An idealized study of coupled atmosphere-ocean 4D-Var in the presence of model error

Article

Accepted Version

Fowler, A. M. ORCID: https://orcid.org/0000-0003-3650-3948 and Lawless, A. S. ORCID: https://orcid.org/0000-0002-3016-6568 (2016) An idealized study of coupled atmosphere-ocean 4D-Var in the presence of model error. Monthly Weather Review, 144 (10). pp. 4007-4030. ISSN 0027-0644 doi: https://doi.org/10.1175/MWR-D-15-0420.1 Available at https://centaur.reading.ac.uk/65663/

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

An idealized study of coupled atmosphere-ocean 4D-Var in the presence of model error

Alison M. Fowler and Amos S. Lawless

School of Mathematical and Physical Sciences, University of Reading, Reading, RG6 6AX, UK.

Abstract

Atmosphere only and ocean only variational data assimilation (DA) schemes are able to use window lengths that are optimal for the error growth rate, non-linearity and observation density of the respective systems. Typical window lengths are 6-12 hours for the atmosphere and 2-10 days for the ocean. However, in the implementation of coupled DA schemes it has been necessary to match the window length of the ocean to that of the atmosphere, which may potentially sacrifice the accuracy of the ocean analysis in order to provide a more balanced coupled state. This paper investigates how extending the window length in the presence of model error affects both the analysis of the coupled state and the initialized forecast when using coupled DA with differing degrees of coupling.

Results are illustrated using an idealized single column model of the coupled atmosphere-ocean system. It is found that the analysis error from an uncoupled DA scheme can be smaller than that from a coupled analysis at the initial time, due to faster error growth in the coupled system. However, this does not necessarily lead to a more accurate forecast due to imbalances in the coupled state. Instead coupled DA is more able to update the initial state to reduce the im-

pact of the model error on the accuracy of the forecast. The effect of model error is potentially most detrimental in the weakly coupled formulation due to the inconsistency between the coupled model used in the outer loop and uncoupled models used in the inner loop.

1 Introduction

Coupling processes between the atmosphere and ocean are known to be important for seasonal and climate prediction, for example for the accurate prediction of the El Niño-Southern Oscillation (Barnett et al. 1993; Jin et al. 2008). In addition to this, there is increasing evidence to suggest the importance of atmosphere-ocean interaction at the weather time scales, for example for the prediction of the Madden-Jullen Oscillation (Shelly et al. 2014), coastal fog and extra-tropical cyclones (Siddorn et al. 2014; Vitart et al. 2008).

Until recently the initialization of coupled models for the prediction of the Earth system has been performed using uncoupled atmosphere and ocean states, produced from separate data assimilation (DA) systems (e.g. Saha et al. (2006); Molteni et al. (2011); Arribas et al. (2011); MacLachlan et al. (2015)). These uncoupled states are then effectively 'stitched' together to

create an initial state for the coupled forecast. There are a variety of known problems with using uncoupled initial conditions. These include

- the generation of imbalances between the atmosphere and ocean systems leading to an unrealistic and sudden adjustment to balanced conditions in the first part of the forecast (Balmaseda and Anderson 2009; Mullholland et al. 2015); and

- a sub-optimal use of observations, particularly those close to the air-sea interface which depend on the physics of both the ocean and atmosphere.

In an attempt to overcome many of these issues, coupled DA methods, in which the atmosphere and ocean are treated simultaneously, are being developed at operational centers worldwide (Saha et al. 2010; Lea et al. 2015; Alves et al. 2014; Laloyaux et al. 2015). As well as for the initialization of coupled models, coupled DA will be used for reanalyses, in which it is necessary to provide a consistent global transport of mass, water, and energy on the relevant time scales (Dee et al. 2014).

Coupled DA is a relatively new field of research. Early studies include those by Galanti et al. (2003), Zhang et al. (2007) and Sugiura et al. (2008). Despite the techniques proposed in these early studies not estimating the full atmosphere and ocean states simultaneously, results showed that using a coupled atmosphere-ocean model during the assimilation significantly improved the accuracy of reanalyses and forecasts of seasonal to interannual climate variations (such as El Niño). Another early study is that of Lu and Hsieh (1998), who used a simple 5 dimensional coupled equatorial model to investigate the use of coupled DA for both state and parameter estimation. Within this study we focus on the use of coupled DA for short-medium range weather forecasting which can be sensitive to the interaction between the atmospheric boundary layer and the oceanic mixed layer.

Implementing a fully coupled DA scheme in practice faces many challenges. An immediate practical consideration is the increase in the size of the state that needs to be estimated, which makes the problem much more computationally expensive. There are also more scientific challenges which are invariably related to the very different nature of the two fluids. The atmosphere is much less dense than the ocean and also much more unstable. There are also fundamental differences in the available observations of the atmosphere and ocean. The atmosphere is relatively densely observed whilst the ocean by its very nature is difficult to observe and subsurface observations generally only come from a sparse array of in-situ observations. These differences impact on essential aspects of the implementation of coupled DA, such as the specification of cross-covariances between errors in the atmosphere and ocean systems (Han et al. 2013).

In previous work, we used an idealized single column model to investigate different approximations to coupled DA in the 4D-Var framework when no model error is present (Smith et al. 2015). It was found that when using a short window length of 12 hours (consistent with that used operationally for atmosphere only DA), a strongly coupled formulation was able to provide the best analysis of the coupled initial state, having both a smaller error and being more balanced than an uncoupled formulation. A weakly coupled approximation was seen to display some of these same benefits as the strongly coupled formulation but was more sensitive to the resolution of the observations. In this paper we con-

tinue this study, by investigating how the different approximations to coupled DA compare as the window length is extended and model error becomes a greater issue.

Due to the different nature of the two fluids, model error in the atmosphere has a much faster growth rate than in the ocean. The growth rate of the model error and the linearity assumption restricts the length of the window in which observations can be assimilated in 4D-Var, and is an important factor in why a short window length of 6-12 hours is typically used in atmosphere only data assimilation (Rawlins et al. 2007), and a long window length of 2-10 days is typically used in ocean only data assimilation (Weaver et al. 2003). Coupling the two systems allows for the model error to interact and may introduce a faster model error growth rate in the upper ocean than in an uncoupled simulation. Therefore the model error in the coupled system restricts the assimilation window to something shorter than the optimal window length for an uncoupled ocean DA scheme (Lea et al. 2015; Laloyaux et al. 2015). As the ocean is poorly observed, this severely limits the number of observations that can be assimilated and has the effect of potentially sacrificing the accuracy of the ocean initial state in order to provide a more balanced coupled initial state. Even with the use of a shortened window length, the ECMWF have found it necessary to constrain the sea surface temperature to a gridded analysis product in order to avoid the rapidly growing bias in the model (Laloyaux et al. 2015). There is therefore motivation to understand the effect of extending the window length, in the presence of increasing model error, within the coupled DA framework.

The model error in the coupled system was studied by both Magnusson et al. (2013) and Smith et al. (2013) using the ECMWF IFS model coupled to the NEMO ocean model (Madec 2008) and the HadCM3 model respectively. Both found cold biases in surface temperatures. Magnusson et al. (2013) believed this is due to imbalances in the energy flux at the top of the atmosphere and a strong uptake of heat in the ocean. Magnusson et al. (2013) also looked at the bias in the 10m wind speed and found it to be large within the western Tropical Pacific. They concluded that this is due to a positive coupled feedback between wind and SST: too strong winds lead to excessive upwelling which produces a colder sea surface temperatures (SST) which in turn produces stronger zonal winds.

Here we examine how model errors in the coupled system affect the analysis and subsequent forecast in coupled DA. The structure of the paper is as follows. In section 2 the different approximations to coupled DA that we consider are given. This is largely an overview of results recently presented in Smith et al. (2015). In section 3 the theoretical impact of model error is presented with discussion of how the different coupling strategies may be affected. In section 4 the design of the idealized experiments is presented and an illustration of the model error for a case study is shown. In section 5 results are shown for a series of assimilation experiments applied to this case study in which the effect of model error on the accuracy of the analysis and initialized forecast are given. Finally conclusions that can be drawn from these experiments are detailed in section 6 along with a discussion of the insight provided on how to account for model error in coupled 4D-Var.

2 Strategies for coupled DA

This present study will focus on incremental 4D-Var methods (Courtier et al. 1994) in line with those being developed at ECMWF and the Met Office (Laloyaux et al. 2015; Lea et al. 2015). We note that methods based on the ensemble Kalman filter are also being developed with interesting results (e.g. Frolov et al. (2016) and Zhang et al. (2007)) but will not be included in the comparison of methods presented here. We consider three different strengths of coupling within the 4D-Var scheme:

1. *strongly coupled*, in which the ocean and atmosphere are updated together. This is the epitome of what coupled DA schemes are trying to achieve, but is currently not practical for operational sized models;

2. *weakly coupled*, an approximation to strongly coupled which makes use of the inner loop structure of incremental 4D-Var to simplify the algorithm. This makes it feasible to implement in current operational settings, particularly when different assimilation schemes have been implemented in the atmosphere and ocean;

3. and lastly *uncoupled* in which, the atmosphere and ocean are updated independently of each other.

An in-depth description and study of these different coupling strategies in the case of no model error is given in Smith et al. (2015). Here a brief summary of the different algorithms is given.

As stated above, each method is based on incremental 4D-Var which minimizes iteratively a series of linear approximations to the full 4D-Var

cost function given by

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^{\text{b}})^{\text{T}}\mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^{\text{b}}) \\ + \frac{1}{2}(\hat{\mathbf{y}} - \hat{\mathcal{H}}(\mathbf{x}_0))^{\text{T}}\hat{\mathbf{R}}^{-1}(\hat{\mathbf{y}} - \hat{\mathcal{H}}(\mathbf{x}_0)).$$

(1)

The two sources of information about the initial state are given by $\mathbf{x}_0^{\text{b}}$, the background state (or first guess, usually provided by a previous forecast valid for the time of interest) and $\hat{\mathbf{y}}$, a vector of all observations throughout the assimilation window. The operator $\hat{\mathcal{H}}(\mathbf{x}_0)$ is the generalized observation operator, a mapping from the initial state to all the observation variables and times. This differs from the observation operator, $\mathcal{H}_i(.)$, which maps the model state at time t_i to observation variables at time t_i. $\mathbf{B}$ is the background error covariance matrix and $\hat{\mathbf{R}}$ is the observation error covariance matrix.

The incremental 4D-Var algorithm can be expressed in terms of a series of outer and inner loops, illustrated in figure 1. In the outer loop the linearization state, $\mathbf{x}_{0:n}^{\text{g}}$ over the time window $[t_0, t_n]$, and the innovations, $\delta\hat{\mathbf{y}}$, are evaluated. The linearization state is given by

$$\mathbf{x}_{0:n}^{\text{g}} = \begin{pmatrix} \mathbf{x}_0^{\text{g}} \\ \mathcal{M}(\mathbf{x}_0^{\text{g}}, t_0, t_1) \\ \vdots \\ \mathcal{M}(\mathbf{x}_0^{\text{g}}, t_0, t_n) \end{pmatrix},$$

(2)

where n is the number of time steps in the assimilation window and $\mathcal{M}(\mathbf{x}_0, t_0, t_i)$ is the propagation of the model state at time t_0 to time t_i. The vector $\mathbf{x}_0^{\text{g}}$ refers to the current estimate of the initial state; initially this will be the background, but it will then be updated on each outer loop. The innovations are then given by the difference between the observations and the estimate of the initial state mapped to observation space,

$$\delta\hat{\mathbf{y}} = \hat{\mathbf{y}} - \hat{\mathcal{H}}(\mathbf{x}_0^{\text{g}}).$$

(3)

The computation of the initial state mapped to observation space, $\hat{\mathcal{H}}(\mathbf{x}_0^{\mathrm{g}})$, utilizes the linearization state in the following way:

$$\left(\hat{\mathcal{H}}(\mathbf{x}_0^{\mathrm{g}})\right)_i = \mathcal{H}_i(\mathcal{M}(\mathbf{x}_0^{\mathrm{g}}, t_0, t_i)) = \mathcal{H}_i(\mathbf{x}_i^{\mathrm{g}}). \quad (4)$$

In the inner loop a linear approximation to the full cost function (1) is then minimized with respect to $\delta\mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_0^{\mathrm{g}}$:

$$J(\delta\mathbf{x}_0) = \quad \tfrac{1}{2}(\delta\mathbf{x}_0 - (\mathbf{x}_0^{\mathrm{b}} - \mathbf{x}_0^{\mathrm{g}}))^{\mathrm{T}}\mathbf{B}^{-1}(\delta\mathbf{x}_0 - (\mathbf{x}_0^{\mathrm{b}} - \mathbf{x}_0^{\mathrm{g}}))$$
$$+ \tfrac{1}{2}(\delta\hat{\mathbf{y}} - \hat{\mathbf{H}}\delta\mathbf{x}_0)^{\mathrm{T}}\hat{\mathbf{R}}^{-1}(\delta\hat{\mathbf{y}} - \hat{\mathbf{H}}\delta\mathbf{x}_0)), \quad (5)$$

where $\hat{\mathbf{H}}\delta\mathbf{x}_0$ is a tangent linear (TL) approximation to $\hat{\mathcal{H}}(\mathbf{x}_0) - \hat{\mathcal{H}}(\mathbf{x}_0^{\mathrm{g}})$, linearized about $\mathbf{x}_{0:n}^{\mathrm{g}}$. This can be separated out in terms of the TL approximation to the observation operator, $\mathbf{H}$, and the dynamical model, $\mathbf{M}$.

$$\hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1\mathbf{M}_{0:1} \\ \vdots \\ \mathbf{H}_n\mathbf{M}_{0:n} \end{pmatrix}, \quad (6)$$

where

$$\mathbf{M}_{0:i} = \frac{\partial\mathcal{M}(\mathbf{x}_0, t_0, t_i)}{\partial\mathbf{x}_0}. \quad (7)$$

Once the minimization of $J(\delta\mathbf{x}_0)$ has met a given criterion, the estimate of the initial state is updated: $\mathbf{x}_0^{\mathrm{g}} := \mathbf{x}_0^{\mathrm{g}} + \delta\mathbf{x}_0$. The 'analysis', which is used to initialize the forecast, is then given by $\mathbf{x}_0^{\mathrm{g}}$ once the outer loop has converged, or a maximum number of iterations has been performed (Courtier et al. 1994; Lawless 2013).

2.1 Strongly coupled

In strongly coupled DA the state vector, $\mathbf{x}_0$, includes both the atmospheric and oceanic variables. The non-linear (NL) model, $\mathcal{M}$, used to calculate (2) is the fully coupled model and the

exact TL approximation to $\hat{\mathcal{H}}(\mathbf{x}_0) - \hat{\mathcal{H}}(\mathbf{x}_0^{\mathrm{g}})$ is used in (5). Within this formulation, the covariance matrix $\mathbf{B}$ may include cross covariances between errors in the atmosphere and ocean and the observation operator may account for the sensitivity of observations to both the atmosphere and ocean. Exactly how to derive the background error cross covariances between the atmosphere and ocean is an area of active research (e.g. Han et al. (2013)).

A benefit of 4D-Var over 3D-Var is that the inclusion of the model dynamics allows for some flow-dependence to develop in the background error covariances (Thépaut et al. 1996). In Smith et al. (2015) strongly coupled DA was found to be able to generate implicit correlations between the atmosphere and ocean states and hence observations of the atmosphere were able to influence the analysis of the ocean and vice versa, even when no explicit correlations in the background errors were specified. This provided a more balanced analysis and allowed for more information to be extracted from the observations.

2.2 Weakly coupled

In weakly coupled DA the coupling between the atmosphere and ocean is only accounted for in the outer loop. In practice, this means that the state vector, $\mathbf{x}_0$, still includes both the atmospheric and oceanic variables and the NL coupled model is still used to calculate (2). However, instead of using the exact TL approximation as in strongly coupled DA, an uncoupled approximation is used. Therefore (5) can be split into two cost functions, one corresponding to finding the increment to the atmospheric state, $\delta\mathbf{x}_0^{\mathrm{atmos}}$, and the other corresponding to finding the increment to the oceanic state, $\delta\mathbf{x}_0^{\mathrm{ocean}}$. This means that although it is possible for the observation
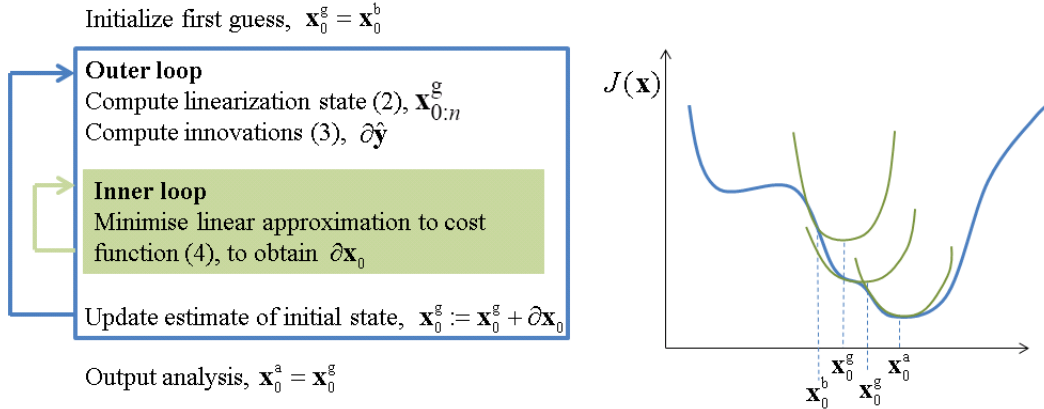
Figure 1: Left: Schematic of the incremental 4D-Var algorithm. Right: Illustration of the the non-linear cost function (blue) and the linear approximations made on each inner loop (green).

operator to account for the sensitivity of observations simultaneously to both the atmosphere and ocean in (3), the calculation of $\delta \mathbf{x}_0^{\mathrm{atmos}}$ and $\delta \mathbf{x}_0^{\mathrm{ocean}}$ in the minimization of (5) does not account for this. Similarly it is not possible to include cross covariances between background errors in the atmosphere and ocean.

As suggested by the name, this formulation reduces the strength of the coupling. Since the innovations (3) are computed in observation space and $\mathcal{H}$ is able to contain contributions from both the updated ocean and atmosphere, the strength of the coupling can be seen to be related to the density of the observations, particularly those which are sensitive to the coupling of the ocean and atmosphere. However as $\mathbf{B}$, $\mathbf{H}$ and $\mathbf{M}$ are not coupled, atmospheric observations cannot update the ocean state directly, and vice versa. Therefore, even if cross correlations between the atmosphere and background errors were not included in strongly coupled DA, it is not simply a case of the weakly coupled scheme converging

to the strongly coupled scheme as more outer loop iterations are performed. This means that compared to strongly coupled DA the risk of initialization shock is increased and weakly coupled DA is particularly sensitive to the frequency and density of observations (Smith et al. 2015).

This approach is the essence of that currently being investigated by the ECMWF (Laloyaux et al. 2015) and Met Office (Lea et al. 2015). However, in the initial implementation at these centres there are some differences to the clean form we present here. Most notably, both the ECMWF and UK Met Office only use 4D-Var for the atmosphere and instead use 3D-Var FGAT (first guess at appropriate time) for the ocean. Another important point is that the UK Met Office currently only perform one outer-loop whilst the ECMWF perform two, increasing the strength of the coupling.

2.3 Uncoupled

In uncoupled DA the state vector is separated between the atmosphere and ocean components. The uncoupled non-linear models are used to compute (2), with the boundary conditions at the interface specified externally, and the exact TL approximation of each uncoupled model is used in the minimization of (5) to approximate the non-linear models used. There is therefore no exchange in information between the atmosphere and ocean and the analysis increments $\delta\mathbf{x}_0^{\text{atmos}}$ and $\delta\mathbf{x}_0^{\text{ocean}}$ may be inconsistent. It is this inconsistency that can lead to initialization shocks.

Our aim is to understand how each of these coupling strategies react to errors in the coupled model equations. In particular, we examine whether there may be benefit to using the uncoupled and weakly coupled formulations if the error growth rate is larger in the coupled model compared with the uncoupled model. In the next section a brief summary of the theory of model error in 4D-Var is given along with an examination of how model error may enter into the different coupling strategies.

3 Model error in 4D-Var

There are many different potential sources of model error in coupled atmosphere-ocean models. In practice it is rare to be able to identify (and correct) a single cause of model error within the model itself. Nevertheless attempts have been made. For example Vannière et al. (2014) demonstrated a systematic approach for identifying sources of error in tropical SSTs. This involved performing 7 separate simulations using coupled and uncoupled models, with differ-ent coupling, forcing and initialization strategies. Such a method shows promise for aiding model development to reduce model error but it cannot eradicate it.

In 4D-Var, model error can be seen to manifest itself through the calculation of the innovations, (3), in which it is assumed that $\hat{\mathcal{H}}(\mathbf{x}_0)$ provides the exact mapping between state and observation space. In the presence of model error this no longer holds. Instead, under the assumption of additive model error, we have the following relationship between the truth in state space, $\mathbf{x}_0^{\text{t}}$, and the truth in observation space, $\hat{\mathbf{y}}^{\text{t}}$:

$$\begin{aligned}\hat{\mathbf{y}}^{\text{t}} \ &= \hat{\mathcal{H}}^{\text{t}}(\mathbf{x}_0^{\text{t}})\\ &= \hat{\mathcal{H}}(\mathbf{x}_0^{\text{t}}) + \boldsymbol{\epsilon}^{\hat{\mathcal{H}}},\end{aligned} \tag{8}$$

where $\hat{\mathcal{H}}^{\text{t}}$ is the exact mapping but $\hat{\mathcal{H}}$ is the mapping used within the assimilation. The error in the generalized observation operator, $\boldsymbol{\epsilon}^{\hat{\mathcal{H}}}$, has the same dimensions as the observation vector and incorporates error in the model equations described in (2). Within this current work it is assumed that the error arises from the dynamical equations rather than the observation operator $\mathcal{H}$. Errors may also be present in the observation operator, which are often considered to be errors of representativity. This is a vast error of research (e.g. Waller et al. (2013), van Leeuwen (2014), Bormann et al. (2014)) but will not be considered further in this current work, that is we will assume the observation operator, $\mathcal{H}$, to be perfect.

If model error is unaccounted for then $\mathbf{B}$, $\hat{\mathbf{R}}$ and $\hat{\mathbf{H}}$ remain unchanged in the 4D-Var algorithm. Therefore the computation of the analysis, $\mathbf{x}^{\text{a}}$, can still be given by the theoretical linear approximation:

$$\mathbf{x}^{\text{a}} = \mathbf{x}^{\text{b}} + \mathbf{K}(\hat{\mathbf{y}} - \hat{\mathcal{H}}(\mathbf{x}^{\text{b}})), \tag{9}$$

where $\mathbf{K}$ is known as the Kalman gain matrix given by $\mathbf{B}\hat{\mathbf{H}}^{\mathrm{T}}(\hat{\mathbf{R}} + \hat{\mathbf{H}}\mathbf{B}\hat{\mathbf{H}}^{\mathrm{T}})^{-1}$. With multiple outer loops $\mathbf{x}^{\mathrm{b}}$ in (9) would be replaced by the current outer loop iterate, and $\mathbf{K}$ computed with a $\hat{\mathbf{H}}$ linearized around the current outer loop trajectory. However, as the generalized observation operator is no longer optimal the error covariance of the analysis will be inflated, with covariance equal to

$$\mathbf{P}^{\mathrm{a}} + \mathbf{K}E[(\boldsymbol{\epsilon}^{\hat{\mathcal{H}}} - E(\boldsymbol{\epsilon}^{\hat{\mathcal{H}}}))(\boldsymbol{\epsilon}^{\hat{\mathcal{H}}} - E(\boldsymbol{\epsilon}^{\hat{\mathcal{H}}}))^{\mathrm{T}}]\mathbf{K}^{\mathrm{T}}, \quad (10)$$

where $E[.]$ is the mathematical expectation and $\mathbf{P}^{\mathrm{a}}$ is the analysis error covariance if $\hat{\mathcal{H}}(\mathbf{x}_0^{\mathrm{t}})$ were correct. In deriving (10), it is assumed that the model error is uncorrelated with the observation and background error. In addition to this, if the random variable $\boldsymbol{\epsilon}^{\hat{\mathcal{H}}}$ is biased then the analysis will be biased, that is

$$E[\boldsymbol{\epsilon}^{\mathrm{a}}] = \mathbf{K}E[\boldsymbol{\epsilon}^{\hat{\mathcal{H}}}], \quad (11)$$

where $\boldsymbol{\epsilon}^{\mathrm{a}} = \mathbf{x}^{\mathrm{a}} - \mathbf{x}_0^{\mathrm{t}}$ (the analysis minus the true initial state). The derivations of (10) and (11) can be found in the appendix. In a similar way, it can be shown that the expected value of the cost function evaluated at the analysis will also be increased, as it becomes impossible to fit to both the background and the observations in a way which is consistent with the prescribed error covariances (Dee 1995).

From equations (10) and (11) we see that the impact that the model error has on the analysis error depends on the Kalman gain matrix. The larger the elements of $\mathbf{K}$ are, the greater the impact of model error, or in other words the more dominant the observations are (due to either their number or their accuracy) in computing the analysis, the greater the impact of model error.

In addition to the non-linear model error present in the outer loop, an error in the TL approximation to the NL model used in the assimilation is also present in the inner-loop of incremental 4D-Var. Let

$$\boldsymbol{\epsilon}^{\mathrm{TL}} = \hat{\mathcal{H}}(\mathbf{x}_0 + \delta\mathbf{x}_0) - \hat{\mathcal{H}}(\mathbf{x}_0) - \hat{\mathbf{H}}\delta\mathbf{x}_0 \quad (12)$$

be the tangent linear model error in observation space. For each of the coupling strategies the choices of $\hat{\mathcal{H}}$ and $\hat{\mathbf{H}}$ differ and so the error in the generalized observation operator and the TL model error is different in each case:

Strongly coupled:

$$\boldsymbol{\epsilon}^{\hat{\mathcal{H}}} = \hat{\mathcal{H}}^{\mathrm{t}}(\mathbf{x}_0^{\mathrm{t}}) - \hat{\mathcal{H}}^{\mathrm{c}}(\mathbf{x}_0^{\mathrm{t}}) \quad (13\mathrm{a})$$
$$\boldsymbol{\epsilon}^{\mathrm{TL}} = \hat{\mathcal{H}}^{\mathrm{c}}(\mathbf{x}_0^{\mathrm{g}} + \delta\mathbf{x}_0) - \hat{\mathcal{H}}^{\mathrm{c}}(\mathbf{x}_0^{\mathrm{g}}) - \hat{\mathbf{H}}^{\mathrm{c}}\delta\mathbf{x}_0 \quad (13\mathrm{b})$$

Weakly coupled:

$$\boldsymbol{\epsilon}^{\hat{\mathcal{H}}} = \hat{\mathcal{H}}^{\mathrm{t}}(\mathbf{x}_0^{\mathrm{t}}) - \hat{\mathcal{H}}^{\mathrm{c}}(\mathbf{x}_0^{\mathrm{t}}) \quad (14\mathrm{a})$$
$$\boldsymbol{\epsilon}^{\mathrm{TL}} = \hat{\mathcal{H}}^{\mathrm{c}}(\mathbf{x}_0^{\mathrm{g}} + \delta\mathbf{x}_0) - \hat{\mathcal{H}}^{\mathrm{c}}(\mathbf{x}_0^{\mathrm{g}}) - \hat{\mathbf{H}}^{\mathrm{uc}}\delta\mathbf{x}_0 \quad (14\mathrm{b})$$

Uncoupled:

$$\boldsymbol{\epsilon}^{\hat{\mathcal{H}}} = \hat{\mathcal{H}}^{\mathrm{t}}(\mathbf{x}_0^{\mathrm{t}}) - \hat{\mathcal{H}}^{\mathrm{uc}}(\mathbf{x}_0^{\mathrm{t}}) \quad (15\mathrm{a})$$
$$\boldsymbol{\epsilon}^{\mathrm{TL}} = \hat{\mathcal{H}}^{\mathrm{uc}}(\mathbf{x}_0^{\mathrm{g}} + \delta\mathbf{x}_0) - \hat{\mathcal{H}}^{\mathrm{uc}}(\mathbf{x}_0^{\mathrm{g}}) - \hat{\mathbf{H}}^{\mathrm{uc}}\delta\mathbf{x}_0 \quad (15\mathrm{b})$$

The superscripts c and uc refer to the coupled and uncoupled assimilation models. We note that the truth, $\hat{\mathcal{H}}^{\mathrm{t}}(\mathbf{x}_0^{\mathrm{t}})$, is always coupled by definition. From these equations it is clear that $\boldsymbol{\epsilon}^{\hat{\mathcal{H}}}$ is the same for the weakly and strongly coupled formulations due to the outer-loop calculations being the same. For the strongly coupled and uncoupled methods the exact tangent

linear of the erroneous nonlinear model is used in the inner loop. In these cases the incremental 4D-Var scheme is equivalent to a truncated Gauss-Newton iteration and, provided that the inner loop is solved to sufficient accuracy, the outer loop iterates should converge to a minimum of the corresponding discrete cost function and the TL error should tend to zero. On the other hand, for the weakly coupled assimilation the uncoupled tangent linear models only provide an approximation to the true linearization of the coupled model used in the outer loop. In this case the incremental 4D-Var scheme is a perturbed Gauss-Newton method and, under certain conditions, will converge to a solution close to the minimum of the discrete nonlinear cost function, but not equal to it (Lawless et al. 2005; Gratton et al. 2007). In practice, however, the tangent linear model always contains some approximations, since it is usually run at a lower resolution than the outer loop nonlinear model and may not include the linearization of all physical parameterizations.

There are also differences in the way the linearization trajectory is defined. For the weakly coupled case, $\hat{\mathbf{H}}^{\mathrm{uc}}$ is linearized about the coupled trajectory given by $\hat{\mathcal{H}}^{\mathrm{c}}(\mathbf{x}_0^{\mathrm{g}})$ and the boundary conditions at the air-sea interface (BCs) used to force $\hat{\mathbf{H}}^{\mathrm{uc}}$ are calculated using the NL coupled model. For the uncoupled case, $\hat{\mathbf{H}}^{\mathrm{uc}}$ is linearized about $\hat{\mathcal{H}}^{\mathrm{uc}}(\mathbf{x}_0^{\mathrm{g}})$. The BCs for the uncoupled model runs, $\hat{\mathcal{H}}^{\mathrm{uc}}(\mathbf{x}_0^{\mathrm{g}})$ and $\hat{\mathbf{H}}^{\mathrm{uc}}$, are prescribed externally.

4   Experimental design

Following on from Smith et al. (2015) we make use of a single column model of the coupled atmosphere and ocean. We aim to create an experimental set up in which we believe the model error to have characteristics of the error seen in an operational coupled atmosphere-ocean model. That is, we wish for the growth rates of the model error in the atmosphere to be much larger than in the ocean and to be complex in nature. Below details of the 'true' and erroneous model are given. We assume the error to originate from missing physics, erroneous parameter values and errors in the large scale forcing.

4.1   'Truth' model

The model which represents the truth in our idealized experiments comprises of the ECMWF single-column model (SCM), which originates from an early cycle of the IFS (Integrated Forecasting System) code, coupled to a single-column ocean mixed layer model. A brief description of the key features of the model follows. A more complete description of the dynamical core equations and discretization is given in Smith et al. (2015).

4.1.1   The atmosphere

The atmospheric component of the coupled model solves the primitive equations for prognostic variables, temperature, T, specific humidity, q, and ageostrophic zonal, u, and meridional, v, wind. The model is forced externally by horizontal advection for each of the prognostic variables and by the geostrophic component of the winds. Tendencies due to sub-grid scale physical processes are also included to represent the effect of radiation, turbulent mixing, moist convection and clouds.

The vertical discretization of the equations for the atmosphere component uses the hybrid vertical co-ordinate scheme developed by Simmons

and Burridge (1981) to describe the atmosphere on 60 model levels. This allows for greater resolution in the planetary boundary layer (maximum resolution is $\approx$15m), with decreasing resolution towards the top of the model domain (minimum resolution is $\approx$4km) at 0.1 hPa.

4.1.2 The ocean

The ocean mixed layer model is based on the K-Profile Parametrization (KPP) vertical mixing scheme of Large et al. (1994). The code was originally developed by the NCAS Centre for Global Atmospheric Modeling at the University of Reading (Woolnough et al. 2007) and incorporated into the ECMWF SCM code by Takaya et al. (2010).

The prognostic variables in the ocean are the mean values of temperature, θ, salinity, s, and zonal, u_o and meridional, v_o currents. The KPP model describes mixing in the boundary layer near to the surface and mixing in the interior ocean. This includes the effects of shear instability, internal wave breaking in the interior of the ocean and double diffusion. The model is forced by solar irradiance at the upper boundary and by externally specified geostrophic currents.

The ocean model uses a stretched vertical grid of Takaya et al. (2010) with 35 levels from the surface to a depth of 250m. The resolution is increased in the upper layers in order to simulate the diurnal SST variability; the top model layer is chosen to be 1m thick and there are 19 levels in the top 25m. Some examples of the model level heights/pressures for the atmosphere and ocean are given in tables 1 and 2 respectively.

model level	model full pressure level (hPa)[1]
17	18.815
22	54.624
25	95.980
30	202.230
33	288.093
39	501.637
49	861.497
56	995.055

Table 1: Locations of the reduced observations in the atmosphere used in experiments in section 55.4.

model level	depth (m)
1	1.000
5	5.277
10	11.406
16	20.173
20	28.100
23	37.366
25	46.985
27	61.498
29	83.818
31	118.214
33	170.778
35	250.00

Table 2: Reduced observation locations in the ocean used in experiments in section 55.4.

4.1.3 Atmosphere-Ocean coupling

Coupling of the atmosphere and ocean components of the model occurs at every time step (15 minutes) via the exchange of latent and sensible heat fluxes and surface momentum flux from the atmosphere to the ocean. The updated ocean model sea-surface temperature is passed back to the atmosphere where it is then used in the computation of the atmosphere lower boundary conditions for the next step. Fluxes are estimated from bulk formulae, with the method of Louis et al. (1981) used to calculate the transfer coefficients.

4.2 Assimilation and forecast models

The coupled model used for the coupled DA experiments and to produce the coupled forecasts is similar to that used in Smith et al. (2015). In comparison with the truth model this has missing physics in the atmosphere, representing just advection, vertical diffusion and turbulent mixing. It also has a positive bias in the large scale forcing of the horizontal advection terms for the atmosphere. In the ocean, perturbed parameters for the diffusion parameters are used (details of which are given in table 3) and there is no non-local mixing. The parameters that are perturbed each affect the mixing in the erroneous model. However their combined effect is minimal compared to the errors propagating down from the air-sea interface.

The uncoupled models used by the uncoupled DA experiments are the same as those used for the coupled forecast with the exception that there is no exchange in information between the two components. Instead the surface fluxes needed to force the ocean component and the SSTs needed to force the atmosphere component

are prescribed externally. In this study we consider two options:

1. *poor BCs*: BCs given by ERA interim (Dee et al. 2011) and Mercator (Lellouche et al. 2013) reanalyses products. These products are inevitably inconsistent with the idealized model used in these experiments. In addition to this the SSTs obtained from the Mercator product have no diurnal cycle, with only a daily averaged value provided.

2. *good BCs*: BCs given by output from the truth model, $\mathcal{M}^{t}(\mathbf{x}_0^t)$. The output is prescribed at every 6 hours with linear interpolation to provide BCs at intermediate times.

4.3 Illustration of model error

The model error that results from this set up is illustrated for a case study relevant for July 2014 for a point in the NW Pacific (188.75°E, 25°N). The initial conditions are obtained by running the truth model for 1 day initialized by data taken from ERA Interim Re-analysis for the atmosphere and Mercator Ocean reanalysis for the ocean valid at 00:00UTC on the 2nd July 2014. Obtaining the initial conditions in this way ensures that they lie on the true model 'attractor'. Forcing fields are also calculated from these reanalysis products, specified 6 hourly throughout the forecast period, with linear interpolation between these times. The true evolution of the atmospheric and oceanic temperature over an integration time of 4 days is shown in figure 2.

In figures 3 to 6 the NL and TL model error for temperature in the atmosphere and ocean are shown, computed using equations (13) to (15). If we assume that the entire state is observed at every time step then the NL model error is

name	description	true value	value used in assimilation and forecast
RRINFTY	Critical Richardson number for shear instability	0.8	0.7
RDIFM_{IW}	background/internal waves viscosity(m^2/s)	1.5×10^{-4}	1.0×10^{-5}
RDIFS_{IW}	background/internal waves diffusivity(m^2/s)	1.5×10^{-5}	1.0×10^{-5}
RDIFM_{MAX}	max viscosity due to shear instability (m^2/s)	5×10^{-3}	1.0×10^{-3}
RDIFS_{MAX}	max diffusivity due to shear instability (m^2/s)	5×10^{-3}	1.0×10^{-3}

Table 3: Parameters modified to create model error in the ocean component of the coupled model.

equivalent to $\epsilon^{\hat{\mathcal{H}}}$ and the TL error is equivalent to ϵ^{TL}. In this case the TL error has been computed for a perturbation equal to the truth minus the background to be used in the assimilation experiments presented in section 5.

In figures 3 and 4 we see that in the atmosphere the NL and TL model error are fairly insensitive to the coupling strategy above the boundary layer (level 50 and above) and within the boundary layer only small differences can be seen. This suggests that for our model set up the lower boundary is not a great source of error. This could also be expected in general as the ocean acts as a 'slave' to the atmosphere and so changes to the ocean will not have an immediate effect on the atmosphere at short time-scales. In each case we see a large warm bias forming in the NL model between levels 40-50, which corresponds to approximately 1.2-4.7km. There is also a cool bias developing at level 30 which corresponds to approximately 12km, roughly the top of the troposphere. Compared to the NL model error the TL model error is small (approximately 20% of its value) and should reduce throughout the minimization procedure.

Figures 5 and 6 show that the behavior of the model error in the ocean is much more sensi-tive to the upper boundary, suggesting that the atmosphere is a large source of error over the two day forecast window. For the coupled NL model (used in the strongly and weakly coupled scheme) we see that there is a cool bias at the surface, peaking at the diurnal maximum (roughly 24 hours and 48 hours into the forecast), overlying a warm bias. This suggests that in the coupled model the heat originating from the atmosphere is being mixed down too quickly. This is seen to result in a large warm bias at level 20 (approximately 25m) after 1 day, corresponding to the thermocline being deepened too quickly. This hypothesis is consistent with the error in the lower winds which are also seen to have a positive bias (not shown) and are therefore causing too much momentum to be passed to the ocean and hence too much mixing in the upper ocean. Experiments (not shown) in which the uncoupled ocean model is run with the true heat fluxes but erroneous momentum fluxes (either computed from the coupled erroneous model or prescribed externally from the ERA interim product) show that the errors in the momentum fluxes explain a large part of the errors in the NL ocean model component seen in figure 5.
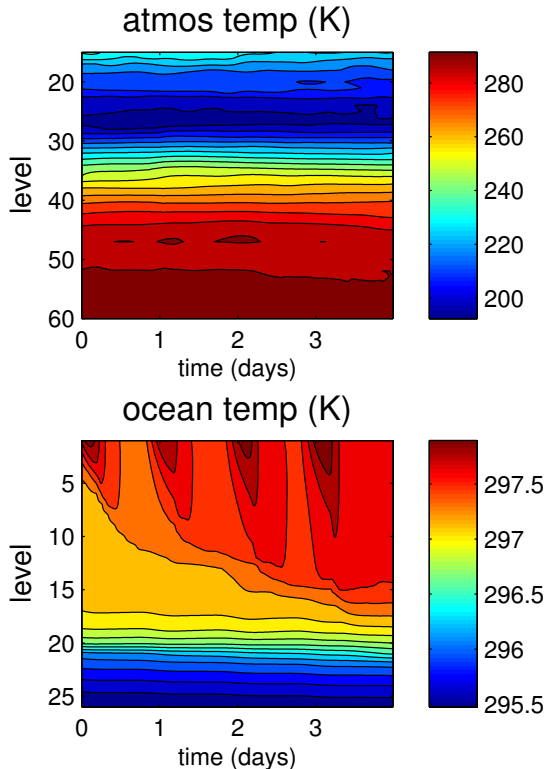
atmos temp (K)

ocean temp (K)

Figure 2: The simulated true evolution of temperature in the atmosphere (top) and ocean (bottom) over an integration time 4 days.

notably there is an underestimation of the mixing of the heat into the ocean caused by an underestimation of the prescribed surface momentum flux (not shown), whereas in the coupled model there is an overestimation of the mixing as discussed above.

In contrast to the atmosphere, the TL error in the ocean (figure 6) is of comparable magnitude to the error in the NL model (figure 5). However, both are substantially smaller than the errors seen in the atmosphere.

In both the atmosphere and ocean it is interesting to note that the error in the approximate TL used by the weakly coupled formulation is only slightly larger than using the exact TL in the strongly coupled formulation. The main difference in the atmosphere occurs in the lowest 10 levels (corresponding to the BL) where the errors are seen to be maintained longer into the simulation. In the ocean the greatest differences are at around level 20 (corresponding to the thermocline). Although these difference are small, we will see in section 55.4 that in some circumstances they can lead to differences in the balance between the ocean and atmosphere analyses.

4.4 Assimilation experiment design

Within this section the details of the assimilation experiments setup are given. The aim of the experiments is to study the effect of the model error on the assimilation of observations. We concentrate on the case study presented above so that the link between the resulting analysis and the realization of the model error is explicit. Experiments using data from a June 2013 case study for the same location gave qualitatively similar results so are not included but give us the confidence that our results are robust.

For the uncoupled model, figures 5 and 6 show the error in the ocean to be very sensitive to the prescribed BCs. In particular we note that when the error in the BCs is negligible (last panel) the NL and TL model error is substantially reduced. When poor BCs are used, the error in the oceanic temperature may be larger than in the coupled model, but because the errors are not dynamically coupled to the atmosphere they are of a completely different nature to the errors in the ocean component of the coupled model. Most

The experimental design is essentially a biased twin experiment, in which the truth is known (see section 44.1) and observations are made directly from this known truth. A biased model (see section 44.2) is then used to assimilate these observations.

In the following experiments the background error covariance matrix $\mathbf{B}$ is assumed to be diagonal. This is a large simplification but allows for a clean comparison between the three different coupling strategies introduced in section 2. The variances of the background error are estimated from the variance in a time-series of model output as described in Smith et al. (2015). The background is then computed by generating white noise with the background error variances, adding it to the true profile at 24 hours prior to the initial time of interest and running the coupled forecast model forward 24 hours. This ensures that the background profile lies on the coupled model attractor. As the errors in the background will have grown throughout the forecast the background error variances are then inflated so that they are consistent with the errors in the background profile. The background error standard deviations are shown in figure 7.

In the following experiments observations are made of the truth at every model level at 3 hourly intervals in the atmosphere and at 6 hourly intervals in the ocean. The spatial density of the observations has been chosen to accentuate the effect of the model error, which as seen in (10) and (11) is greatest when the observations play a dominant role in calculating the analysis, implying that the effect of model error will be greater with denser observations, especially when coinciding with the region of large model error. The frequency of the observations has been chosen to mimic the reduced availability of observations in the ocean compared to the

atmosphere.

The observations, simulated from the truth model, are consistent with a prescribed error variance which is assumed to be known exactly in the assimilation. The observation error standard deviations, the values of which are plotted in figure 7, are constant for each variable and independent of height. We note that humidity is not observed.

The inner loop is stopped when the relative change in the gradient is less than 0.001. The number of iterations needed depends on the assimilation scheme used and the window length. In general more iterations are necessary with the strongly coupled scheme. To improve convergence a simple preconditioning of the control vector is used. Instead of minimizing (5) with respect to $\delta\mathbf{x}_0$ it is instead minimized with respect to a transformed variable equal to $\mathbf{B}^{-1/2}\delta\mathbf{x}_0$. Such a transformation is commonly used in operational data assimilation (see, for example, Bannister (2008) and references therein). In additional tests (not shown) it has been found that for all DA strategies three outer loop iterations are sufficient for convergence, however results will also be shown for the weakly coupled scheme in which only one outer loop is performed.

5 Assimilation results

5.1 Sensitivity to window length

It can be expected that the assimilation results will be sensitive to the assimilation window length. If no model error is present and the TL error remains small then increasing the window length can be expected to reduce the analysis error as more observations become available for assimilation. However, in the presence of model error the analysis error will increase in ac-

cordance with (10) and (11) as the model error grows throughout the window and the greater number of observations accentuates its effect. In figures 8 and 9 the absolute analysis errors (computed from the difference from the true state) at the initial time for the different coupling strategies are given for assimilation window lengths of 6 and 48 hours.

In figure 8 we see an increase in the analysis error of atmospheric temperature for all coupling strategies as the window length increases (note the change in the x-axis scale). This is most noticeable at levels 45 and 25-30 where the analysis becomes significantly poorer than the background (gray line). These levels coincide with the large biases seen in the assimilation model which develop after approximately 12 hours (see figure 3).

In figure 9 we see that for the analysis of oceanic temperature the error does not increase in the same way as the window length is increased. Instead it appears in places that the analysis error reduces, particularly between levels 6 to 10, with the uncoupled DA schemes (red lines) showing the greatest reduction in error. This is consistent with the longer window length allowing for the observations to provide more information about the true ICs and the diurnal cycle of the evolution of the mixed layer. The fact that the uncoupled analyses (especially with the good BCs) outperforms the coupled analyses was expected from figure 5. The opposite is true closer to the surface at around level 2 despite the improvement in temperature at level 1. This could be indicative of the inability of the uncoupled methods to utilize information from the atmospheric observations to constrain the ocean analysis.

It is interesting to note the difference between the analyses using the weakly coupled scheme when 1 and 3 outer loops are performed (green dashed-dot and green dashed lines respectively). We see that for both window lengths an improved analysis is given when only one outer loop is performed around levels 6-9. The reason for this could be due to the fact that the model error due to the coupling is only experienced in the weakly coupled scheme during the outer loop update, and so more outer loops implies that the coupled model has a greater influence on the analysis (see Smith et al. (2015)). Therefore performing fewer outer loops reduces the effect of the model error allowing for a more accurate analysis.

We also observe differences between the coupled and uncoupled analyses at level 18, just above the thermocline, where the coupled analyses are more accurate than the uncoupled. From figure 5 we see that the model error in this region propagates down from the surface. Therefore the larger analysis error seen in the coupled analyses at the surface down to level 10, for a 48 hour window, may in fact be correcting for the error deeper down and hence allowing for the model prediction in this region to become more consistent with the observations and the information in the observations to be interpreted correctly. No scheme is able to correct the large error at the thermocline due to the relatively small background error variances in this region (caused by the fact that there was little variation in this feature in the forecast used to estimate the background error variances, see section 44.4).

5.2 Forecast error when initialized from the analyses

The potential for the coupled DA scheme to produce a poorer ocean analysis in the presence of model error has been demonstrated. However,

often the aim of data assimilation is not to produce the most accurate initial conditions but the most accurate forecast. In this section we look at the error in the forecast produced using the coupled erroneous model when initialized by the different analyses computed with the 48hr assimilation window in the previous section. Figure 10 shows the error in the forecast of atmospheric temperature. The top two panels show the error in the forecast using the erroneous model when initialized with the true ICs (left) and the background (right). The other panels show the error in the forecast when initialized with the different coupling DA strategies (due to the similarity between the weakly coupled results when 1 and 3 outer loops are performed, only the results with 3 outer loops are shown). It can be seen that the error in the analysis at the initial time (seen in figure 8 in which the error in the analyses was larger than for the background between levels 25 and 30 and at level 45 in the case of a 48 hour window length) has helped to restrict the growth of the warm bias around level 45 and the cool bias around level 30 in the forecast initialized by the analyses at later times, compared with using the true ICs. This is because, in order to minimize the cost function an analysis was found which gave a good fit to the observations throughout the 48hr assimilation window and not just at the beginning of the window. This effect of model error in variational data assimilation has been noted previously in the work of Wergen (1992) and Griffith and Nichols (2000).

In figure 11 the forecast error for the atmospheric variables is summarized by the root mean square error (RMSE) averaged over all atmospheric levels, plotted as a function of time. The reduction in the forecast error for temperature initialized using the analyses (colored lines) compared to the true ICs (black line) is clear after 12 hours into the assimilation window and is maintained throughout the 48 hour assimilation widow and the following 2 day forecast. The forecast error is also seen to be mostly reduced for the wind fields, which displays an inertial oscillation in magnitude. Note that humidity is unobserved which explains why the forecast error is not reduced when initialized with the analyses compared to initialing with the truth.

In figure 12 the error in the forecast of oceanic temperature when initialized by the different states is shown. The reduction in the forecast error when initialized using the different analyses compared with initialization from the truth is not as clear as for the atmosphere. However, the forecasts initialized using the coupled DA strategies (middle row) do result in an improved forecast in the region of the thermocline compared to the forecast initialized with the true state, which is particularly noticeable beyond 2 days. This is because, although using the coupled non-linear model in the assimilation resulted in a larger analysis error in some regions of the ocean, it is consistent with the coupled model used to produce the forecast. Therefore, as in figure 10, errors at the initial time have helped to restrict the growth of the errors due to the imperfect model. It is interesting to note that this is also the case for the weakly coupled analysis with only 1 outer loop which was seen to give a more accurate analysis of the ocean temperature than when more outer loops were performed (see figure 9). We can therefore speculate that in this case even with just one outer loop the weakly coupled scheme, by linearizing around the coupled trajectory, has allowed for an analysis consistent with the coupled model. The improvement seen in the error in the thermocline suggests that the assimilation has reduced the amount of heat being mixed down from the

surface. From figure 13 we can see that this is not due to a substantial improvement in the momentum fluxes. In fact the fluxes initialized by the strongly coupled analysis are worse in many places, and may support the idea that surface heat fluxes are compensating for errors elsewhere in the coupled model, as found in the studies of de Szoeke and Xie (2008) and Zheng et al. (2011).

The results presented in figure 12 clearly demonstrate the advantage of using the coupled scheme over the uncoupled scheme in reducing the error in the ocean forecast. Despite the model error it is important that the assimilation and the forecast models are consistent with one another. In the next section we show that this is particularly true when forecasting using the coupled system, as it is essential not only that the analysis allows for a good fit to the observations but also that the atmosphere and ocean analyses are in balance with one another.

5.3 Initialization shock

A benefit of coupled DA in the absence of model error is its ability to reduce the occurrence of initialization shock by updating the atmosphere and ocean as a coherent system and in turn find an analysis which lies on the forecast model attractor. In previous work (Smith et al. 2015) initialization shock was found to be evident in the first few hours of forecast of sea surface temperature (SST). We now examine whether coupled assimilation can still reduce the shock when model error is present and so the true attractor and model attractor differ.

In figure 14, 72-hour and 3-hour forecasts of the sea surface temperature (SST) are given using the coupled model initialized using the different assimilation schemes with a 48hr assimilation window. It is clear that although the uncoupled analyses are closer to the true SST at the initial time they quickly deviate from the truth and have a much less realistic forecast during the first hour than the coupled analyses. Beyond the first few hours the forecasts initialized using the uncoupled analyses continue to be poorer than the forecasts initialized using the coupled analyses. We can therefore conclude that, in this case, the presence of model error at the atmosphere-ocean interface does not adversely affect the ability of the coupled DA to produce a state consistent with the coupled forecast model. The fact that the forecasts initialized with the coupled analyses still have a large cool bias in the SST, comparative to the bias when initialized by the uncoupled analyses, is most likely related to the reduction in the warm bias seen in the thermocline see figure 12.

5.4 Effect of strength of coupling in 'weakly coupled' DA

The strength of the coupling in weakly coupled DA is controlled by the number of outer loops performed and the resolution of the observations (Smith et al. 2015). In the previous experiment (using 3 outer loops and dense observations) the weakly coupled scheme is seen to perform in a similar way to the strongly coupled scheme.

Within this section we wish to understand how the model error affects the analysis when the strength of coupling in weakly coupled DA is reduced. Unfortunately the number of outer loops and the resolution of the observations also have a significant impact on the analysis in other ways too, making it difficult to perform a clean experiment showing just the effect of the reduced coupling. For example reducing the number of outer loops reduces the ability to find the mini-

mum of the cost function and reducing the number of observations means that the effect of the observations and model error on the analysis is reduced (see (10) and (11)), so that there will be less deviation from the background no matter which assimilation scheme is used.

Given these caveats the assimilation experiments are now repeated using sparser observation (both temporally and spatially). In the atmosphere the frequency of the observations is matched to the 6 hour frequency of the oceans (the frequency of the ocean observations remains at 6 hourly) and in both systems the vertical resolution of the observations is reduced to the levels given in tables 1 and 2. With this set up little difference was seen in the results between using 1 or 3 outer loops in the weakly coupled formulation. This is because reducing the observations means that the effect of the outer loop update of the innovations is reduced. Therefore all the results that follow use 3 outer loops.

In figures 15 and 16 we see that the analysis error is more comparable to the denser observation case with a 48 hour window (see figures 8 and 9) if we increase the window length to 96 hours. In particular we see that we still get a large spike in the analysis error at level 45 in the atmospheric temperature and we see a reduction in the analysis error (compared to the background error) when the uncoupled methods are used around levels 5-10 in the oceanic temperature. This is because we are still assimilating a similar number of observations so that the effect of the observations and model error is still significant allowing for the analysis to divert from the background.

We expect this setup to reduce the strength of the coupling within the weakly coupled scheme because although observations are available for a longer period of time the spatial and temporal frequency of the observations is reduced and so there is less information in observation space about the coupling between the atmosphere and ocean. In figure 17 this is illustrated by again looking at the initialized forecast of SST. Compared to figure 14, we see that the initialization shock has increased for both the strongly and weakly coupled schemes, but the difference is much greater for the weakly coupled scheme, with the effect of the initial imbalance seen beyond the first hour of the forecast.

These experiments illustrate the potential risks of the weakly coupled scheme in the presence of significant model error. As it is in the outer loop that the observations and model are compared, the discrepancy due to the coupled model error will be similar for both the strongly and weakly coupled formulations. However, because the uncoupled TL models used in the inner loop are inconsistent with the coupled NL model used in the outer loop, the weakly coupled scheme is unable to find an analysis increment which allows for an agreement between the observations and the model as successfully as the strongly coupled scheme if there is not enough information about the coupling from the observations. This means that not only can we expect a poorer analysis with the weakly coupled scheme but also a larger forecast error and a greater initialization shock.

6 Conclusions and discussion

There is strong motivation for the development of coupled DA methods, namely their ability to produce a more balanced coupled analysis state and to make better use of near surface observations. However, a limiting factor in the implementation of coupled DA is the model error in the atmosphere which restricts the length of the

18

assimilation window that can be used to something shorter than in an uncoupled ocean only scheme. Within this study we have aimed to give insight into the effect of lengthening the window when model error is present to see if the benefits of coupled DA are still evident. A summary of our key findings follows.

The effect of the model error in coupled DA depends not only on the nature of the error in the coupled model but also on the coupling strategy used within the DA scheme. It is possible for errors in the coupled system to introduce an error in the ocean component near to the surface which has faster timescales than in the uncoupled ocean model (as seen in figure 5 comparing the coupled model error to the uncoupled model error with good BCs). This new source of error means that the accuracy of the ocean analysis may be degraded using a coupled scheme compared to using an uncoupled DA scheme which (in the absence of this fast error growth) is able to utilize a longer window length . This was shown to be evident for a case study in which we found the errors in the analysis of the ocean to be smaller in some regions when using an uncoupled scheme and a 48 hour window (figure 9).

The clear problem with an uncoupled scheme, however, is that despite allowing a smaller error in the analysis of the ocean, the atmospheric and oceanic analyses are inconsistent with the forecast model. This means that the error growth rate in the coupled model forecast may actually be larger when initialized using the uncoupled analyses and an initialization shock may become apparent in the forecast of the SSTs (see figure 14).

With dense observations and 3 outer loops, the weakly coupled scheme was seen to perform in a very similar manner to the strongly coupled scheme, responding in a similar way as the assimilation window was increased (figure 9) and reducing the error in the forecast by a similar degree (figure 12). When the number of outer loops was reduced to one and observations were dense, it was seen that weakly coupled scheme gave a better analysis than the strongly coupled scheme in some regions due to the reduction of the impact of the coupled error in the outer loop calculation of the innovations, although the coupling was still strong enough to allow for the analysis to be consistent with the coupled model and so the forecast was better than that initialised from the uncoupled analyses. For example, the error in the thermocline was reduced (figure 12) and initialisation shock was smaller (figure 14). However, if the density of the observations is reduced then the strength of coupling is also seen to reduce, but instead of tending towards the uncoupled scheme (as is the case when no model error is present (Smith et al. 2015)) it can give a much poorer analysis and forecast than both the uncoupled and strongly coupled schemes. This is because, although the information in the observations about the coupling is reduced, the inconsistency between the observations and the NL coupled model seen in the outer loop remains. Therefore, unlike with the strongly coupled scheme which used the coupled model within both the inner and outer loop, the weakly coupled scheme is unable to find an update to the background state which allows for the initialized model to become more consistent with the observations. Therefore the problem of model error is likely to be much more problematic in the weakly coupled schemes, which are currently being implemented at centers such as the UK Met Office and ECMWF, than in a strongly coupled scheme. To address this issue, the implementation of coupled DA has been

forced to use a short assimilation window length (6 hours at the Met Office and 24 hours at the ECMWF).

To conclude, the benefits of a coupled DA scheme are still evident even in the presence of model error. However, if the aim is to find the best analysis then it is important to choose an assimilation window length in which the model error remains negligible. In practice this means choosing a window length consistent with an atmosphere only assimilation which will severely limit the number of ocean observations available for assimilation. If the purpose is to give an improved forecast then using a longer window length may help to reduce the model error growth (particularly in the observed fields) by finding the initial conditions which limit it. However, in order to use coupled 4D-Var to its full potential it is necessary to take into account model error in the assimilation, allowing for the window length and number of ocean observations available for assimilation to be increased and theoretically a more accurate analysis to be found. Allowing for model error in variational data assimilation greatly increases the complexity of the data assimilation problem and is an active area of research (e.g. Fisher et al. (2011), Moore et al. (2011)).

One method, known as weak constraint 4D-Var, aims to estimate the model error along with the initial conditions (Griffith and Nichols 2000; Trémolet 2006). In theory this allows for the model to be corrected using the observations. However, in practice it is very difficult to obtain accurate results from weak constraint 4D-Var, as it is necessary to have a good understanding of the elusive model error statistics (see Todling (2014) and references therein). This is particularly challenging in coupled data assimilation as the model error statistics do not only need to

be specified for the ocean and atmosphere but an understanding of the cross-correlations is also needed for strongly coupled DA. One possibility is to only estimate the error in the atmospheric component assuming that the error in the ocean is comparatively negligible and has its origins in the atmosphere for the timescale of the assimilation window.

An alternative method is to use parameter estimation as well as initial state estimation, essentially tuning the model parameters to give a better fit to the observations via the assimilation. Kondrashov et al. (2008) argue that systematic errors in many tropical ocean-atmosphere general circulation models are caused by incorrect parameter values. Even if the model error is not entirely due to erroneous parameter values, parameter estimation can be used to reduce model error if the model is sensitive to the parameters and the observations are sufficient (Navon 1997). This idea was used by Liu et al. (2014) with the Fast Ocean Atmosphere Model (FOAM) and a coupled ensemble adjustment Kalman filter (Anderson 2001). They successfully estimated the solar penetration depth (SPD) in twin experiments. SPD is thought to be a parameter that may have significant impact on the surface climate (see Liu et al. (2014) for references). They also tried to estimate two additional parameters related to the momentum and latent heat fluxes. This was less successful due to the nonlinear relationships between the parameters and state variables weakening the correlations between forecast error and parameter uncertainty. Estimation of the bulk adjustment factors was also performed by Mochizuki et al. (2009) using a 4D-Var technique. This was found to reduce model biases in climatological fields.

With both the weak constraint and parameter estimation methods it is unclear if the estimates

20

of respectively model error and the parameter values should be used in the subsequent forecast. If they are not used then the models used in the assimilation and the forecast will be inconsistent, and so, in a similar way to the uncoupled DA scheme, the analyses will not lie on the forecast model attractor. Hence the forecast error growth may be large even if the analysis error is small.

A   Derivation of analysis error covariance and bias in the presence of model error

The analysis was given in (9) in terms of the background, $\mathbf{x}^{\mathrm{b}}$, the observations over the assimilation window , $\hat{\mathbf{y}}$, the generalized non-linear observation operator, $\hat{\mathcal{H}}$ and the Kalman gain matrix, $\mathbf{K}$. In the presence of model error the mapping $\hat{\mathcal{H}}$ is erroneous as described by (8). The Kalman gain matrix is therefore no longer optimal and this has an impact on the analysis error. Let $\tilde{\epsilon}^{\mathrm{a}}$ be the analysis error when there is no error present in $\hat{\mathcal{H}}$ (i.e. $\hat{\mathcal{H}} = \hat{\mathcal{H}}^{\mathrm{t}}$) and ϵ^{a} be the analysis error when there is error present.

$$
\begin{aligned}
\epsilon^{\mathrm{a}} \quad &= \mathbf{x}^{\mathrm{a}} - \mathbf{x}_0^{\mathrm{t}} \\
&= \qquad\quad \mathbf{x}^{\mathrm{b}} - \mathbf{x}_0^{\mathrm{t}} + \mathbf{K}(\hat{\mathbf{y}} - \hat{\mathcal{H}}(\mathbf{x}^{\mathrm{b}})) \\
&= \qquad\quad \mathbf{x}^{\mathrm{b}} - \mathbf{x}_0^{\mathrm{t}} \\
&\qquad + \mathbf{K}(\hat{\mathbf{y}} - \hat{\mathcal{H}}^{\mathrm{t}}(\mathbf{x}_0^{\mathrm{t}}) + \hat{\mathcal{H}}^{\mathrm{t}}(\mathbf{x}_0^{\mathrm{t}}) \\
&\qquad - \hat{\mathcal{H}}(\mathbf{x}_0^{\mathrm{t}}) + \hat{\mathcal{H}}(\mathbf{x}_0^{\mathrm{t}}) - \hat{\mathcal{H}}(\mathbf{x}^{\mathrm{b}})) \\
&= \qquad\quad \epsilon^{\mathrm{b}} + \mathbf{K}(\epsilon^{\mathrm{o}} + \epsilon^{\hat{\mathcal{H}}} - \mathbf{H}\epsilon^{\mathrm{b}}) \\
&= \qquad\quad \tilde{\epsilon}^{\mathrm{a}} + \mathbf{K}\epsilon^{\hat{\mathcal{H}}}
\end{aligned}
\tag{16}
$$

If we assume that ϵ^{b} and ϵ^{o} are unbiased then $\tilde{\epsilon}^{\mathrm{a}}$ is also unbiased and the expected value of ϵ^{a} is

$$
E[\epsilon^{\mathrm{a}}] = \mathbf{K}E[\epsilon^{\hat{\mathcal{H}}}] \tag{17}
$$

as stated in (11).

Similarly if we assume that ϵ^{b} and ϵ^{o} are uncorrelated with $\epsilon^{\hat{\mathcal{H}}}$, we can compute the analysis error covariance as

$$
\begin{aligned}
&E[(\epsilon^{\mathrm{a}} - E(\epsilon^{\mathrm{a}}))(\epsilon^{\mathrm{a}} - E(\epsilon^{\mathrm{a}}))^{\mathrm{T}}] \\
&= E[(\tilde{\epsilon}^{\mathrm{a}} + \mathbf{K}\epsilon^{\hat{\mathcal{H}}} - \mathbf{K}E[\epsilon^{\hat{\mathcal{H}}}])(\tilde{\epsilon}^{\mathrm{a}} + \mathbf{K}\epsilon^{\hat{\mathcal{H}}} - \mathbf{K}E[\epsilon^{\hat{\mathcal{H}}}])^{\mathrm{T}}] \\
&= E[\tilde{\epsilon}^{\mathrm{a}}(\tilde{\epsilon}^{\mathrm{a}})^{\mathrm{T}}] + \mathbf{K}E[\epsilon^{\hat{\mathcal{H}}} - E[\epsilon^{\hat{\mathcal{H}}}])(\epsilon^{\hat{\mathcal{H}}} - E[\epsilon^{\hat{\mathcal{H}}}])^{\mathrm{T}}]\mathbf{K}^{\mathrm{T}} \\
&= \mathbf{P}^{\mathrm{a}} + \mathbf{K}E[\epsilon^{\hat{\mathcal{H}}} - E[\epsilon^{\hat{\mathcal{H}}}])(\epsilon^{\hat{\mathcal{H}}} - E[\epsilon^{\hat{\mathcal{H}}}])^{\mathrm{T}}]\mathbf{K}^{\mathrm{T}}
\end{aligned}
\tag{18}
$$

where $\mathbf{P}^{\mathrm{a}}$ is the analysis error covariance if no model error were present, as stated in (10).

References

Alves, O., Y. Yin, L. Shi, R. Wedd, D. Hudson, P. Okley, and H. Hendon, 2014: A coupled ensemble ocean data assimilation system for seasonal prediction and its comparison with other state-of-the-art-systems. *EGU General Assembly Conference Abstracts*, **16**, 9487.

Anderson, J. L., 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.

21

Arribas, A., M. Glover, A. Maidens, M. Peterson, M. Gordon, C. MacLachlan, and coauthors, 2011: The GlosSea4 ensemble prediction system for seasonal forecasting. *Mon. Wea. Rev.*, **139**, 1891–1910.

Balmaseda, M., and D. Anderson, 2009: Impact of initialization strategies and observations on seasonal forecast skill. *Geophys. Res. Lett.*, **36**, L01 701.

Bannister, R. N., 2008: A review of forecast error covariance statistics in atmospheric variational data assimilation. ii: Modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, **134 (637)**, 1971–1996.

Barnett, T. P., N. Graham, S. Pazan, W. White, M. Latif, and M. Fl ugel, 1993: ENSO and ENSO-related Predictability. Part I: Prediction of Equatorial Pacific Sea Surface Temperature with a Hybrid Coupled Ocean-Atmosphere Model. *J. Climate*, **6**, 1545–1566, doi:http://dx.doi.org/10.1175/1520-0442(1993)006¡1545:EAERPP¿2.0.CO;2.

Bormann, N., A. Hernandez-Carrascal, R. Borde, H.-J. Lutz, J. A. Otkin, and S. Wanzong, 2014: Atmospheric Motion Vectors from Model Simulations: Part I: Methods and Characterizations as Single-Level Estimates of Wind. *Journal of Applied Meteorology and Climatology*, **53**, 47–64.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Met. Soc.*, **120**, 1367–1387.

de Szoeke, S. P., and S. P. Xie, 2008: The tropical Eastern Pacific seasonal cycle: assessment of errors and mechanisms in IPCC AR4 coupled ocean atmosphere general circulation models. *J. Clim.*, **21**, 392–409.

Dee, D., 1995: Testing the perfect-model assumption in variational data assimilation. *WMO Symposium on Assimilation of Observations in Meteorology and Oceanography, Tokyo.*

Dee, D. P., M. Balmaseda, G. Balsamo, R. Engelen, A. J. Simmons, and J.-N. Thépaut, 2014: Toward a consistent reanalysis of the climate system. *BAMS*, 1235–1248.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, and co authors, 2011: The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Q. J. Roy. Meteorol. Soc.*, **137**, 553–597.

Fisher, M., Y. Trémolet, H. Auvinen, D. Tan, and P. Poli, 2011: Weak constraint and long-window 4D-Var. Technical Memorandum 655, European Centre for Medium-Range Weather Forecasts, Reading UK.

Frolov, S., C. H. Bishop, T. Holt, J. Cummings, and D. Kuhl, 2016: Facilitating strongly coupled ocean-atmosphere data assimilation with an interface solver. *Mon. Wea. Rev.*, **144**, 3–20.

Galanti, E., E. Tziperman, M. Harrison, A. Rosati, and Z. Sirkes, 2003: A study of ENSO prediction using a hybrid coupled model and the adjoint method for data assimilation. *Mon. Wea. Rev.*, **131**, 2748–2764.

Gratton, S., A. S. Lawless, and N. K. Nichols, 2007: Approximate Gauss-Newton methods

for nonlinear least squares problems. *SIAM Journal on Optimization*, **18 (1)**, 106–132.

Griffith, A., and N. Nichols, 2000: Adjoint methods in data assimilation for estimating model error. *Flow, Turbulence and Combustion*, **65**, 469–488.

Han, G., X. Wu, S. Zhang, Z. Liu, and W. Li, 2013: Error covariance estimation for coupled data assimilation using a Lorenz atmosphere and simple pycnocline ocean model. *J. Climate*, **26**, 10 218–10 231.

Jin, E. K., and Coauthors, 2008: Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Climate Dynamics*, **31 (6)**, 647–664, doi:10.1007/s00382-008-0397-3, URL http://dx.doi.org/10.1007/s00382-008-0397-3.

Kondrashov, D., C. Sun, and M. Ghil, 2008: Data assimilation for a coupled ocean-atmosphere model: part II: Parameter estimation. *Mon. Wea. Rev.*, **136**, 5062–5076.

Laloyaux, P., M. Balmaseda, D. Dee, K. Mogensen, and P. Janssen, 2015: A coupled data assimilation system for climate reanalysis. *Q. J. Roy. Meteor. Soc.*, doi:10.1002/qj.2629, URL http://dx.doi.org/10.1002/qj.2629.

Large, W. G., J. C. McWilliams, and S. Doney, 1994: Oceanic vertical mixing: a review and a model with non-local boundary layer parameterization. *Rev. Geophys.*, **32**, 363–403.

Lawless, A., S. Gratton, and N. Nichols, 2005: An investigation of incremental 4D-Var using non-tangent linear models. *Quarterly Journal of the Royal Meteorological Society*, **131 (606)**, 459–476.

Lawless, A. S., 2013: Variational data assimilation for very large environmental problems. *Large scale Inverse Problems: Computational Methods and Applications in the Earth Sciences, Radon series on Computational and Applied Mathematics*, M. J. P. Cullen, M. A. Freitag, S. Kindermann, and R. Scheichl, Eds., De Gruyter, 55–90.

Lea, D., I. Mirouze, M. Martin, R. King, A. Hines, and D. Walters, 2015: Assessing a new coupled data assimilation system based on the Met Office coupled atmosphere, land, ocean, sea ice model . *Mon. Wea. Rev., in press*, doi:10.1175/MWR-D-15-0174.1.

Lellouche, J.-M., O. Le Galloudec, M. Drévillon, C. Régnier, E. Greiner, and co authors, 2013: Evaluation of global monitoring and forecasting systems at Mercator Océan. *Ocean Sci.*, **9**, 57–81.

Liu, Y., Z. Liu, S. Zhang, R. Jacob, F. Lu, X. Rong, and S. wu, 2014: Ensemble-based parameter estimation in a coupled general circulation model. *J. Clim.*, **27**, 7151–7162.

Louis, J. F., M. Tiedtke, and J. F. Geleyn, 1981: A short history of the PBL parameterization at ECMWF. *ECMWF Workshop on Boundary Layer Parameterization. Reading, 25-27 November 1981.*

Lu, J., and W. W. Hsieh, 1998: On determining initial conditions and parameters in a simple coupled atmosphere-ocean model by adjoint data assimilation. *Tellus A*, **50A**, 534–544.

MacLachlan, C., A. Arribas, K. A. Peterson, A. Maidens, D. Fereday, A. A. Scaife, and coauthors, 2015: Global Seasonal Forecast System version 5 (GloSea5): a high resolution

seasonal forecast system. *Q. J. Roy. Meteor. Soc.*, **141**, 1072–1084.

Madec, G., 2008: NEMO ocean engine. *Note du Pole de Modalisation, Institut Pierre-Simon Laplace (IPSL): Paris.*

Magnusson, L., M. Balmaseda, S. Corti, F. Molteni, and T. Stockdale, 2013: Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors. *Clim. Dyn.*, **41**, 2393–2409.

Mochizuki, T., S. Nozomi, T. Awaji, and T. Toyoda, 2009: Seasonal climate modeling over the Indian Ocean by employing a 4D-VAR coupled data assimilation approach. *J. Geophys. Res.*, **114**.

Molteni, F., T. Stockdale, M. Balmaseda, G. Balsamo, R. Buizza, L. Ferranti, and coauthors, 2011: The new ECMWF seasonal forecast system (System 4). *European Centre for Medium-range Weather Forecasting Technical Memorandum.*, **656**.

Moore, A. M., H. G. Arango, G. Broquet, B. S. Powell, A. T. Weaver, and J. Zavala-Garay, 2011: The regional ocean modeling system (roms) 4-dimensional variational data assimilation systems part i - system overview and formulation. *Progress in Oceanography*, **91**, 34–49.

Mullholland, D. P., P. Laloyaux, K. Haines, and M. Balmaseda, 2015: Origin and impact of initialisation shocks in coupled atmosphere-ocean forecasts. *Mon. Wea. Rev., in press*, doi:http://dx.doi.org/10.1175/MWR-D-15-0076.1.

Navon, I. M., 1997: Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dyn. Atmos. Oceans.*, **27**, 55–79.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Q. J. R. Met. Soc.*, **133**, 347–362.

Saha, S., S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, and coauthors, 2010: The NCEP climate forecast system reanalysis. *BAMS*, **91**, 1010–1057.

Saha, S., S. Nadiga, C. Thiaw, W. Wang, J. Wang, Q. Zhang, and coauthors, 2006: The NCEP climate forecast system. *J. Climate*, **19**, 3483–3517.

Shelly, A., P. Xavier, D. Copsey, T. Johns, J. M. Rodriguez, S. Milton, and N. Klingaman, 2014: Coupled versus uncoupled hindcast simulations of the Madden-Julian Oscillation in the Year of Tropical Convection. *Geophysical Research Letters*, **41 (15)**, 5670–5677, doi:10.1002/2013GL059062, URL `http://dx.doi.org/10.1002/2013GL059062`.

Siddorn, J., C. Harris, H. Lewis, and W. Tennant, 2014: Coupling strategy for weather. *Met Office MOSAC and SRG Meetings 2014*, http://www.metoffice.gov.uk/learning/library/publications/s office–scientists/mosac.

Simmons, A. J., and D. M. Burridge, 1981: An Energy and Angular-Momentum Conserving Vertical Finite-Difference Scheme and Hybrid Vertical Coordinates. *Mon. Wea. Rev.*, **109 (4)**, 758–766.

Smith, P. J., A. M. Fowler, and A. S. Lawless, 2015: Exploring strategies for cou-

pled 4D-Var data assimilation using an idealised atmosphere-ocean model. *Tellus A*, doi: 10.3402/tellusa.v67.27025.

Smith, S. M., R. Eade, and H. Pohlmann, 2013: A comparison of full-field and anomaly inialization for seasonal to decadal climate prediction. *Clim. Dyn.*, **41**, 3325–3338.

Sugiura, N., T. Awaji, S. Masuda, T. Mochizuki, T. Toyoda, T. Miyama, H. Igarashi, and Y. Ishikawa, 2008: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *J. Geophys. Res..*, **113**, C10 017.

Takaya, Y., F. Vitart, G. Balsamo, M. Balmaseda, M. Leutbecher, and F. Molteni, 2010: Implementation of an ocean mixed layer model in IFS. Technical Memorandum 622, European Centre for Medium-Range Weather Forecasts, Reading UK.

Thépaut, J.-N., P. Courtier, G. Belaud, and G. Lemaître, 1996: Dynamical structure functions in a four dimensional variational assimilation: A case study. *Q. J. R. Met. Soc.*, **122**, 535–561.

Todling, R., 2014: A lag-1 smoother approach to system-error estimation: sequential method. *Q. J. R. Met. Soc.*, doi:10.1002/qj.2460.

Trémolet, Y., 2006: Accounting for an imperfect model in 4d-var. *Quarterly Journal of the Royal Meteorological Society*, **132 (621)**, 2483–2504.

van Leeuwen, P. J., 2014: Representation errors and retrievals in linear and nonlinear data assimilation. *Q. J. R. Met. Soc.*, **141**, 1612–1623.

Vannière, B., E. Guilyargi, T. Toniazzo, G. Madec, and S. Woolnough, 2014: A systematic approach to identify the sources of tropical SST errors in coupled models using the adjustment of initialised experiments. *Clim. Dyn.*, **43**, 2261–2282.

Vitart, F., and Coauthors, 2008: The new VarEPS-monthly forecasting system: A first step towards seamless prediction. *Q. J. Roy. Meteorol. Soc.*, **134 (636)**, 1789–1799, doi:10.1002/qj.322, URL http://dx.doi.org/10.1002/qj.322.

Waller, J. A., S. Dance, A. S. Lawless, N. K. Nichols, and J. Eyre, 2013: Representativity error for temperature and humidity using the Met Office high-resolution model. *Q. J. R. Meteorol. Soc.*, **140**, 1189 –1197.

Weaver, A. T., J. Vialard, and D. L. T. Anderson, 2003: Three- and four-dimensional variational assimilation with a general circulation model of the tropical pacific ocean. part I: formulation, internal diagnostics, and consistency checks. *Mon. Wea. Rev.,*, **131**, 1360???1378.

Wergen, W., 1992: The effect of model errors in variational data assimilation. *Tellus*, **44A**, 297–313.

Woolnough, S., F. Vitart, and M. A. Balmaseda, 2007: The role of the ocean in the Madden-Julian oscillation: implications for MJO prediction. . *Q. J. Roy. Meteorol. Soc.*, **133**, 117–128.

Zhang, S., M. J. Harrison, A. Rosati, and A. Wittenberg, 2007: System design and evaluation of coupled ensemble data assimilation

for global oceanic climate studies. *Mon. Wea. Rev*, **135**, 3541–3564.

Zheng, Y., T. Shinoda, J. L. Lin, and G. N. Kiladis, 2011: Sea surface temperature biases under the stratus cloud deck in the Southeast Pacific Ocean in 19 IPCC AR4 coupled general circulation models. *J. Clim.*, **24**, 4139–4164.
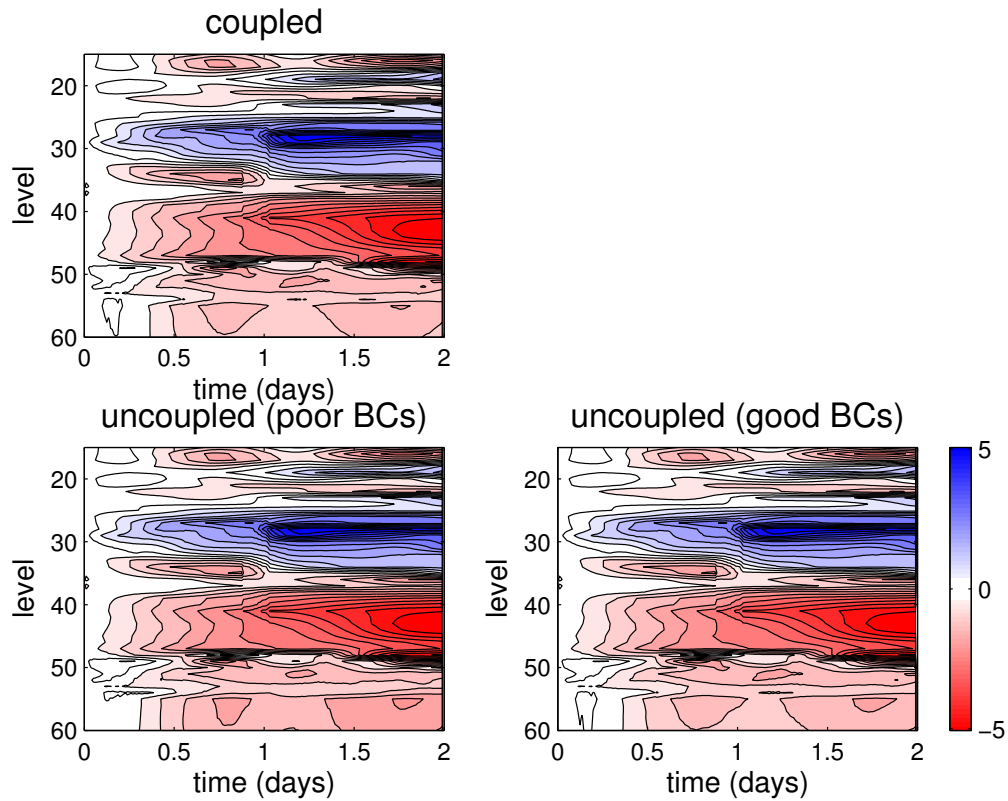
Figure 3: Nonlinear model error $\epsilon^{\hat{\mathcal{H}}}$ for atmospheric temperature (K). The different panels show the errors present in the four different coupling strategies, from left to right: strongly coupled DA, weakly coupled DA, uncoupled DA (poor BCs) and uncoupled DA (good BCs).
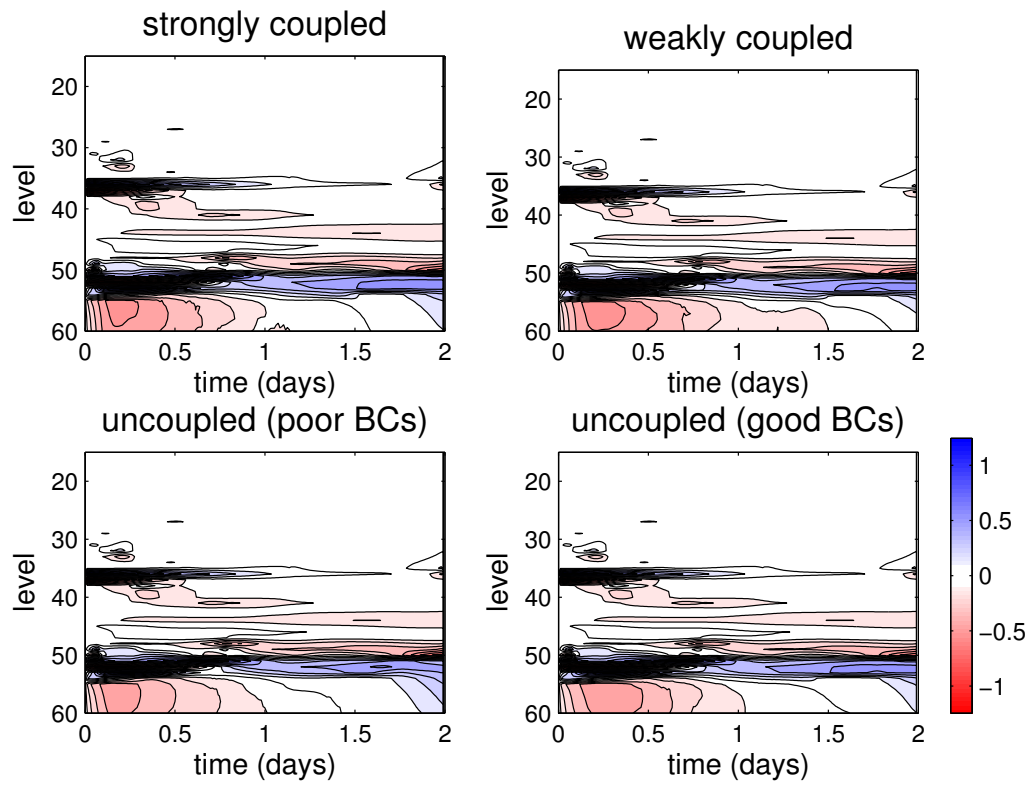
Figure 4: Tangent linear model error $\boldsymbol{\epsilon}^{\mathrm{TL}}$ for atmospheric temperature (K). The different panels are as in Figure 3.
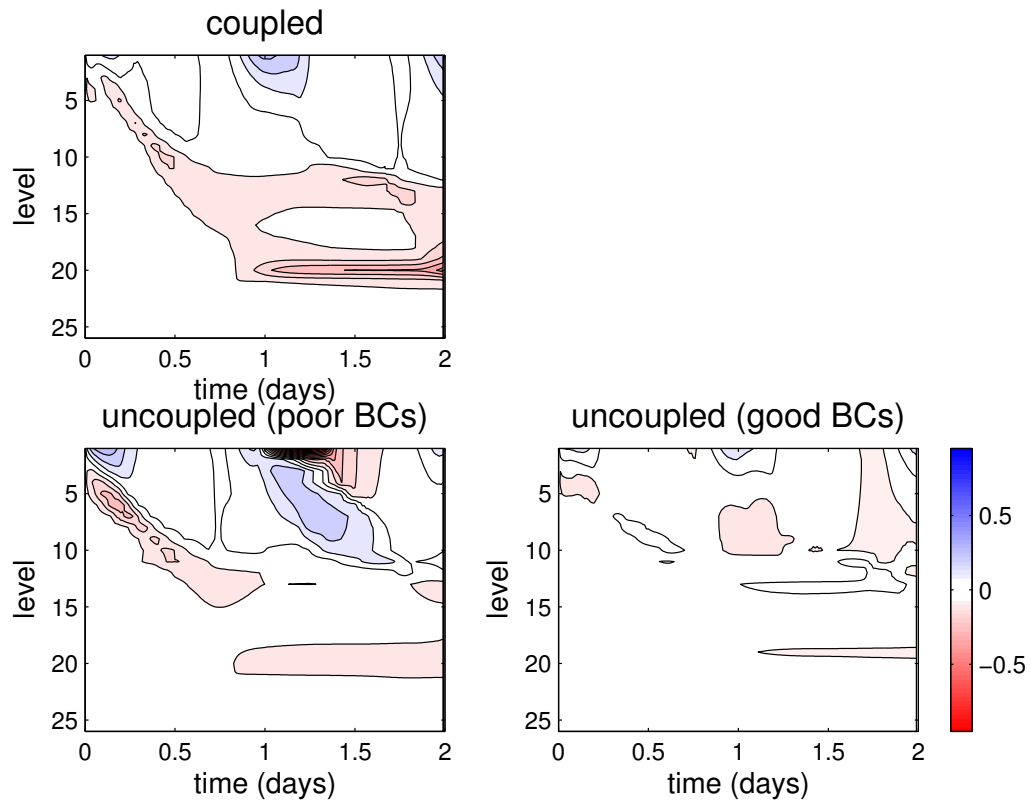
Figure 5: Nonlinear model error $\epsilon^{\hat{\mathcal{H}}}$ for oceanic temperature (K). The different panels are as in Figure 3.
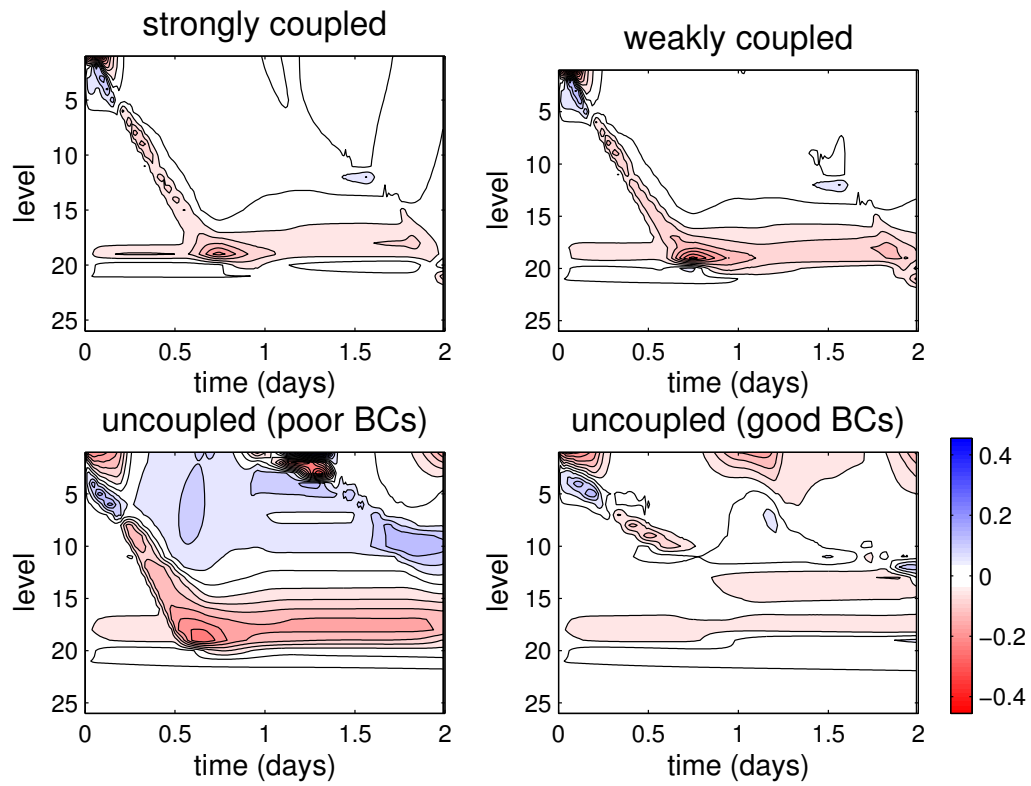
Figure 6: Tangent linear model error ϵ^{TL} for oceanic temperature (K). The different panels are as in Figure 3.
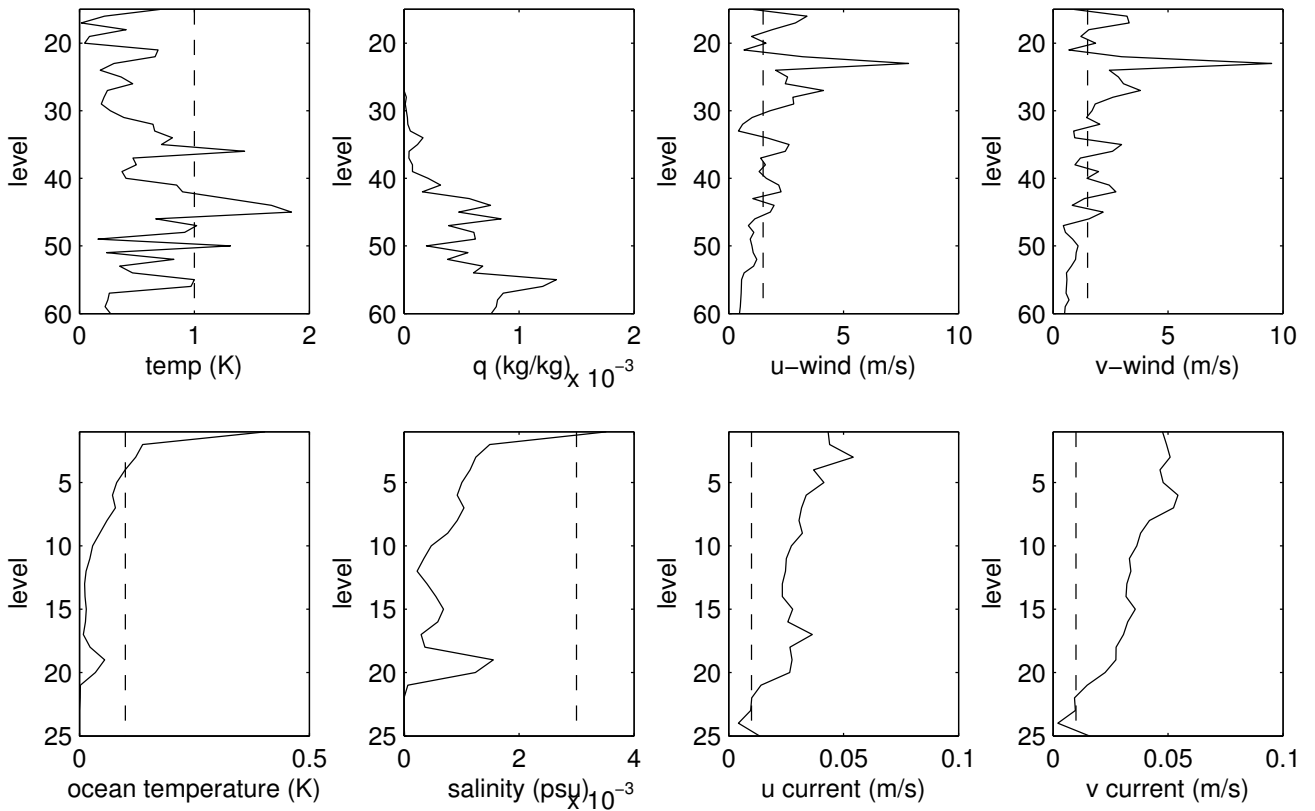
Figure 7: Profile of background error standard deviations for each of the prognostic variables (solid line) and the observation error standard deviations (dashed line). Note that humidity is unobserved.
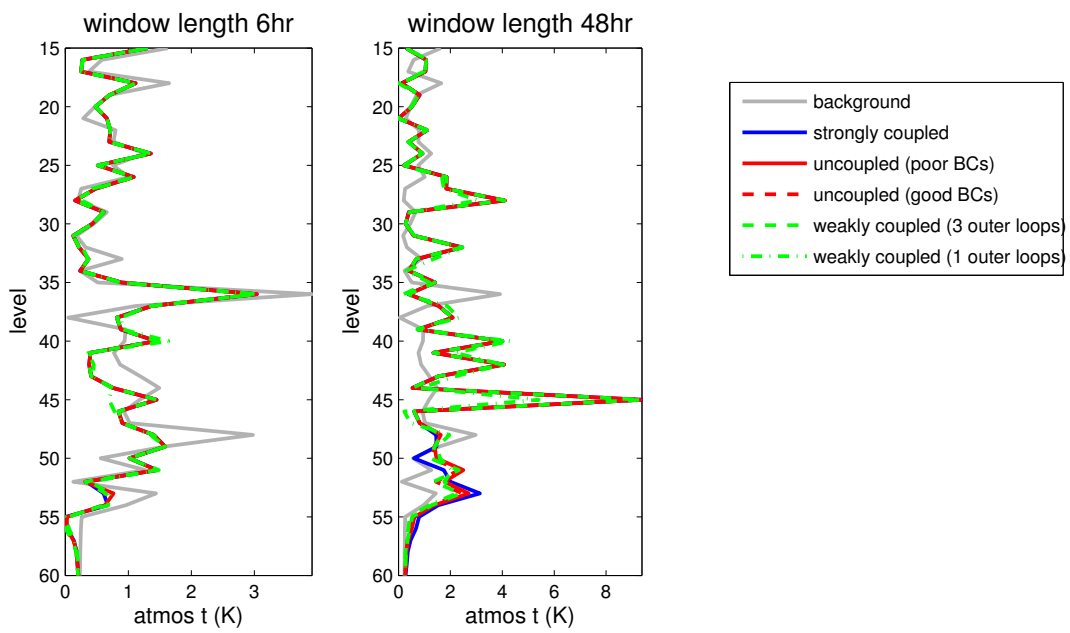
Figure 8: Absolute error in the analysis of atmospheric temperature at initial time using the four different coupling strategies. The window length increases from 6hr (left) to 48hrs (right).
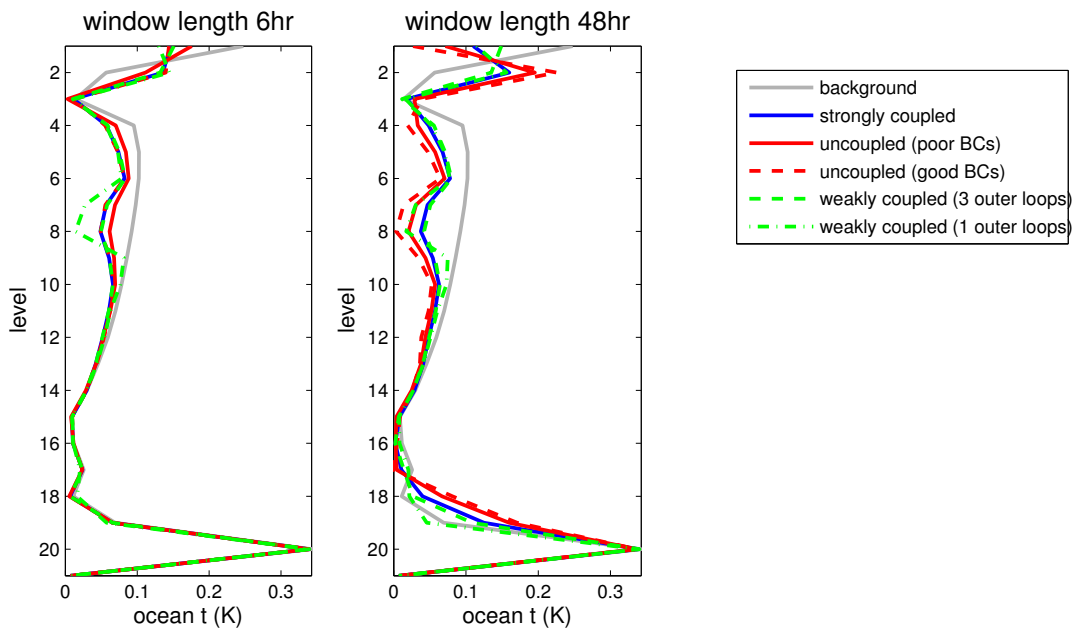
Figure 9: As in figure 8 but for ocean temperature.
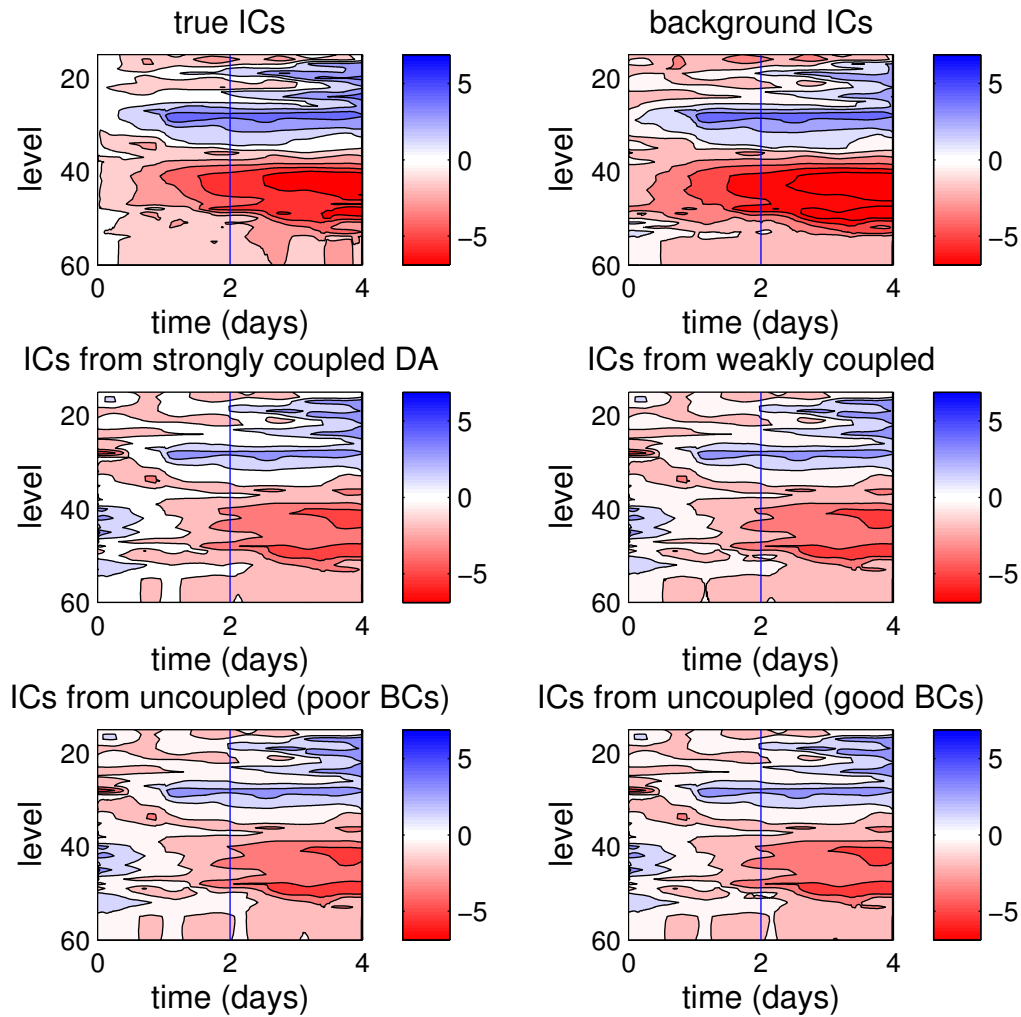
Figure 10: Forecast error in forecasts of atmospheric temperature produced using the erroneous coupled model initialized using different initial conditions. The top two panels show results when the model is initialized using the true state (left) and the background state (right). The other four panels show results when the assimilation model is initialized using the analysis produced by the different coupling strategies using a 48hr window and 3 hourly atmospheric observations and 6 hourly oceanic observations. The blue vertical line indicates the end of the 2 day assimilation window.

34
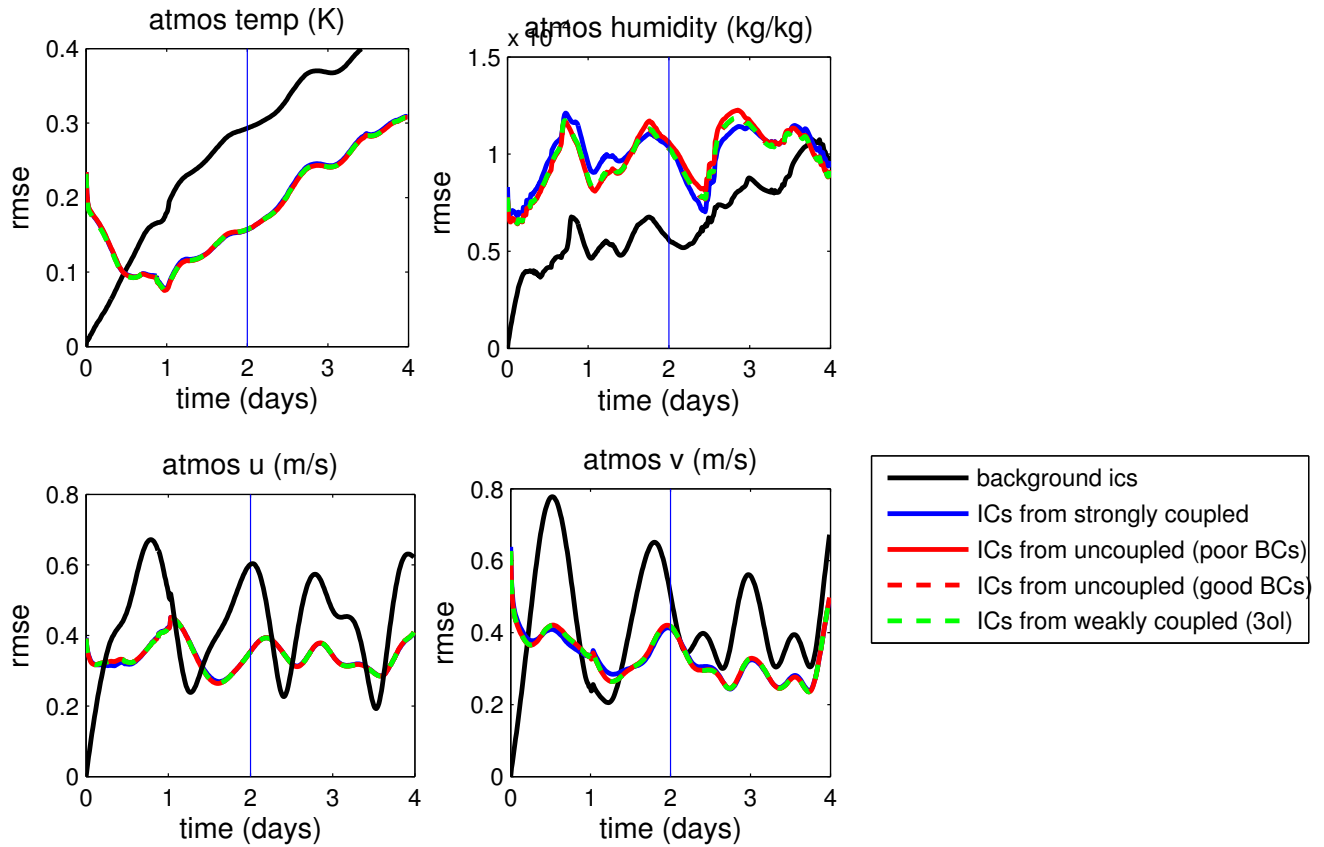
Figure 11: RMSE during the 48 hour assimilation window and subsequent 48 hour forecast window using the coupled model for the atmospheric variables. Colored lines are the RMSE in the forecast when initialized with the different analyses (see figure 8 for legend), the black line is the RMSE in the forecast when initialized with the truth.
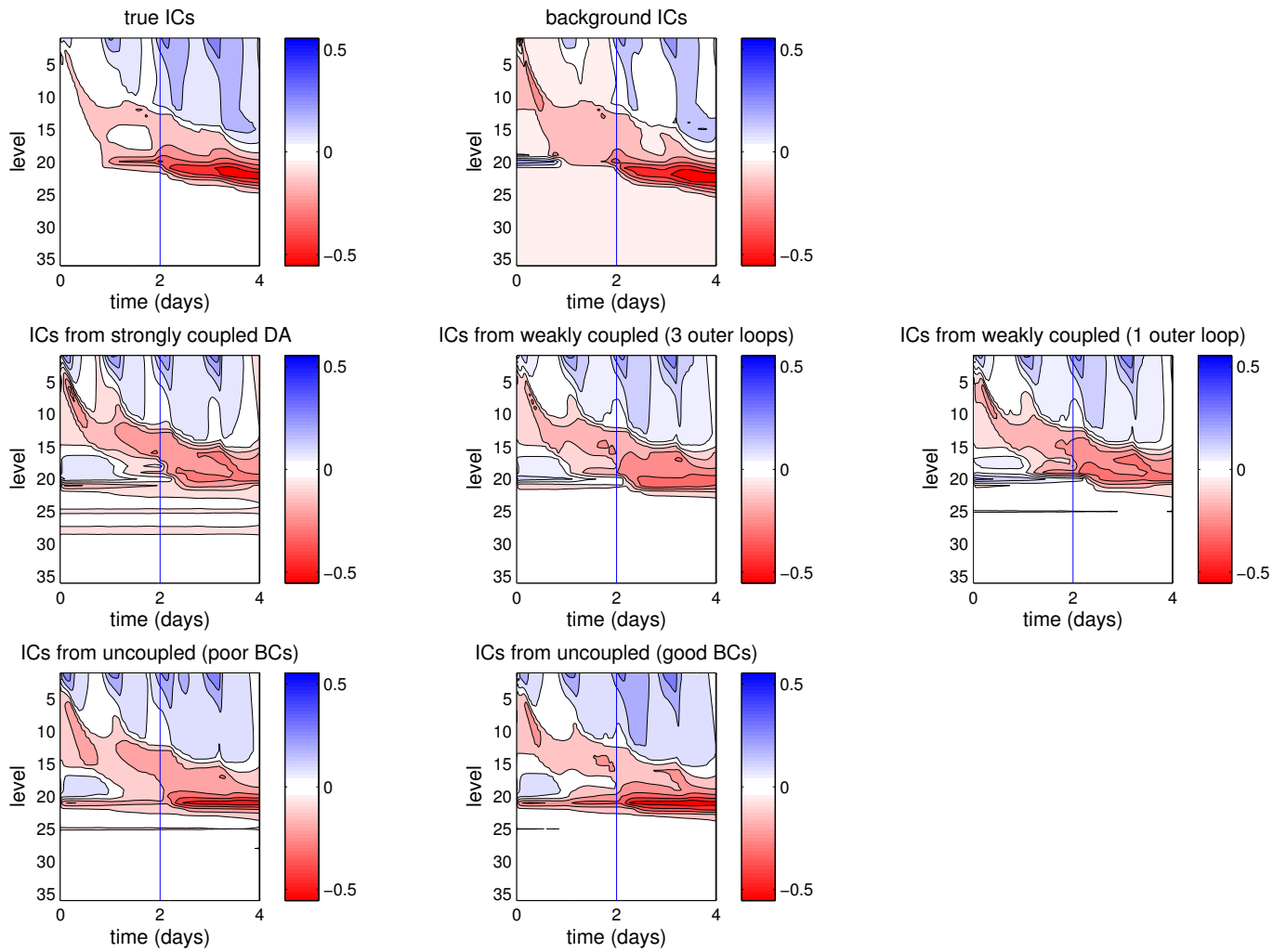
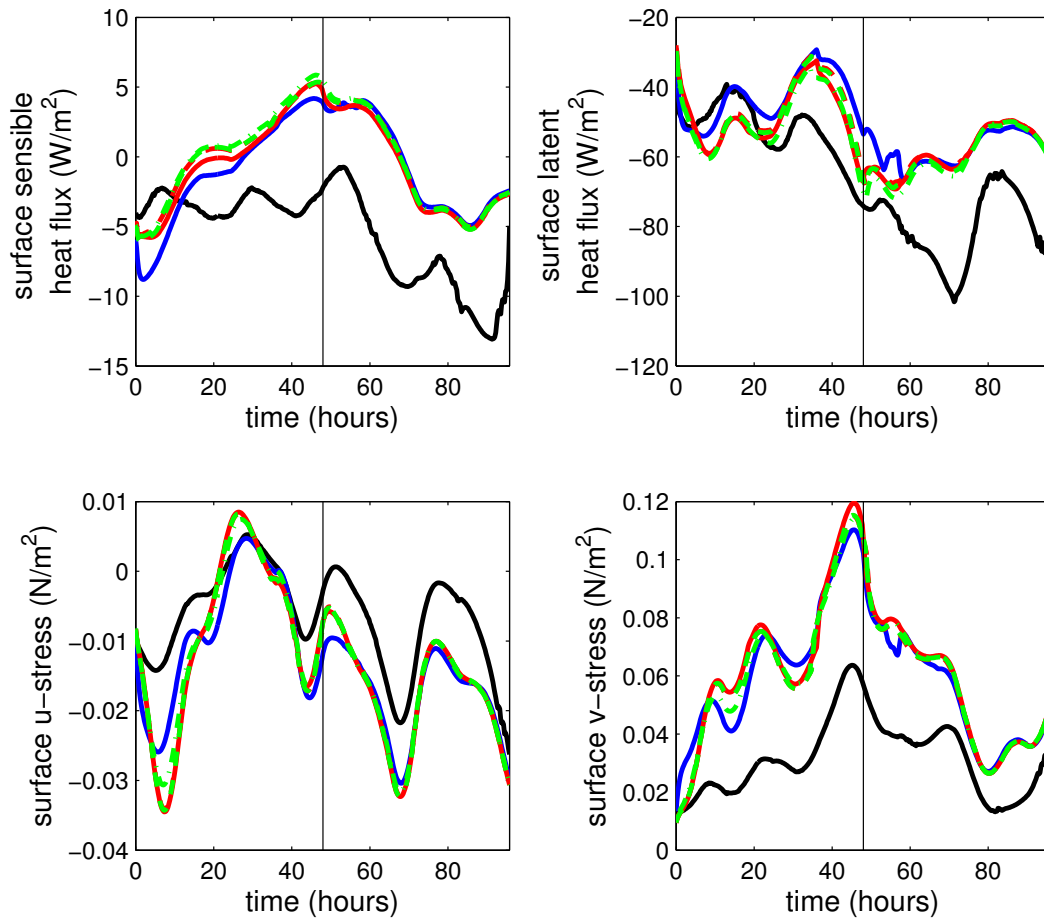Figure 12: As in figure 10 but for ocean temperature.

Figure 13: Forecasts of the surface fluxes when initialized using the different analyses (colored lines as in figure 8). These can be compared to the truth (thick black line).

37
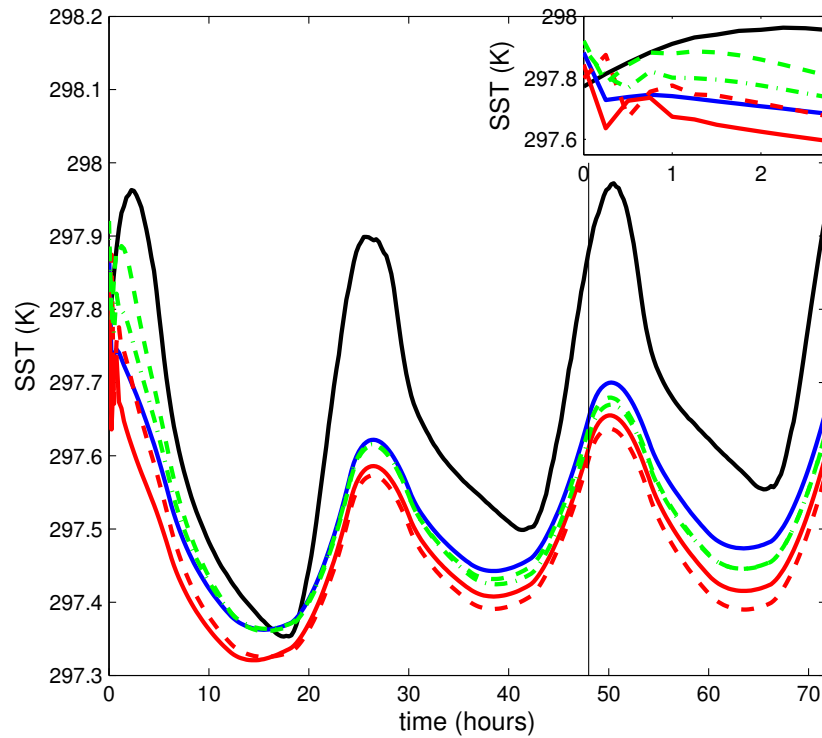
Figure 14: Forecasts of SST initialized using the different coupling DA strategies (colored lines as in figure 8). These can be compared to the truth (thick black line). Inset shows just the first three hours to highlight any initialization shock.
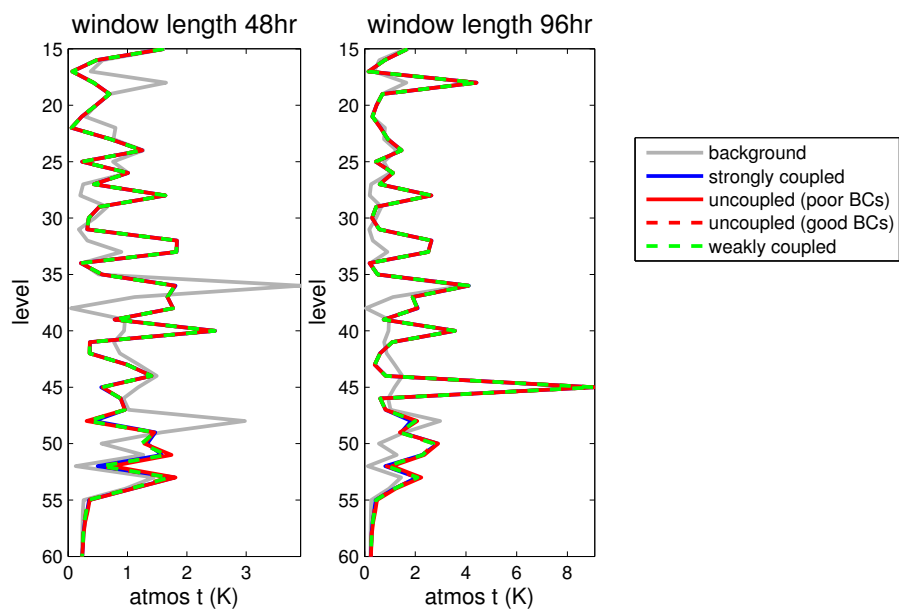
Figure 15: Analysis error for atmospheric temperature using reduced observations with a 48 and 96 hour assimilation window.
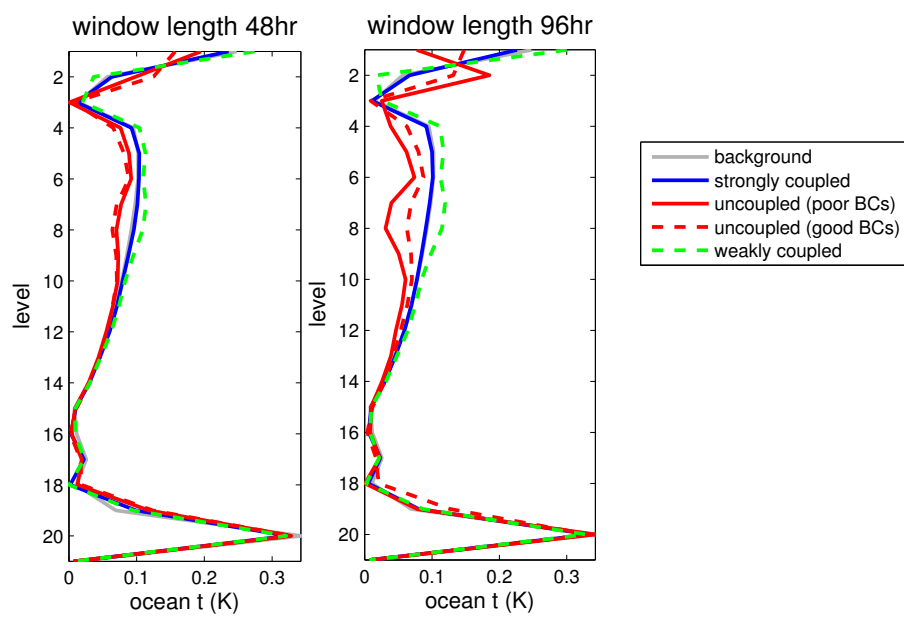
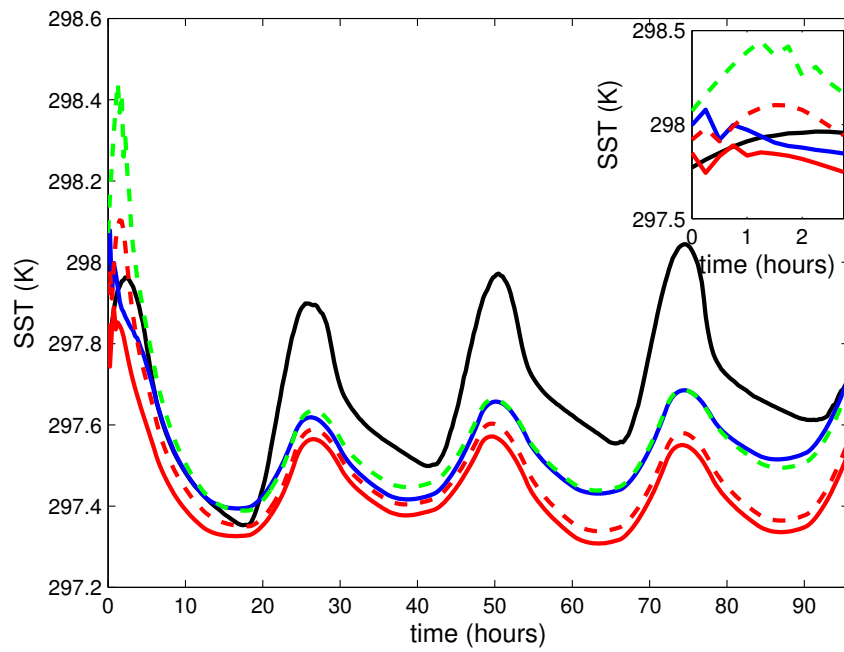Figure 16: As in figure 15 but for ocean temperature.

Figure 17: As in figure 14 but for a reduced observation resolution and a 96 hour assimilation window.