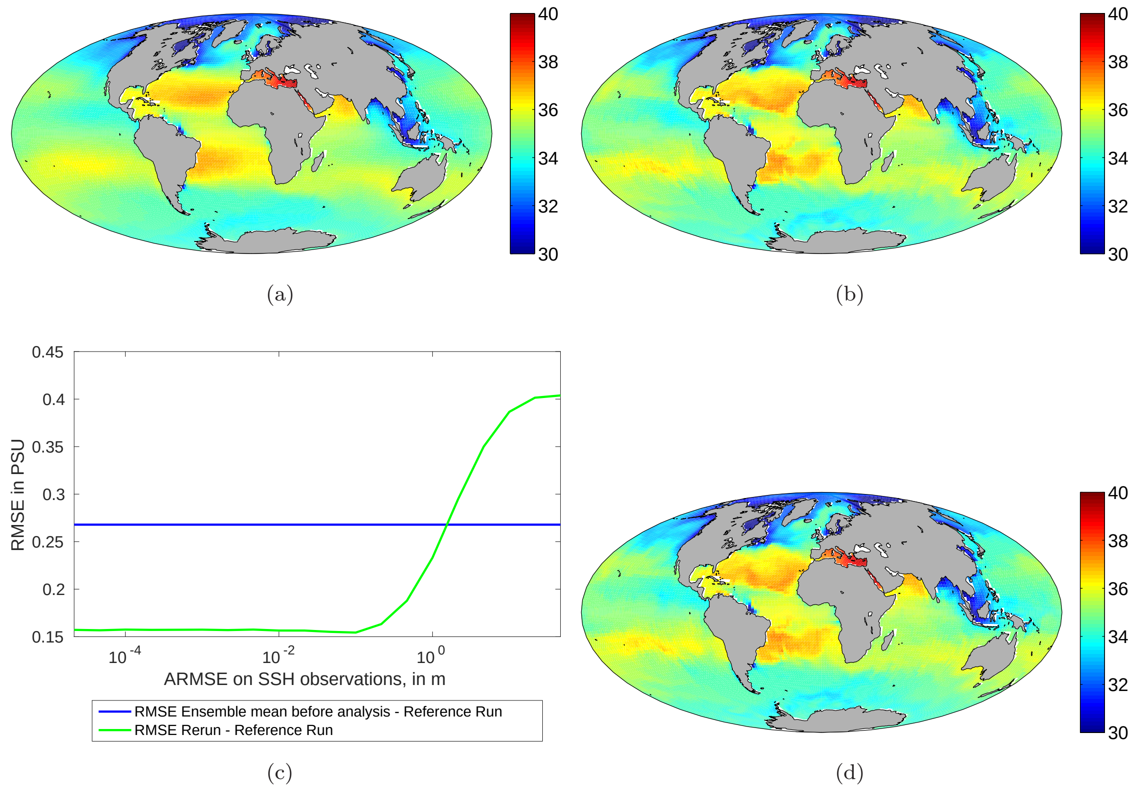


Figure 13: (a) Yearly mean sea surface salinity of the ensemble mean (in PSU). (b) Sea surface salinity of the rerun with analysed forcing (in PSU). (c) RMSE on sea surface salinity from Ensemble Mean after analysis, and Rerun, with True Run. (d) Yearly mean sea surface salinity of the twin experiment true run (in PSU).

6.2 Realistic case

The efficiency of this bias correction method has been successfully tested on a twin experiment test case in the previous section. The following covers the results of this method in a realistic case experiment.

The same setup as the twin experiment is taken for the NEMO model configuration. Observations are however taken from the mean dynamic topography (MDT) of CNES (Centre National d'Etudes Spatiales) (Rio et al., 2011). The SSH provided by the MDT of CNES is interpolated on the ORCA2 grid. Again, an ensemble of forced model runs is created. The observations are assimilated with a range of RMSE fields to find the best compromise between the ensemble and the observations. This procedure provides a forcing which is used to rerun the model. The same parameters as for the twin experiment are taken: a correlation length of 5000 km, 100 ensemble members, and an $ARMSE$ on the SSH observations of 4.6 cm.

The different relevant RMSE are shown in Fig. 14. One can notice that the RMSE between the ensemble mean and ensemble members shows a sufficient enough variability on the model to cover the RMSE between the model free run and the CNES observations. Like in the previous section, the RMSE of the analysed SSH field is significantly reduced compared to the RMSE between the ensemble mean before analysis and the CNES observations. Finally, the rerun of the model with the assimilated forcing shows a significant improvement on the SSH RMSE when compared to the free run. This means that the analysed forcing effectively removes a part of the error of the model on the SSH, through the forcing on the zonal and meridional currents.

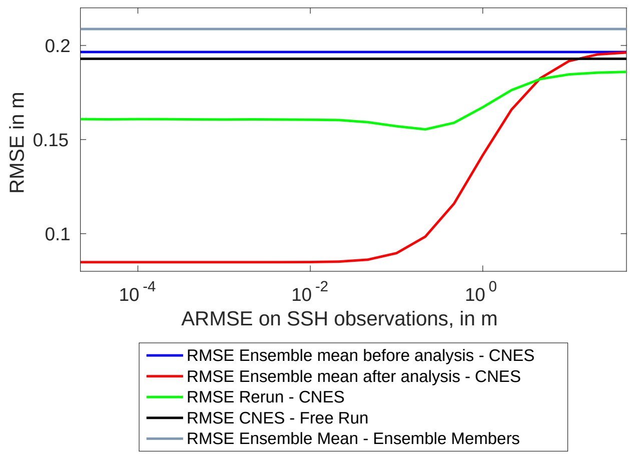
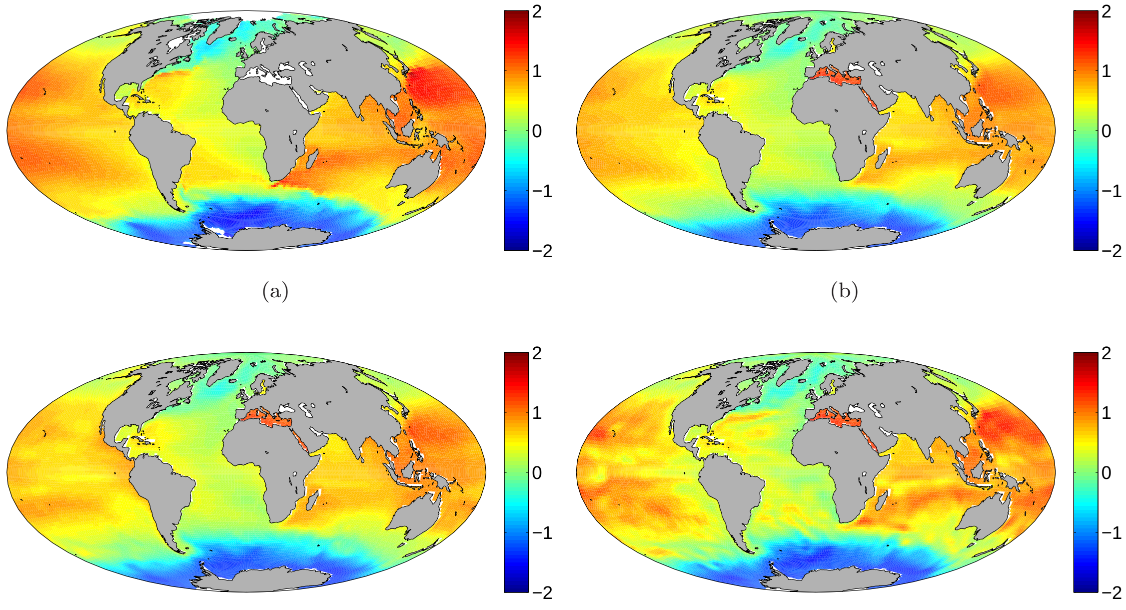


Figure 14: RMSE on SSH from the ensemble mean before and after analysis with CNES observations, from the forced rerun with the observations, from the model free run with the observations, and the internal variability of the ensemble.

More extensive results are shown in the following figures. Fig. 15a shows the interpolated yearly mean SSH of the CNES observations on the ORCA2 grid. Fig. 15b show the yearly mean SSH of the model free run, for the year 1984-1985. in Fig. 15c, the yearly mean SSH of the ensemble mean is shown. One can notice the differences between the model free run and the ensemble mean of forced runs on the yearly mean SSH. This is due to the fact that, even though the ensemble of zonal and meridional forcings has a close to zero mean, the presence of those forcings do increase the currents in the ocean, producing a non-zero mean SSH modification. Finally, 15d shows the yearly mean SSH of the rerun with the analysed forcing.

When comparing figures 15a, 15b and 15d, one can notice the differences on the SSH between the observations, the free model run and the forced rerun. The SSH of the model free run appears to be very smooth and does not show the same variability as the CNES observations. This property, directly influenced by strong, localised, currents, shows to be improved in the forced rerun. In particular, the SSH variations caused by the Gulf Stream are absent from the free run but present in the forced run. Other similar improvements are present around the Cape of Good Hope and along the coast of Chili.



(a) (b) (c) (d)

Figure 15: (a) Yearly mean SSH of the CNES observations (in m). (b) Yearly mean SSH of the model free run (in m). (c) Yearly mean SSH of the ensemble mean (in m). (d) Yearly mean SSH of the lowest RMSE model forced rerun (in m).

The final forcing field produced by this procedure is shown in Fig. 16, in vector form. It is the optimal forcing resulting from the analysis with the CNES SSH observations, applied to the rerun of the NEMO model, a single time, producing the rerun SSH field from Fig. 15d. One must remember that even though the initial perturbations did contain some specific physical constraints, especially regarding the currents perpendicular to the coasts, the correlation lengths and the depth of the forcing, no other properties of the oceanic currents was present in the ensemble of forcings. However, Fig. 15a clearly shows some specific real currents, like the Gulf Stream in the North Atlantic Ocean, the Humboldt Current, in the South Pacific Ocean, or the Antartic Circumpolar current. This result is coherent with the limitations inherent with the low resolution of the NEMO model, which tends to underestimate the strength of those strong currents. The forcing reinforces those currents with a specific correction, effectively accounting for the limitations of the non-corrected model. This forcing, intended to correct current biases in the NEMO model, could thus be used in the future as an additional forcing on the currents to provide a better and more realistic ocean dynamic climatology for NEMO.

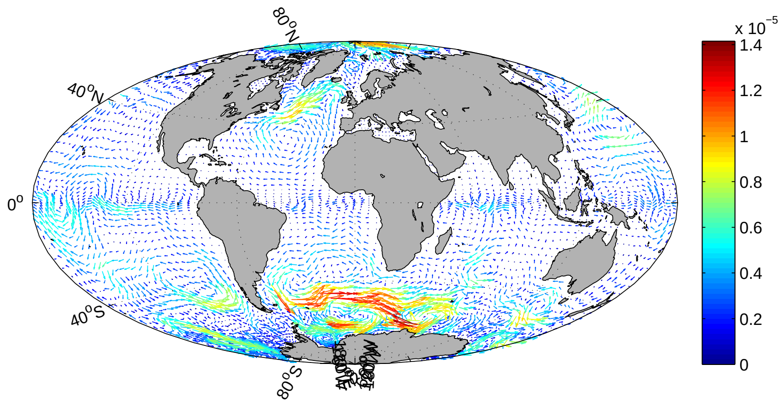
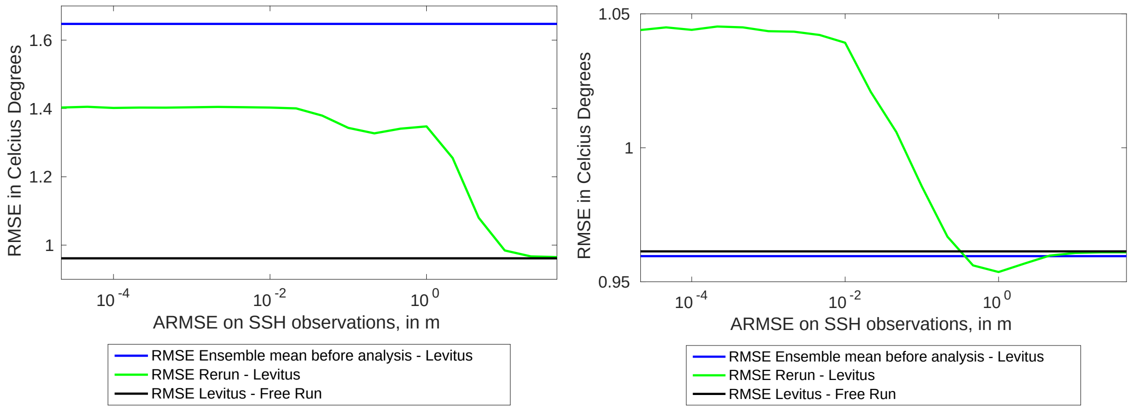


Figure 16: Analysed forcing from CNES observations, used to rerun the model (in ms^{-2}).

In order to validate the final correction field from Fig. 16, the model rerun mean SST is compared against a mean SST climatology (hence observations) from NODC-WOA94 data provided by the NOAA-OAR-ESRL PSD, Boulder, Colorado, USA (Levitus and Boyer, 1994). The RMSE of the model free run, the ensemble mean before assimilation, and the model rerun are shown on Fig. 17a. One can see that the optimal forcing from Fig. 16 does deteriorate the SST. The origin of this behaviour lies in the origin of the model bias. In this work, the bias is only corrected for the ocean circulation, whereas in reality multiple other bias sources also affect the model and the SST. However, with other parameters for the bias correction on the ocean currents, in particular with weaker currents forcing and a correlation length of 10000 km, the effect on the SST climatology of the model rerun shows slight improvements, with RMSE as low as on Fig. 17b. Those results show that a slight improvement can be obtained on other non-assimilated variables, but the complicated relations between the different variables and the model bias renders those improvement particularly difficult to obtain.



(a) (b)

Figure 17: RMSE on SST from Ensemble Mean after analysis, Model Free Run, and Rerun, with Levitus observations. (a) are values for the optimal SSH correction from Fig. 16, (b) are the best obtained results for the SST with weaker forcing and a 10000 km correlation length.

7 Summary and conclusions

In this study, a new method of bias correction through stochastic forcing using data assimilation has been developed. It has first been developed and tested in a fully controlled environment with the Lorenz '96 model. Some properties of this model have also been studied in order to test its responsiveness and the behaviour of the model mean. Due to the successful results, this method was then applied to a twin experiment using the NEMO-LIM2 model with the ORCA2 grid. An effective method for constructing a physically constrained forcing term was used. The assimilation method used allowed the reconstruction of the reference forcing, which showed the efficiency and stability of the assimilation procedure. The method also showed significant improvements on variables that were not included in the assimilated observations.

Finally, this method was tested with real observations on the NEMO-LIM2 model, in order to improve the classic configuration of the model. The assimilation procedure provided a significant improvement on the free run, introducing more variability in the SSH structure, especially around the Gulf Stream. The specific and physical structure of the forcing resulting from the analysis shows the ability of the assimilation procedure to extract, reproduce and correct existing currents on which the NEMO-LIM2 model induces errors. However, those corrections deteriorated other variables, such as the SST.

The encouraging results of both twin experiments shows that as long as the model is able to reproduce the behaviour of the pseudo-observations, the bias correction term is able to effectively improve and diminish the model bias. It is however no longer the case when confronted with real observations due to the model inability to reproduce realistic behaviours. The limitations of the structure of the forcing, as well as the calibration of the different parameters has been pointed out.

One must note though that is was not the objective of this work to find optimal parameters for the bias correction, but rather prove the feasibility of this method. A specific search for optimal parameters, in particular for the real experiment using the CNES MDT and independent SST validation, should provide better results.

Subsequent studies should concentrate on the possibility of assimilating other variables, as well as creating spatially more complex or time-varying forcings to improve the forcing structure. The forcings should also be interpreted in terms of physical processes. The effect of the forcing both on the assimilated and independent variables needs to be examined. This method should also be coupled with other traditional bias estimation schemes of high-frequency variability to provide a dual-estimation of the correction to apply.

8 Acknowledgments

This work was funded by the project PREDANTAR (SD/CA/04A) from the federal Belgian Science policy and the Sangoma FP7-SPACE-2011 project (grant 283580). Alexander Barth is an F.R.S. - FNRS Research Associate. Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11. We also thank the CNES and CLS for the mean dynamic topography. This is a MARE publication.

The final publication is available at Springer via http://dx.doi.org/10.1007/s10236-016-1022-3

References

Anderson, J. L., 2009. Spatially and temporally varying adaptive covariance inflation for ensemble filters. Tellus A 61 (1), 72–83.

Annan, J., Lunt, D., Hargreaves, J., Valdes, P., 2005. Parameter estimation in an atmospheric GCM using the ensemble Kalman filter. Nonlinear Processes Geophysics 12, 363–371.

Baek, S.-J., Hunt, B. R., Kalnay, E., Ott, E., Szunyogh, I., 2006. Local ensemble kalman filtering in the presence of model bias. Tellus A 58 (3), 293–306.

Baek, S.-J., Szunyogh, I., Hunt, B. R., Ott, E., 2009. Correcting for surface pressure background bias in ensemble-based analyses. Monthly Weather Review 137 (7), 2349–2364.

Barth, A., Alvera-Azcárate, A., Beckers, J.-M., Rixen, M., Vandenbulcke, L., 2007. Multigrid state vector for data assimilation in a two-way nested model of the Ligurian Sea. Journal of Marine Systems 65 (1-4), 41–59.
URL http://hdl.handle.net/2268/4260

Barth, A., Alvera-Azcárate, A., Beckers, J.-M., Weisberg, R. H., Vandenbulcke, L., Lenartz, F., Rixen, M., 2009. Dynamically constrained ensemble perturbations - application to tides on the West Florida Shelf. Ocean Science 5 (3), 259–270.

Barth, A., Alvera-Azcárate, A., Gurgel, K.-W., Staneva, J., Port, A., Beckers, J.-M., Stanev, E. V., 2010. Ensemble perturbation smoother for optimizing tidal boundary conditions by assimilation of High-Frequency radar surface currents - application to the German Bight. Ocean Science 6 (1), 161–178.

Barth, A., Beckers, J.-M., Troupin, C., Alvera-Azcárate, A., Vandenbulcke, L., 2014. divand-1.0: n-dimensional variational data analysis for ocean observations. Geoscientific Model Development 7 (1), 225–241.
URL http://www.geosci-model-dev.net/7/225/2014/

Barth, A., Canter, M., Van Schaeybroeck, B., Vannitsem, S., Massonnet, F., Zunz, V., Mathiot, P., Alvera-Azcárate, A., Beckers, J.-M., 2015. Assimilation of sea surface temperature, sea ice concentration and sea ice drift in a model of the southern ocean. Ocean Modelling 93, 22–39.

Bell, M. J., Martin, M., Nichols, N., 2004. Assimilation of data into an ocean model with systematic errors near the equator. Quarterly Journal of the Royal Meteorological Society 130 (598), 873–893.

Bishop, C. H., Etherton, B. J., Majumdar, S. J., 2001. Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. Monthly weather review 129 (3), 420–436.

Bouillon, S., Maqueda, M. A. M., Legat, V., Fichefet, T., 2009. An elastic-viscous-plastic sea ice model formulated on Arakawa B and C grids. Ocean Modelling 27, 174–184.

Broquet, G., Moore, A., Arango, H., Edwards, C., 2011. Corrections to ocean surface forcing in the california current system using 4d variational data assimilation. Ocean Modelling 36 (1), 116–132.

Carton, J. A., Chepurin, G., Cao, X., Giese, B., 2000. A simple ocean data assimilation analysis of the global upper ocean 1950-95. Part I: Methodology. Journal of Physical Oceanography 30 (2), 294–309.

Chepurin, G. A., Carton, J. A., Dee, D., 2005. Forecast model bias correction in ocean data assimilation. Monthly weather review 133 (5), 1328–1342.

Dee, D. P., 2004. Variational bias correction of radiance data in the ECMWF system. In: Proceedings of the ECMWF workshop on assimilation of high spectral resolution sounders in NWP. Vol. 28. pp. 97–112.

Dee, D. P., 2005. Bias and data assimilation. Quarterly Journal of the Royal Meteorological Society 131 (613), 3323–3344.

Dee, D. P., Da Silva, A. M., 1998. Data assimilation in the presence of forecast bias. Quarterly Journal of the Royal Meteorological Society 124 (545), 269–295.

Dee, D. P., Todling, R., 2000. Data assimilation in the presence of forecast bias: The geos moisture analysis. Monthly Weather Review 128 (9), 3268–3282.

Derber, J., Rosati, A., 1989. A global oceanic data assimilation system. Journal of Physical Oceanography 19, 1333–1347.

Derber, J. C., Wu, W.-S., 1998. The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. Monthly Weather Review 126 (8), 2287–2299.

Evensen, G., 2007. Data assimilation: the Ensemble Kalman Filter. Springer, 279pp.

Fertig, E. J., BAEK, S.-J., Hunt, B. R., Ott, E., Szunyogh, I., Aravéquia, J. A., Kalnay, E., Li, H., Liu, J., 2009. Observation bias correction with an ensemble kalman filter. Tellus A 61 (2), 210–226.

Fichefet, T., Maqueda, M. A. M., 1997. Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. Journal of Geophysical Research 102, 12609–12646.

Friedland, B., 1969. Treatment of bias in recursive filtering. Automatic Control, IEEE Transactions on 14 (4), 359–367.

Gelb, A., 1974. Applied optimal estimation. MIT Press, Cambridge, MA, 374 pp.

Gerbig, C., Körner, S., Lin, J., 2008. Vertical mixing in atmospheric tracer transport models: error characterization and propagation. Atmospheric Chemistry and Physics 8 (3), 591–602.

Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J., Sauer, T., Szunyogh, I., Yorke, J. A., Zimin, A. V., 2004. Four-dimensional ensemble Kalman filtering. Tellus 56A, 273–277.

Hunt, B. R., Kostelich, E. J., Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. Physica D 230, 112–126.

Jazwinski, A. H., 1970. Stochastic Processes and Filtering Theory. Academic, San Diego, California.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-Year Reanalysis Project. Bulletin of the American Meteorological Society 77, 437–471.

Keppenne, C. L., Rienecker, M. M., Kurkowski, N. P., Adamec, D. A., 2005. Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. Nonlinear Processes In Geophysics 12, 491–503.

Leeuwenburgh, O., 2008. Estimation and correction of surface wind-stress bias in the tropical pacific with the ensemble kalman filter. Tellus A 60 (4), 716–727.

Levitus, S., Boyer, T., 1994. World ocean atlas 1994. volume 4. temperature. Tech. rep., National Environmental Satellite, Data, and Information Service, Washington, DC (United States).

Li, H., Kalnay, E., Miyoshi, T., 2009. Simultaneous estimation of covariance inflation and observation errors within an ensemble kalman filter. Quarterly Journal of the Royal Meteorological Society 135 (639), 523–533.

Lorenz, E. N., 1963. Deterministic nonperiodic flow. Journal of the Atmospheric Sciences 20, 130–141.

Lorenz, E. N., 1996. Predictability: A problem partly solved. In: Proc. Seminar on predictability. Vol. 1.

Lorenz, E. N., Emanuel, K. A., 1998. Optimal sites for supplementary weather observations: Simulation with a small model. Journal of the Atmospheric Sciences 55, 399–414.

Madec, G., 2008. NEMO ocean engine. No. 27 in Note du Pole de modélisation. Institut Pierre-Simon Laplace (IPSL), France.

Massonnet, F., Goosse, H., Fichefet, T., Counillon, F., 2014. Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble kalman filter. Journal of Geophysical Research: Oceans 119 (7), 4168–4184.

Massonnet, F., Mathiot, P., Fichefet, T., Goosse, H., Beatty, C. K., Vancoppenolle, M., Lavergne, T., 2013. A model reconstruction of the Antarctic sea ice thickness and volume changes over 1980-2008 using data assimilation. Ocean Modelling 64, 67–75.

Mathiot, P., Goosse, H., Fichefet, T., Barnier, B., Gallée, H., 2011. Modelling the seasonal variability of the Antarctic Slope Current. Ocean Science 7 (4), 455–470.

Nerger, L., Gregg, W. W., 2008. Improving assimilation of seawifs data by the application of bias correction with a local seik filter. Journal of marine systems 73 (1), 87–102.

Radakovich, J. D., Bosilovich, M. G., Chern, J.-d., da Silva, A., Todling, R., Joiner, J., Wu, M.-l., Norris, P., 2004. Implementation of coupled skin temperature analysis and bias correction in the NASA/GMAO finite-volume data assimilation system (FvDAS). In: P1. 3 in Proceedings of the Eighth AMS Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface. pp. 12–15.

Radakovich, J. D., Houser, P. R., da Silva, A., Bosilovich, M. G., 2001. Results from global land-surface data assimilation methods. In: AGU Spring Meeting Abstracts. Vol. 1.

Rio, M., Guinehut, S., Larnicol, G., 2011. New CNES-CLS09 global mean dynamic topography computed from the combination of GRACE data, altimetry, and in situ measurements. Journal of Geophysical Research: Oceans (1978–2012) 116 (C7).

Sakov, P., Evensen, G., Bertino, L., 2010. Asynchronous data assimilation with the EnKF. Tellus 62A, 24–29.

Timmermann, R., Goosse, H., Madec, G., Fichefet, T., Ethe, C., Dulière, V., 2005. On the representation of high latitude processes in the ORCA-LIM global coupled sea ice-ocean model. Ocean Modelling 8 (1-2), 175–201.

van Leeuwen, P. J., 2001. An Ensemble Smoother with Error Estimates. Monthly Weather Review 129, 709–728.

van Leeuwen, P. J., 2010. Nonlinear Data Assimilation in geosciences: an extremely efficient particle filter. Quarterly Journal of the Royal Meteorological Society 136, 1991–1996.

Zunz, V., Goosse, H., Massonnet, F., 2013. How does internal variability influence the ability of CMIP5 models to reproduce the recent trend in Southern Ocean sea ice extent? The Cryosphere 7 (2), 451–468.

9 Appendix

One can show that the analysis using the average model state (Eq. (13)) provides the same analysed bias $\widehat{\mathbf{b}^a}$ as when the full trajectory is included in the estimation vector (Eq. (6)).

Using $i = 1, \ldots, N$ to refer to the ensemble members, the forecast of the model trajectory can be defined as

$$\mathbf{x'}_i^f = \begin{bmatrix} \mathbf{x}_i^{f(1)} \\ \mathbf{x}_i^{f(2)} \\ \vdots \\ \mathbf{x}_i^{f(m_{\max})} \\ \widehat{\mathbf{b}}_i^f \end{bmatrix}, \qquad \mathbf{x'}_i^a = \begin{bmatrix} \mathbf{x}_i^{a(1)} \\ \mathbf{x}_i^{a(2)} \\ \vdots \\ \mathbf{x}_i^{a(m_{\max})} \\ \widehat{\mathbf{b}}_i^a \end{bmatrix}. \tag{37}$$

The analysis is provided by

$$\mathbf{x'}^a = \mathbf{x'}^f + \frac{1}{N-1}\mathbf{X'}^f \underbrace{(\mathbf{X'}^f)^T \mathbf{H'}^T (\mathbf{H'}\mathbf{P'}^f \mathbf{H'}^T + R)^{-1}(\mathbf{y}^o - \mathbf{H'}\mathbf{x'}^f)}_{\mathbf{W'}}, \tag{38}$$

where

$$\mathbf{x'}^f = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x'}_i^f, \qquad \mathbf{x'}^a = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x'}_i^a, \tag{39}$$

$$\mathbf{P'}^f = \frac{1}{N-1}\sum_{i=1}^{N} (\mathbf{x'}_i^f - \mathbf{x'}^f)(\mathbf{x'}_i^f - \mathbf{x'}^f)^T \tag{40}$$

$$= \frac{1}{N-1}\mathbf{X'}^f (\mathbf{X'}^f)^T. \tag{41}$$

The observation operator $\mathbf{H'}$ applied to the trajectory $\mathbf{x'}$ also includes a time average and an extraction operator $\mathbf{H}$ of the observed part of the model state

$$\mathbf{H'}\mathbf{x'} = \sum_{m=1}^{m_{max}} \mathbf{H}\mathbf{x}^{(m)} = \mathbf{H}\overline{\mathbf{x}}, \tag{42}$$

$$\overline{\mathbf{x}} = \frac{1}{m_{max}} \sum_{m=1}^{m_{max}} \mathbf{x}^{(m)}. \tag{43}$$

Hence, the ensemble mean of the analysed bias correction term $\widehat{\mathbf{b}'^a}$ is contained in the analysed model trajectory $\mathbf{x}'^a$. One can also first take the time average of the trajectory, defined as

$$\mathbf{x}''^f_i = \begin{bmatrix} \overline{\mathbf{x}}^f_i \\ \widehat{\mathbf{b}}^f_i \end{bmatrix}, \qquad \mathbf{x}''^a_i = \begin{bmatrix} \overline{\mathbf{x}}^a_i \\ \widehat{\mathbf{b}}^a_i \end{bmatrix}. \tag{44}$$

The analysis is then given by

$$\mathbf{x}''^a = \mathbf{x}''^f + \frac{1}{N-1}\mathbf{X}''^f \underbrace{(\mathbf{X}''^f)^T \mathbf{H}''^T (\mathbf{H}''\mathbf{P}''^f\mathbf{H}''^T + R)^{-1}(\mathbf{y}^o - \mathbf{H}''\mathbf{x}''^f)}_{\mathbf{W}''}, \tag{45}$$

where

$$\mathbf{x}''^f = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}''^f_i, \qquad \mathbf{x}''^a = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}''^a_i, \tag{46}$$

$$\begin{aligned} \mathbf{P}''^f &= \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}''^f_i - \mathbf{x}''^f)(\mathbf{x}''^f_i - \mathbf{x}''^f)^T & (47)\\ &= \frac{1}{N-1}\mathbf{X}''^f(\mathbf{X}''^f)^T. & (48) \end{aligned}$$

The ensemble mean of the analysed bias correction term $\widehat{\mathbf{b}''^a}$ is contained in the analysed mean model state $\mathbf{x}''^a$. Given that

$$\mathbf{H}'\mathbf{x}' = \mathbf{H}''\mathbf{x}'', \tag{49}$$

it follows that $\mathbf{W}' = \mathbf{W}''$. Hence, $\widehat{\mathbf{b}''^a} = \widehat{\mathbf{b}'^a}$, since they are both constrained by the same linear combination of $\widehat{\mathbf{b}}^f_i$.