THE SEA: THE SCIENCE OF OCEAN PREDICTION
Journal of Marine Research, 75, 331–359, 2017

Data assimilation for initialization of seasonal forecasts

by Magdalena A. Balmaseda[1,2]

ABSTRACT

This article reviews the requirements for a data assimilation system from the perspective of initializing seasonal forecasts. It provides a historical perspective of the developments in ocean data assimilation and ocean observing systems. It also discusses the differences between state estimation and initialization, and presents a brief assessment of different initialization strategies.

The value of assimilating ocean data to estimate the ocean state and to initialize seasonal forecasts is demonstrated. However, it is also shown that the assumption of unbiased models in conventional data assimilation methods is not suitable for the production of long temporal records of ocean initial states. This is due to the combined effect of model-forcing error and the changing nature of the observing system. Bias correction algorithms are therefore important in the estimation of long records of ocean states. In the equatorial ocean, the delicate balance between the mass and the velocity fields should be preserved in order to maintain realistic circulations.

The most common approach for initializing seasonal forecasts is the so-called full uncoupled initialization, which basically consists of producing an ocean reanalysis by assimilating ocean observations into an ocean model driven by atmospheric fluxes. Alternative approaches are the so-called anomaly initialization, which only attempts to initialize the anomalous state without any attempt of correcting mean; the latter is usually conducted in coupled mode, but coupled and anomaly initialization are not synonymous, and there are approaches where the initialization of the full state is done in coupled mode. The relative value of the approaches is system dependent, but as a long-term strategy the full initialization in coupled mode is more promising.

Keywords: initialization seasonal forecasts, coupled models, initialization strategies, full initialization, anomaly initialization, ocean data assimilation, ocean observing system, flux correction, bias correction

1. Introduction

There is clear demand for reliable forecasts of climate at seasonal timescales for a variety of societal applications. Good-quality seasonal forecasts with reliable uncertainty estimates are of great value to society, allowing institutions and governments to plan actions to minimize risks, manage resources, and increase prosperity and security. Human and economic losses (e.g., famine, epidemics) that may be caused by adverse climate events

1. European Centre for Medium Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, United Kingdom
2. Corresponding author: ***e-mail: magdalena.balmaseda@ecmwf.int***

can be mitigated with early warning systems and disaster preparedness. Equally, adequate planning can aid the exploitation of favorable climate conditions.

Seasonal forecasting is currently a routine activity in several operational centers, with a growing number of economic and societal applications in fields such as agriculture, health, and energy.

Seasonal forecasts predict variations in the atmospheric circulation in response to anomalous boundary forcing, which significantly changes the probability of occurrence of weather patterns (Palmer and Anderson 1994). In order to extend the predictability horizon, these boundary conditions need to be either slow-varying or predictable given the initial conditions. Examples of boundary forcing are variations of sea surface temperature (SST), land conditions (snow depth, soil moisture), sea-ice, and radiative gases. This paper deals in particular with the initialization of the ocean for successful predictions of SST at seasonal timescales. Of special importance are the variations of the tropical SST associated with El Niño Southern Oscillation (ENSO), which have the potential to alter the large-scale atmospheric circulation associated with tropical convective cells.

In this paper, we discuss the application of data assimilation for the initialization of seasonal forecasts, organized as follows: a brief overview of the impact of SST variations in climate, presentation of different elements of a seasonal forecasting system, and in-depth consideration of the initialization problem. The latter includes introduction of data assimilation systems used for ocean initialization, the impact of assimilating data on forecast skill, and an assessment of initialization strategies. We conclude with a discussion of future prospects and directions regarding the initialization of coupled model forecasts.

2. Impact of SST on climate

The dominant climate fluctuations at interannual timescales are related to ENSO, a quasi-periodic warming of sea surface temperatures in the eastern and central equatorial Pacific that affects the patterns of temperature and rainfall in much of the world (Bjerknes 1969). ENSO plays a dominant role in the climate anomalies over the land areas surrounding the entire Pacific basin. The effects of ENSO are also noticeable in other tropical and extratropical regions via the so-called atmospheric bridge (Lau et al. 1996; Klein et al. 1999)—for example, in the Indian monsoon, Atlantic hurricanes, and the climate of southern and eastern Africa. It has been shown that the most predictable variations in worldwide precipitation at interannual timescales are related to ENSO (Goddard and Dilley 2005).

The central role ENSO plays for seasonal forecasting is enhanced by its relatively high potential predictability, which is inherently dependent on the ocean's initial conditions (Zebiak and Cane 1987), and in particular on the precursor provided by the equatorial heat recharge (Jin 1997). Realizing the potential depends critically on the adequacy of initial conditions for the ocean component of the coupled models used for prediction. It is now accepted that linear wave dynamics is insufficient for predicting the evolution of SST: not every eastward propagating Kelvin wave leads to an SST anomaly of the expected sign, and

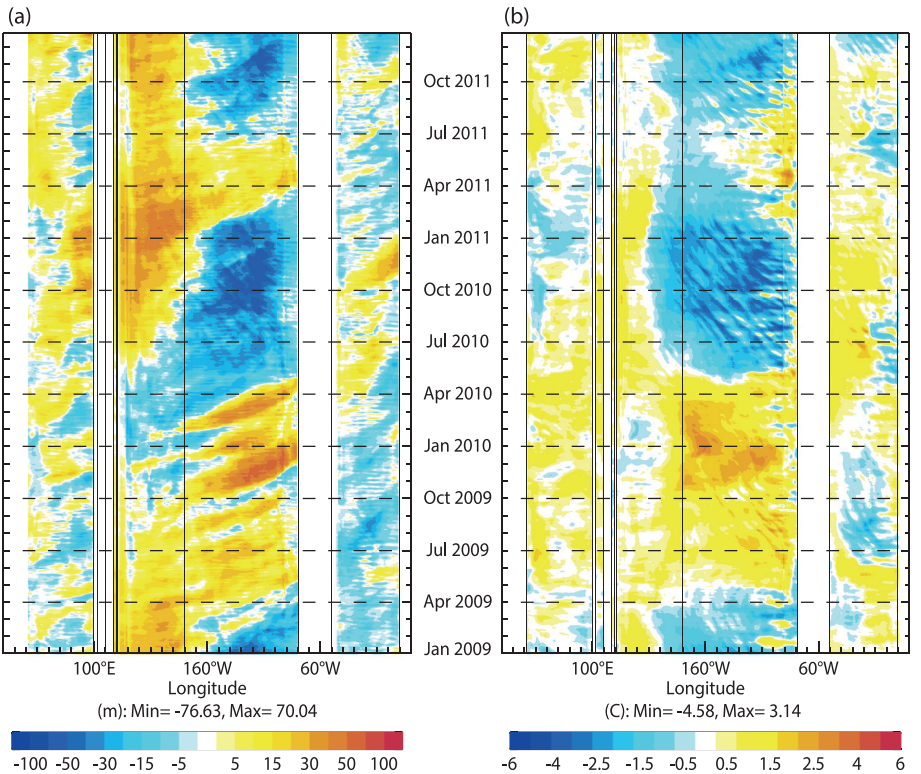

Figure 1. Longitude–time diagrams of equatorial anomalies of (a) thermocline depth and (b) sea surface temperature (SST), for 2009–2011 daily anomalies (1981–2009 climate) ending 31 December 2011. The thermocline depth is represented by the depth of the 20-degree Isotherm (D20). The anomalies are computed with respect to the 1989–2008 climatology. The eastward propagation of equatorial Kelvin waves is visible in D20 usually preceding the appearance of SST anomalies in the eastern Pacific (from the ORAS4 ocean reanalysis [Balmaseda et al. 2013]).

intensity varies from one event to another (Fig. 1). Indeed, the observed ENSO diversity is the result of the different mean-state/wave interactions, the role of the equatorial and extra-equatorial ocean and other tropical ocean basins in modulating the large-scale atmospheric convection, the response of the ocean to different aspects of the so-called Westerly Wind Events (intensity, timing, and fetch, among others), and the role of salinity in the vertical mixing and horizontal pressure gradients (Zhu et al. 2014).

Figure 1 shows time–longitude diagrams illustrating the cross-equatorial eastward propagation of thermocline anomalies (left) preceding the onset of the ENSO-related SST anomalies (right) in the eastern Pacific. The figure also shows that a thermocline anomaly associated with an individual Kelvin wave does not always translate into a large-scale SST anomaly; for example, in the "failed" El Niño of 2011, in spite of a substantial propagation of the

thermocline anomalies the warm SST anomaly was very short lived, and the El Niño did not materialize.

Variations in tropical SST other than those related to ENSO can also drive temperature and precipitation anomalies on seasonal timescales. Examples include the connection of the tropical Atlantic with rainfall in northeastern Brazil (Folland et al. 2001) and with rainfall in west Africa and Sahel (Giannini et al. 2003), the impact of the extratropical Atlantic (e.g., Rodwell et al. 2002) on European climate, and the tropical Indian Ocean impact on east African rainfall and the Indian monsoon (Goddard and Graham 1999), in particular the mode of variability known as the Indian Ocean Dipole IOD, (Saji et al. 1999). Especially important is the role of SST variations due to the IOD that drives rainfall variations across southern Australia during ENSO (Behera et al. 2005; Cai et al. 2011).

3. Elements of seasonal forecasting systems

Seasonal forecasting systems are based on coupled ocean–atmosphere general circulation models that predict both surface boundary forcing and its impact on the atmospheric circulation. The chaotic nature of the atmosphere is taken into account by issuing probabilistic forecasts, obtained by performing an ensemble of coupled integrations. Because of the deficiencies in coupled models, the forecasts need calibration before they are issued. The calibration is done by conducting a series of retrospective seasonal forecasts over past years (or hindcasts), which in turn require initial conditions for a historical period (typically 15–25 years), usually obtained from reanalyses. The hindcasts are also needed for skill assessment. Figure 2 illustrates schematically the different stages in the generation of a probabilistic forecast.

Figure 3 illustrates the production of probabilistic forecasts and the calibration procedure. At a given time, an ensemble of coupled forecasts is produced to sample the likely range of occurrence of atmospheric states in the following seasons (forecast Probability Density Function or PDF). Due to model error, this PDF needs calibration, and this is achieved by comparing "today's" model PDF with the model climatological PDF. The latter is estimated by a series of coupled model "hindcasts" initialized using ocean initial conditions from a historical record (i.e., from an ocean reanalysis). The seasonal forecast is issued as the difference between the climatological PDF and today's PDF. Differences can show in the mean, the spread, or the tails of the distribution. The ensemble information also allows the estimation of statistical significance. The hindcasts produced for the estimation of model climatological PDF are also used for skill assessment. The quality of seasonal forecasts is therefore determined by the various components of the system (the ocean initialization, the coupled model, the ensemble generation, and the calibration strategy).

The need for calibration and uncertainty information has two important implications for the design of the initialization of operational seasonal forecasts. In addition to the near-real-time knowledge of the ocean state given by an ocean analysis, initialization of operational seasonal forecasts requires an ocean reanalysis, which is as consistent as possible with the real-time ocean analysis system used in the production of today's forecast. The

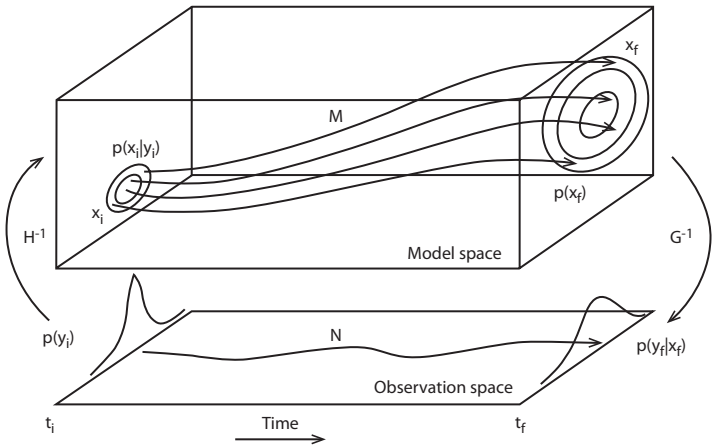

Figure 2. Schematic diagram representing the stages in the production of a calibrated probabilistic forecast. The first stage is the initialization, which merges the physical world observations (y_i observations (y_i, with probability $p(y_i)$)) and model (x_i) information by adequate use of probabilities at initial time t_i. The initialization is also equivalent to a transformation H^{-1} from observation to model spaces, and provides $p(x_i|y_i)$, or the probability of a model state given the observations. The model M is using an approximation of the laws of physics N, thus producing a probability forecast in model space $p(x_f)$. The probability is estimated by means of an ensemble of forecasts. The final stage at time t_f is the calibration G^{-1}, which estimates the probability $p(y_f|x_f)$ of occurrence of some event in the physical world (y_f) given the model forecast probability.

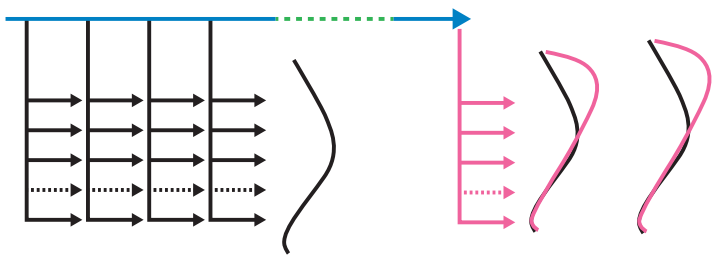

Figure 3. Schematic procedure for forecast calibration of probabilistic forecasts. An ensemble of coupled integrations initialized for the date of interest is used to sample the forecast probability (magenta). A series of hindcasts initialized for the historical period sample the model climatology (black). The differences between forecast and climatology provide insight about possible climate anomalies and their statistical significance. In the figure, the horizontal blue arrow represents time: the end of the arrow is real time while the length of the arrow represents the re-forecast period, of arbitrary length (dotted line; typically 20–30 years or longer). The different short arrows (black and magenta) represent the individual forecasts comprising the ensemble. A typical number of ensemble members is 50 for a real time forecast, and 11 for the re-forecast.

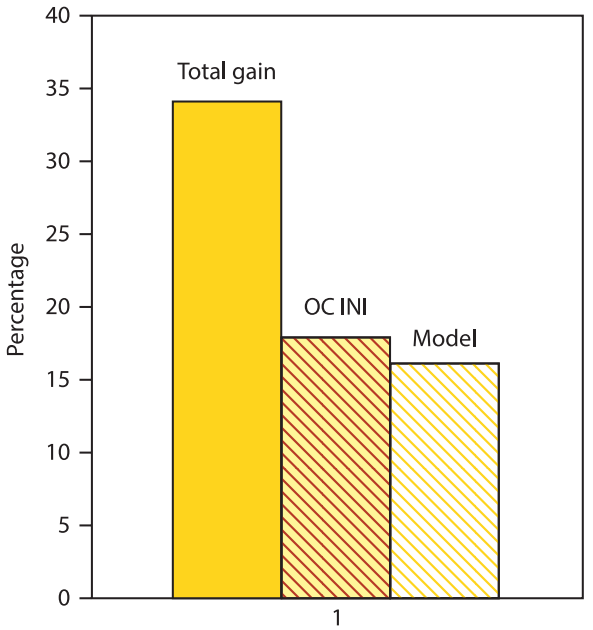

Figure 4. Progress in the seasonal forecast skill of the European Centre for Medium Range Weather Forecasts (ECMWF) operational system during a decade. The solid bar shows the relative reduction in mean absolute error of forecast of sea surface temperature in the eastern Pacific (NINO3). The brown-striped bar shows the contribution from the ocean initialization, and the white-striped bar is the contribution from model improvement (from Balmaseda, Fujii, Alves and Lee et al. 2010).

uncertainty in the ocean initial conditions—required for probabilistic coupled forecasts—can be sampled by an ensemble of ocean (re-)analyses. The ocean reanalyses produced for the initialization of seasonal forecasts are a valuable resource for climate variability studies (Balmaseda et al. 2013), and have the advantage of being continuously brought up to real time, which allows monitoring of relevant climate variables (Xue et al. 2010).

The production of seasonal forecasts is resource-demanding: the need for ensembles and calibration implies the integration of the coupled model for several hundreds of years. This computational burden limits the practical resolution of the ocean model, which is typically of the order of one to one-quarter of a degree in the horizontal, and about 1–10 meters in the vertical in the upper ocean.

The consolidation of seasonal forecasting as a routine operational activity during the last decades has been possible thanks to the improvement in coupled models, data assimilation methods, availability of forcing fluxes from atmospheric reanalysis, and the development of the ocean observing system. Figure 4 shows the improvements in ENSO forecasts at the European Centre for Medium Range Weather Forecasts (ECMWF) over a decade. The improvements can be attributed equally to better initialization of the ocean and improved coupled models (Balmaseda, Fujii, Alves and Lee et al. 2010).

4. Initialization

The most common approach for initializing seasonal forecasts is the so-called full uncoupled initialization (Balmaseda and Anderson 2009). This basically consists of producing a long ocean reanalysis (typically 20 years or longer) by assimilating ocean observations into an ocean model driven by atmospheric fluxes. More recently, with the advent of decadal forecast the so-called anomaly initialization (Smith et al. 2007) has become popular. In this approach, only the anomalous state is assimilated without any attempt to correct the model mean state in the initial conditions. The anomaly initialization is usually conducted in coupled mode, but coupled and anomaly initialization are not synonymous, and there are approaches where the initialization of the full state is done in coupled mode. The discussion in this and the next section (Impact on forecast skill) assumes full initialization. In the section below entitled "Assessment of initialization strategies," we discuss the differences and merits of full versus anomaly initialization.

a. Full initialization

Assimilation of observations into an ocean model forced by prescribed atmospheric fluxes is the most common practice for initialization of the ocean component of a coupled model (Balmaseda, Fujii, Alves and Awaji et al. 2010). The assimilation should improve the estimation of the ocean state (reducing uncertainty and improving the mean interannual variability), but ultimately it should improve the skill of the seasonal forecasts. These objectives are challenged by the paucity of ocean observations, by the abrupt changes in the ocean observing system, and by presence of model errors. (In other words, it is not guaranteed that a good estimation of the real world can project in the model attractor and thus evolve successfully into the forecasts.) In what follows we discuss why the initialization of seasonal forecasts needs data assimilation. We will describe the ocean observing system, the impact of data assimilation in correcting model error, and the need to apply bias corrections for reliable representation of the interannual variability. The section ends with an example of a seasonal forecast initialization system (that used to initialize the ECMWF System4 [S4] seasonal forecasting system, operational at the time of writing).

b. Why do we need to assimilate subsurface ocean data?

In seasonal forecasts the emphasis is on the initialization of the upper ocean thermal structure, particularly in the tropics, where SST anomalies have a strong influence on the atmospheric circulation.

A simple way of providing initial conditions would be to run an ocean model forced with winds and freshwater fluxes from atmospheric reanalyses and with a strong relaxation of observations of SST.

But the quality of the models and/or surface fluxes is usually not sufficient to provide an accurate estimation of the ocean state. The uncertainty induced in the upper ocean by using different wind products can be as large as the interannual variability (Ji et al. 1995). Figure 5a

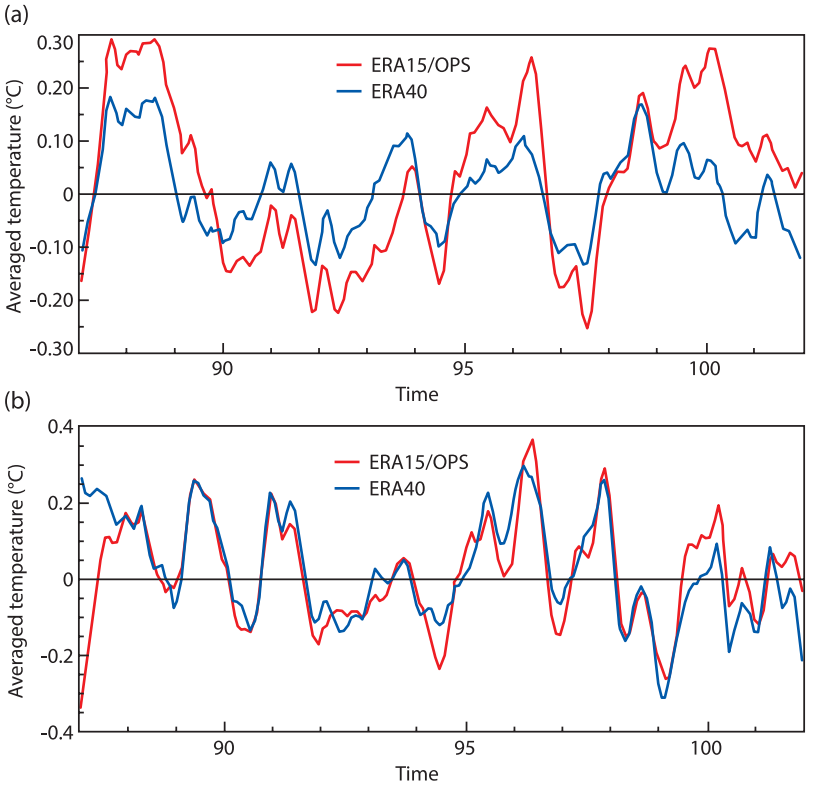

Figure 5. Time evolution of the upper 300 m averaged temperature in the equatorial Atlantic from two ocean model simulations forced by different atmospheric fluxes (a). The magnitude of the differences is comparable to the interannual variability. ERA15 and ERA40 are two consecutive versions of the ECMWF Re-analyses for the Atmosphere. OPS refers to the fluxes from the ECMWF operational output. (b) As above, but assimilating subsurface observations. The assimilation of ocean observations is efficient in reducing the uncertainty in the ocean estimate resulting from uncertainty in surface fluxes (although differences remain at the beginning of the record, since the observations are scarce there). The time axis units are years in the 20th century.

shows the evolution of upper-300-m averaged temperature in the equatorial Atlantic from two ocean-only simulations forced by different atmospheric fluxes. The magnitude of the differences is comparable to the interannual variability. By assimilating ocean observations it is possible to reduce the uncertainty in the ocean estimate. Figure 5b shows the equivalent quantity in two equivalent ocean reanalyses, where the first guess is given by an ocean model forced by the different surface fluxes, but this time subsurface ocean observations are assimilated.

A distinctive aspect of data assimilation in the initialization of seasonal forecasts is the use of observations to constrain the large-scale forced ocean variability. In other applications, such as short-range weather or marine forecasts, data assimilation is used to constrain

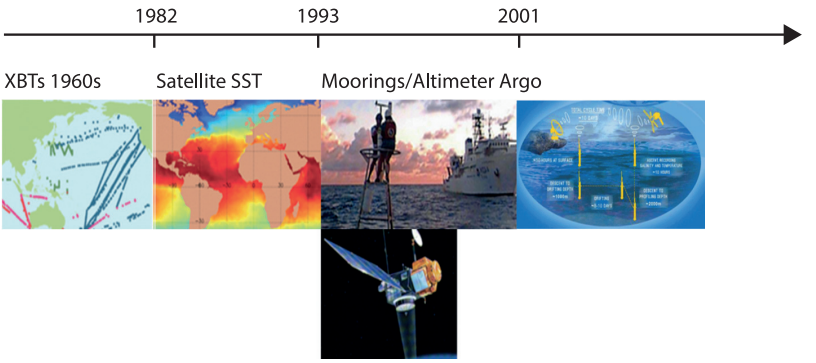

Figure 6. Time evolution of the ocean observing system by type of instrument (XBT, expendable bathythermograph; SST, sea surface temperature).

mainly their chaotic nature; in the case of high-resolution marine forecasting, this implies constraining the small (eddy) scales. In contrast, the observations in seasonal (and decadal) forecasts are used to a large extent to initialize the wind-driven circulation, which often occurs in large-scale modes.

In addition to reducing the uncertainty in the ocean state, the data assimilation of ocean observations often also improves the estimation of the ocean thermal structure and the skill in forecasting the tropical SST at interannual timescales (see for instance Balmaseda et al. 2013).

c. Ocean observing system

Figure 6 shows schematically the different components of the ocean observing system and their availability in time. Sea surface temperature observations are essential for seasonal forecasts. Most of the initialization systems also use subsurface temperature from XBTs (Expendable bathythermograph; Goni et al. 2010); CTDs (Conductivity, Temperature, and Depth), usually from scientific cruises; moored buoys (TAO/TRITON in the Pacific, PIRATA in the Atlantic, RAMA in the Indian Ocean; see McPhaden et al. 2010); and Argo floats (Freeland et al. 2010). Salinity (mainly from Argo and CTDs) and altimeter-derived sea-level anomalies (SLAs, since approximately 1993; Wilson et al. 2010) are also assimilated. The latter usually need a prescribed external Mean Dynamic Topography (MDT), which can be derived indirectly from gravity missions such as GRACE (Gravity Recovery and Climate Experiment) and, in the near future, GOCE (Gravity field and steady-state Ocean Circulation Explorer). Recent seasonal forecasting systems also include interactive sea-ice, which is initialized using sea-ice concentration estimates from space-born instruments (primarily built up from passive microwave data); these provide a record from 1979 until the present. Determination of sea-ice thickness from space is less mature. Estimates based on ice freeboard are currently provided from high-inclination orbit by the CryoSat

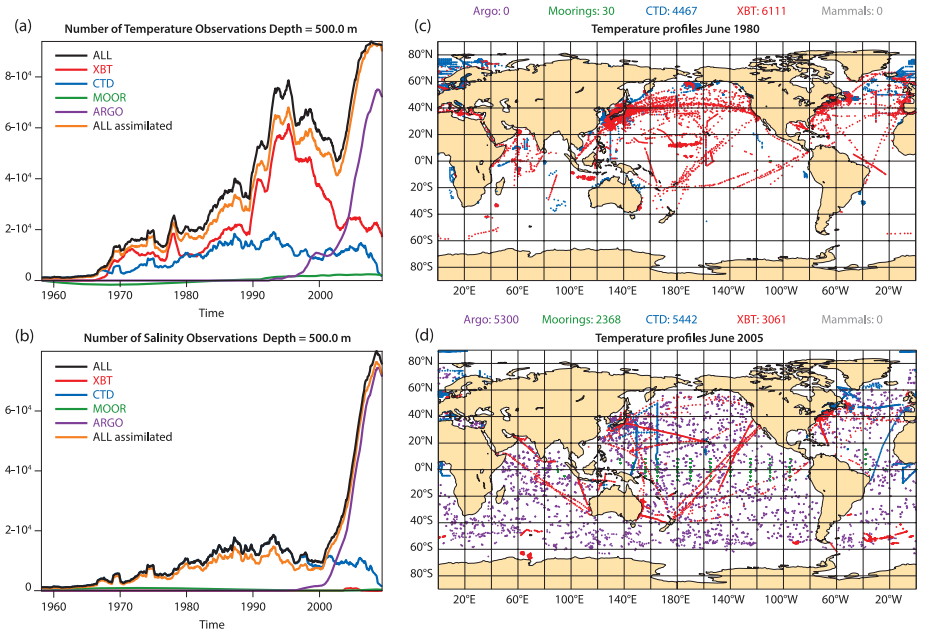

Figure 7. Left: Number of temperature (a) and salinity (b) observations within the depth range 400–600 meters as a function of time per instrument type. The black curve is the total number of observations. The orange curve shows the number of assimilated observations in the ORAS4 reanalysis (from Mogensen et al 2012). Right: Typical observation coverage in June 1980 (c) and in June 2005 (d). Note that the color coding for the instruments is not the same in the left and right panels.

radar altimeter; data from laser altimetry was provided earlier by ICESat-1 and in the future will come from ICESat-2. Brightness temperature from L-band radiometry such as those in SMOS also provides estimates of thickness for thin-ice.

Figure 7 (left column) shows the number of subsurface temperature (top) and salinity (bottom) observations in the depth range 400–600 m as a function of time, illustrating the large increase in observations associated with the advent of Argo. The right panels of Figure 7 show the spatial observation coverage in June 1980 (top) and in June 2009 (bottom). The properties of spatial and temporal sampling vary substantially between instruments: the XBTs usually follow commercial ship routes, CTDs are associated with intense scientific missions, and moored arrays sample the equatorial oceans at a few selected fixed positions. Argo is the only observing system that samples the subsurface of the ocean uniformly, measuring temperature and salinity up to a depth of 2000 m. Altimeter sea-level (not shown) also samples the surface of the ocean quite uniformly, but a good relation between sea level variations and subsurface structure is only possible in regions of strong stratification (the tropics).

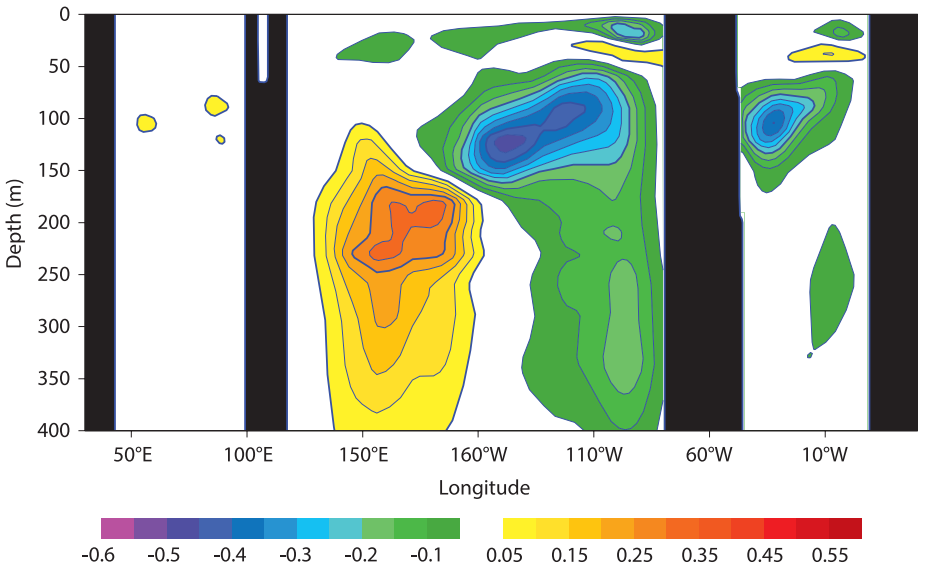

Figure 8. Equatorial longitude-depth section of mean assimilation temperature increment from a previous European Centre for Medium Range Weather Forecasts (ECMWF) ocean analysis system (ORAS2). The contour interval is 2 × 10-4 °C/hr. The mean corresponds to the time-average during the period 1987–2001 (from Balmaseda et al. 2008).

d. Assimilation increment and model error: Need for bias correction

Figure 8 shows the 1987–2001 average of a longitude–depth section of the assimilation increments along the equator from a previous ECMWF ocean analysis (system 2). The nonzero mean increment is indicative of a systematic model error. In this particular case, the large-scale dipolar structure of the increment can be interpreted as a correction in the slope of the thermocline (making it deeper in the western Pacific and shallower in the eastern Pacific). This kind of error could appear if the equatorial winds were too weak, although it may be due to incorrect ocean mixing.

Notwithstanding the source of error, Figure 8 shows that the data assimilation is correcting the system bias, whereas the scheme assumes the first guess given by the model background is unbiased. In practice, the presence of systematic error may introduce spurious temporal variability in regions where the observation coverage is not uniform in time, which may be a serious problem when the ocean analysis is used to predict interannual variability (Balmaseda et al. 2007). This is illustrated in Figure 9, which shows the evolution of 20-degree isotherm in the equatorial Atlantic for three experiments. The control CNTL, in black, corresponds to an ocean simulation forced by atmospheric reanalysis fluxes. The red curve (ASM) is an equivalent experiment where subsurface temperature is being assimilated. There is a clear jump around 1999 caused by the advent of the PIRATA moorings. The data from PIRATA are correcting the position of the model thermocline, and in doing

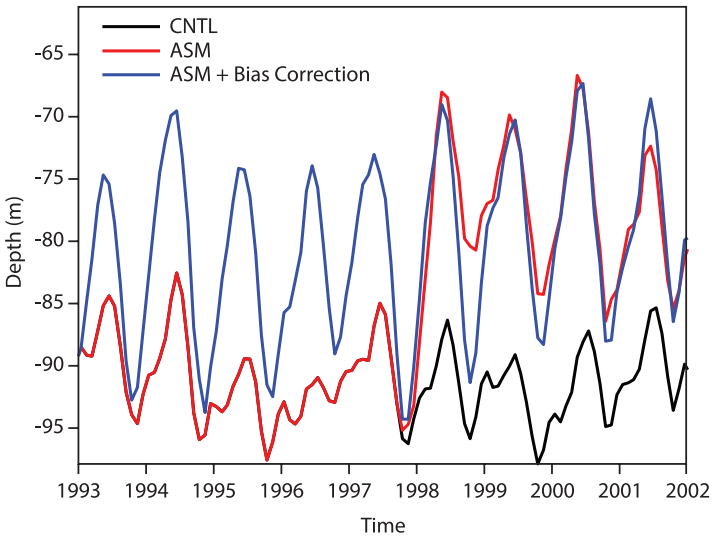

Figure 9. Time evolution of the depth of the 20-degree isotherm in the equatorial Atlantic in an ocean model simulation (CNTL, black), in a data assimilation experiment without bias correction (ASM, red), and in a data assimilation experiment with bias correction (blue). The PIRATA observations are correcting for the too-deep model thermocline, introducing a spurious interannual variability. By applying a bias correction, the depth of the thermocline it is corrected for and the changes in the observing system do not translate into spurious signals.

so introduce a large jump contaminating the interannual variability. The blue curve corresponds to an assimilation experiment where a bias correction scheme (see later) is being applied to the model (or first guess) before each analysis cycle. The bias scheme corrects the model mean error, thus avoiding spurious signals each time that there are changes in the ocean observing system.

The problem of bias is not exclusive to the ocean data assimilation systems; it is also an important issue in atmospheric reanalysis (Dee 2005). Dee and Da Silva (1998) (DdS in what follows) developed an algorithm for the online estimation and correction of the bias in sequential data assimilation. But the general algorithm is costly, since it requires an extra assimilation step to estimate the bias. A simplified DdS algorithm using a single step is also possible, by assuming proportionality between the bias and state variables.

Bell et al. (2004) (BMN in what follows) implemented for the first time a bias correction scheme in ocean data assimilation for tropical oceans. The BMN scheme was quite innovative in that the bias correction is not applied directly to the temperature field, but applied as a correction to the pressure gradient. In this regard, the BMN scheme deviates from the DdS algorithm. Balmaseda et al. (2007) introduced a more general framework in which the bias control variables were linked to the state variables by multivariate bias relationships. In this way, the BMN scheme could be interpreted as a particular choice of multivariate constraint. Thus, the analysis equations of a generalized bias correction were given by

$$\begin{aligned} \mathbf{x}^a &= \mathbf{x}^f + \mathbf{b}^f + \mathbf{K}[\mathbf{y} - \mathbf{H}(\mathbf{x}^f + \mathbf{b}^f)] \\ \mathbf{b}^a &= \mathbf{b}^f + \mathbf{L}[\mathbf{y} - \mathbf{H}(\mathbf{x}^f + \mathbf{b}^f)] \end{aligned} \tag{1}$$

where $\mathbf{x}^a$ and $\mathbf{x}^f$ are the *analysis* and *forecasts state* variables respectively, and $\mathbf{b}^a$ and $\mathbf{b}^f$ are the *analysis and forecast bias* variables. The latter are ignored in the standard data assimilation. The last term on the right-hand side of the equation at the top represents the *assimilation increment*, a correction applied to the forecast variables in order to obtain the analysis. This is proportional to the differences between the observations $\mathbf{y}$ and model variables $\mathbf{x}$, often called *innovation*. Observations and model are compared by collocating the model variables with the observations (in space, time, or other transformation) via the so-called *observation operator* $\mathbf{H}$. In the case represented by Eq. (1), the model variables are also bias-corrected with the first estimate of the bias $\mathbf{b}^f$. The innovation is projected back into the model space with the so-called *gain matrices* $\mathbf{K}$ (for the state variables) and $\mathbf{L}$ (for the bias variables), which can in principle be different. Balmaseda et al. (2007) suggest a linear relationship $\mathbf{A}$ between $\mathbf{L}$ and $\mathbf{K}$:

$$\mathbf{L} = \mathbf{A}\mathbf{K} \tag{2}$$

In Eq. (2) $\mathbf{A}$ can be in general different from the identity, and represents different multivariate relationships in the state and bias variables. Figure 10, from Balmaseda et al. (2007), illustrates the importance of the multivariate relations between state and bias variables: while correcting the bias in the pressure field reduces the bias in temperature and velocity fields, the direct correction of the bias in the temperature field ($\mathbf{A}$ equals identity) reduces the temperature bias, but significantly increases the error in the velocity field.

The bias correction algorithm also requires the prescription of a model for the time evolution of the bias. The simplest and most widely-used model is that of constant-in-time bias. Dee and Todling (2000) discuss this assumption, pointing out the pitfall that a constant bias allows a single observation to influence the bias estimation indefinitely. The introduction of a memory term may thus be desirable. Moreover, the systematic error may not be constant in time: it may be flow dependent (e.g., depend on the diurnal or seasonal cycle), or it may be associated with the nonstationary errors of the external forcing (such as discontinuities in the atmospheric analysis system that provides the surface fluxes).

Balmaseda et al. (2007) propose a generic model for the time evolution of the bias. The model bias correction scheme (Eq. 3) comprises an adaptive component ($\mathbf{b}'$), evolving and updated online, and a-priori component ($\bar{\mathbf{b}}$), which has been estimated from a previous iteration.

$$\begin{aligned} \mathbf{b}_t^f &= \bar{\mathbf{b}}_t + \mathbf{b}_t'^f \\ \mathbf{b}_t'^f &= \alpha \mathbf{b}_{t-1}'^f \end{aligned} \tag{3}$$

The a-priori component ($\bar{\mathbf{b}}$) can represent, for instance, a climatological error estimated from a recent period with more complete observation coverage. This allows the extrapolation

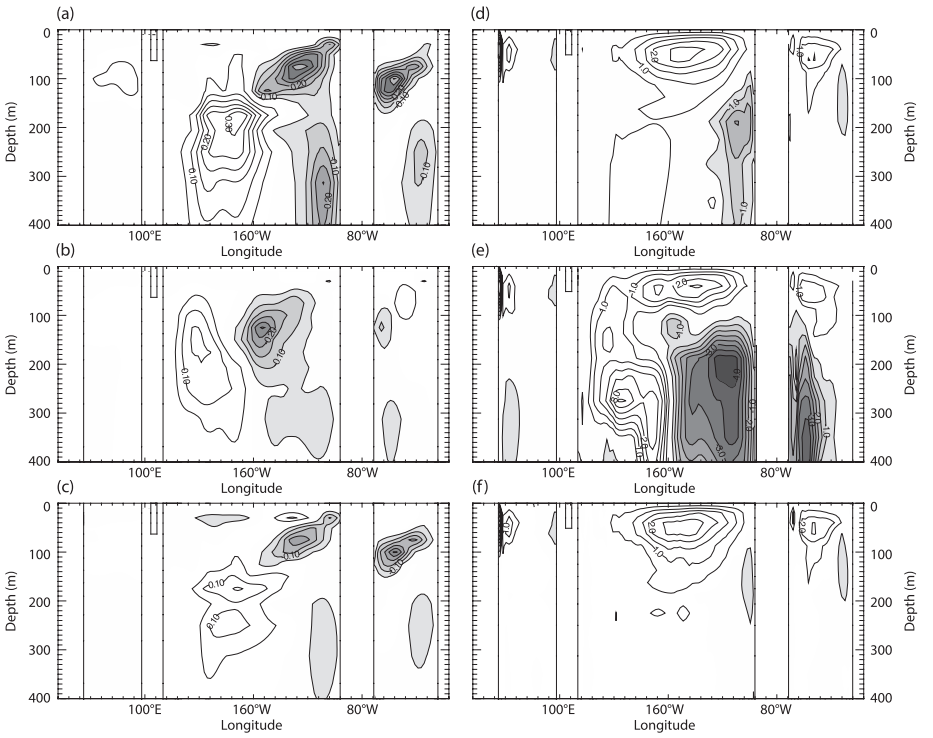

Figure 10. Equatorial longitude-depth sections of mean temperature increment (left) and vertical velocity (right) in data assimilation experiments without bias correction (a, d), with bias correction applied directly on temperature and salinity (b, e), and with bias correction applied on pressure gradient (c, f). At the equator the adiabatic pressure gradient bias correction is preferred, since the direct correction of temperature and salinity degrades the circulation (from Balmaseda et al 2007).

of recent data into the past, assuming that the dominant component of model error is stationary (or just depends on the seasonal cycle). The online component ($\mathbf{b}'$) acts as a correction to the term ($\bar{\mathbf{b}}$); it is modeled as an autoregressive model with a memory term α (typically with a multiyear timescale), and it is updated with information derived from the assimilation increments according to Eq. (1–2).

The following section provides an example of ocean data assimilation of seasonal forecasts. It summarizes the current ECMWF reanalysis system (ORAS4), and provides a specific example for the application of the bias correction algorithm.

e. An example of an ocean initialization system: the ECMWF ORAS4

The ORAS4 (Ocean Reanalysis System 4; Mogensen et al. 2012; Balmaseda et al. 2013) provides ocean initial conditions for the ECMWF Seasonal Forecasting System 4 (Molteni

et al. 2011). ORAS4 has been produced by combining, every 10 days, the output of an ocean model forced by atmospheric reanalysis fluxes with quality-controlled ocean observations.

ORAS4 uses the NEMO ocean model (Madec 2008) with a horizontal resolution of approximately one degree, and the NEMOVAR (Daget et al. 2009; Mogensen et al. 2012) data assimilation system in its 3Dvar configuration. The ocean model is forced by daily atmospheric-derived surface fluxes of solar radiation, total heat flux, evaporation-minus-precipitation, and surface wind stress. These are from the ERA-40 reanalysis (Uppala et al. 2005) from September 1957 to December 1989, and ERA-Interim (Dee et al. 2011) thereafter. The heat fluxes are adjusted using a strong relaxation to gridded SST products. As the relaxation coefficient is equivalent to about a two- to three-day timescale over a depth of 10 m, this is a strong constraint to track observed SSTs. The freshwater flux is also adjusted using ocean observations: i) globally, by constraining the global model sea-level changes to the variations in the altimeter-derived value, and ii) locally, via a weak relaxation (one-year timescale) to a monthly climatology of surface salinity.

The analysis cycle is 10 days, and it is summarized in Eq. (1). The state $\mathbf{x}^f$ produced by integrating the NEMO model forced by daily surface fluxes and relaxed to SST is bias-corrected with a bias term $\mathbf{b}^f$ (as in Eq. 1–3) to produce the first guess ($\mathbf{x}^f + \mathbf{b}^f$), which is contrasted with each available observation $\mathbf{y}$ at its appropriate time and position via the observation operator $\mathbf{H}$.

The observations consist of temperature and salinity (T/S) profiles from the Hadley Centre's EN3 data collection (Ingleby and Huddleston 2007), which includes XBTs, CTDs, TAO/TRITON/PIRATA/RAMA moorings, Argo profiles, and Autonomous Pinniped Bathythermograph (APBs or elephant seals, T/S). Altimeter-derived along-track sea-level anomalies from AVISO are also assimilated. The quality-controlled model-observations departures are passed to the 3D-Var minimization to compute the optimal assimilation increment, which is applied as a tendency forcing during a second model integration spanning the same time window as for the first guess, thus producing the analysis $\mathbf{x}^a$. The analysis increments are also used to modify the bias first guess $\mathbf{b}^f$, producing an update of the bias state $\mathbf{b}^a$ (Eq. 1 and 2). The model bias correction scheme (Eq. 3) comprises an adaptive component ($\mathbf{b}'$), estimated online from previous observations, and a-priori component ($\bar{\mathbf{b}}$), derived offline from a monthly climatology of model errors estimated during the data-rich Argo period (2000–2008) and applied to ORAS4 from the beginning of the record.

Figure 11 shows the offline temperature bias correction term used in ORAS4, for the depth range 300–700 m. Clearly visible are the corrections needed by the western boundary currents, which the low resolution of the model prevents resolving properly.

ORAS4 consists of five ensemble members spanning the period 1958 to present. The five ensemble members sample plausible uncertainties in the wind forcing, observation coverage, and the deep ocean. To sample the uncertainty in the deep ocean, five different ocean states from an ocean model integration sampled at five-year intervals from 1960 to 1980 are used to initialize each of the ensemble members of ORAS4. The impact of this

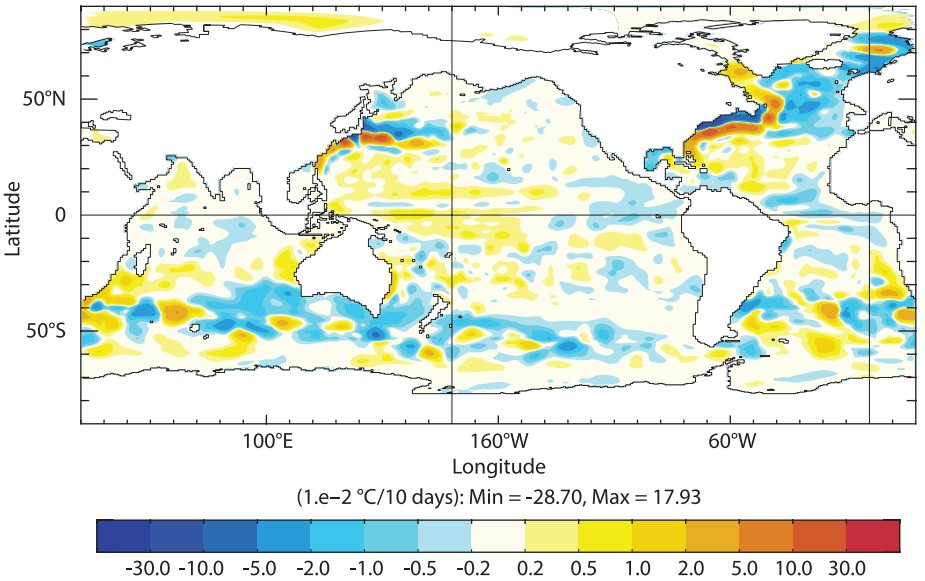

Figure 11. Temperature bias correction term estimated offline from Argo for the depth range 300 to 700 meters. The term is used in ORAS4 as a tendency term in the evolution equation to correct the first guess. Units are degrees Celsius per hour (from Balmaseda et al. 2013).

step in the deep ocean is noticeable in the first 20 years of the reanalysis (see Figure 4 of Balmaseda et al. 2013).

In addition to sampling the uncertainty in reanalysis spin-up, the ensemble of ORAS4 also samples uncertainty in the wind stress and its impact on the ocean state. Monthly wind perturbations are added to the forcing fields during the production of the ocean reanalysis, while the ocean model is integrated forward, resulting in spread in the ocean subsurface, especially along the thermocline. The perturbations are simply differences of monthly anomalies (i.e., the differences in the mean seasonal cycle have been removed) between two data sets—ERA-40 and NCEP-CORE atmospheric reanalyses—for the period 1958–2002 (40 years). The differences are the basis for a repository of perturbations (consisting of 40 realizations per calendar month), from which perturbations are randomly selected for a given month. The perturbations are applied with a plus/minus sign, so the ensemble is centered (see Vialard et al. 2003).

Quality improvements in ORAS4 relative to earlier ocean reanalyses are due to the use of atmospheric surface fluxes from the ERA-Interim reanalysis, various improvements in ocean modeling and data assimilation, and more comprehensive and improved quality-controlled ocean data sets, including important corrections to the ocean observations. Balmaseda et al. (2013) evaluated ORAS4 using different metrics, including comparison with observed ocean currents, RAPID-derived transports, sea-level gauges, and GRACE-derived

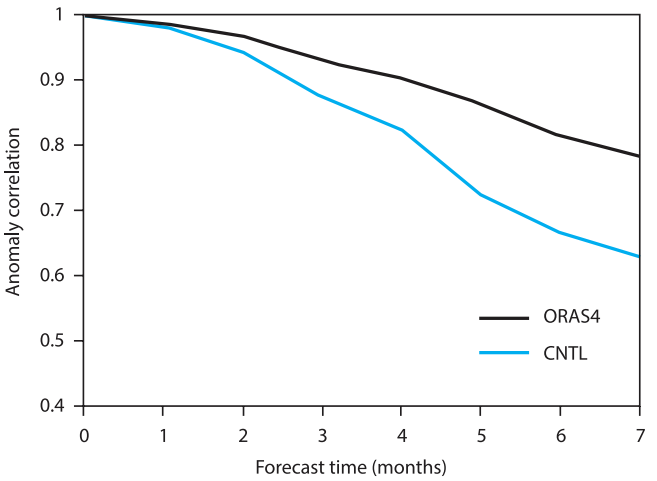

Figure 12. The effect of initialization (ORAS4 compared to CNTL) in seasonal forecast skill. Anomaly correlations between forecast and observed SST anomalies averaged over the central equatorial Pacific (170°W–130°W, 5°S–5°N) are shown as a function of forecast lead time. The CNTL is a nonassimilative ocean run, which is only constrained by surface fluxes and SST (from Balmaseda et al. 2013).

bottom pressure. They showed that compared to a control ocean model simulation, ORAS4 improves the fit to observations and the interannual variability, and consistently results in improved seasonal forecast skill of SST. Figure 12 shows the impact of assimilating data in the skill of the seasonal forecasts. It shows the anomaly correlation skill of seasonal forecast of SST in the central equatorial Pacific (170W–130W, 5S–5N), initialized from ORAS4 (black) and from an equivalent run (CNTL) where neither subsurface observations nor altimeter data are assimilated, so that it is only constrained by surface fluxes and SST. The statistics comprise results from seasonal forecasts initialized in 40 start dates, three months apart, over the period 1989–2008. For each date, an ensemble of five members is integrated for seven months (from Balmaseda et al. 2013).

5. Impact on forecast skill

The skill of seasonal forecasts is often used to gauge the quality of the ocean initial conditions. This may not always be appropriate, since the quality of the coupled model is also important—if the major source of forecast error comes from the coupled model, improvements in ocean initial conditions would have little impact on forecast skill. This is something to bear in mind when interpreting results of the impact of ocean data assimilation on seasonal forecasts. Several studies have demonstrated the benefit of assimilating ocean data on the prediction of ENSO (Alves et al. 2004 and Balmaseda et al. 2008, among others). The benefits are less clear in other areas, such as the equatorial Atlantic, where model errors

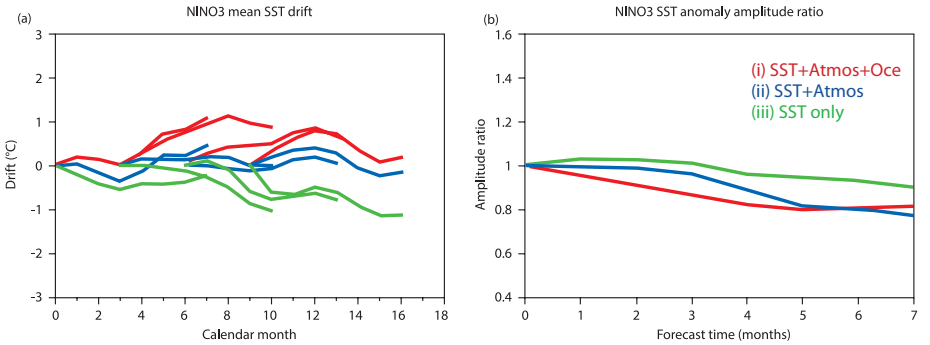

Figure 13. Impact initialization strategy in forecast drift (a) and normalized interannual variability (b). Green uses SST (sea surface temperature) only, blue uses SST and atmospheric reanalyses, and red uses SST, atmospheric reanalyses, and ocean observations. All the forecasts have been conducted with the same coupled model (from Balmaseda and Anderson 2009). Although data assimilation improves skill scores (not shown), it can also introduce initialization shocks. See text for further discussion.

are large and there is no long history of moored observations, as in the Pacific. Ultimately, the impact of initialization in a seasonal forecasting system will depend on the quality of the coupled model (Stockdale et al. 2006; Balmaseda and Anderson 2009).

The contribution of the different sources of observational information in seasonal forecast skill has been quantified by Balmaseda and Anderson (2009). They used a previous version of the ECMWF seasonal forecasting system (S3) to evaluate three different initialization strategies, each of which uses different observational information. Strategy (i) uses ocean, atmospheric, and SST information; strategy (ii) uses atmospheric and SST information; and strategy (iii) uses only SST, as in Keenlyside et al. (2008). The results from the different strategies are summarized in Figure 13: strategy (i) is labelled as SST+atmos+ocean (red curves), strategy (ii) is labelled as SST+atmos (blue curves), while strategy (iii) is labelled as SST-only (green curves). In strategy (i) the coupled system thus starts close to the observed state but it is not obvious that this leads to the most skillful forecasts as the method can have undesirable initialization shocks. Strategy (iii) can reduce the initialization shock since the atmospheric and ocean models will be in closer balance at the start of the coupled integrations. Results show that the initialization strategy has an impact on both the mean state and the interannual variability of coupled forecasts (Fig. 13). The left panel of Figure 13 shows how the model bias develops as a function of lead times for forecasts initialized at four different starting dates (January, April, July, and October) for a given region in the equatorial eastern Pacific. The results show that the way the model bias develops does indeed depend on the initialization procedure. (The forecast coupled model is the same in all the experiments.) The forecasts initialized with SST only develop a strong cold bias, while the ones initialized with SST+atmos+ocean observations develop a warm bias. The right panel

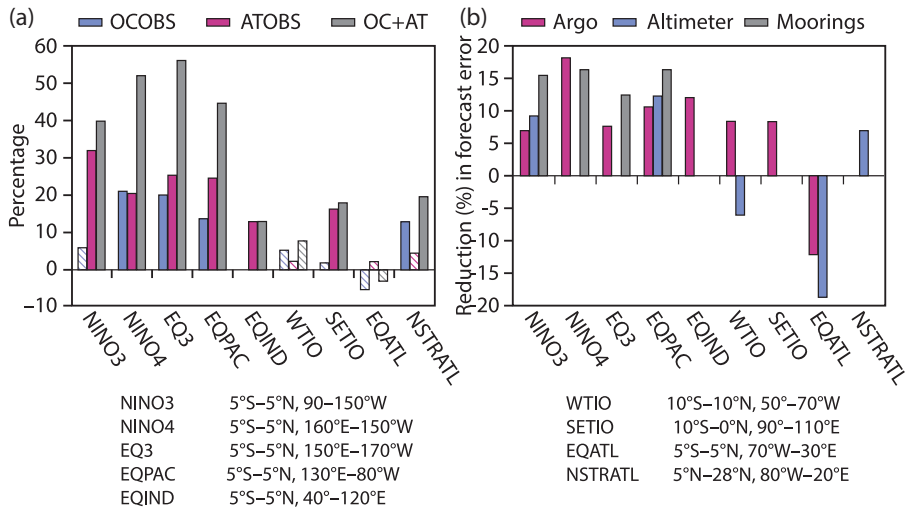

Figure 14. Effect of observations in forecast skill for different regions, as measured by the reduction in mean absolute error for the forecast range. (a) Ocean observations (OCOBS), atmospheric observations (ATOBS), and both; for the forecast range 1–3 months, period 1987–2008. (b) Effect of Argo, altimeter, and moorings for the period 2001–2006.

in Figure 13 shows the ratio between the interannual SST variability of the model versus observations, illustrating that the initialization also affects the amplitude of the interannual variability of the forecast anomalies. The interpretation of these results is fully discussed in Balmaseda and Anderson (2009).

The relation between initialization shock and forecast skill is also discussed in Balmaseda and Anderson (2009). They show that, in the particular system of study, initialization shock does not preclude forecast skill, and the most skillful forecasts are those obtained when the initial conditions are closer to the "real ocean state," even if this causes sizable adjustment processes.

The three experiments above can also be seen as observing system experiments. Differences between strategies (i) and (ii) are indicative of the impact of ocean observations, and those between (ii) and (iii) are indicative of the impact of the atmospheric observations that were used to produce the atmospheric reanalyses. Figure 14 (upper panel) shows the relative reduction in the monthly mean absolute error (MAE) resulting from adding information from the ocean and/or atmospheric observations for the one- to three-month forecast range in the regions defined in the inset table. Observational information has the largest impact in the western Pacific (EQ3), where the combined information of ocean and atmospheric observations can reduce the MAE more than 50%. With the exception of the equatorial Atlantic (EQATL), the best scores are achieved by strategy (i). This means that for the ECMWF system, the benefits of ocean data assimilation and the use of fluxes from

atmospheric (re)analyses more than offset possible problems arising from initialization shock. Seasonal forecast skill can also be used to evaluate the ocean observing system.

The lower panel of Figure 14 shows the impact on forecast skill of Argo, moorings, and altimeters. The statistics have been calculated only for the (rather short) Argo period 2001–2006 and so the impacts are best considered as indicative rather than definitive. The figure shows that no observing system is redundant. Argo has a dominant impact in the western Pacific (NINO4) and equatorial Indian Ocean. Argo is the only observing system with a significant positive impact on the WTIO (Western Tropical Indian Ocean) and SETIO (Southeastern Tropical Indian Ocean) regions. The information from the moorings is still dominant in most of the equatorial Pacific, although in the NINO4 region it is less important than that from Argo. Meanwhile, altimetry has a significant positive impact in the equatorial Pacific, and is the only observing system with positive impact in the north subtropical Atlantic. Again, for this period, all the observing systems have a negative impact on the EQATL region, probably because the assimilation method is deficient here. The impact of the TAO/TRITON array and Argo float data has also been evaluated with the Japan Meteorological Agency (JMA) seasonal forecasting system (Fuji et al. 2008) by conducting data retention experiments for 2004–2007. The results (not shown) are consistent with those above, indicating that TAO/TRITON data improves the forecast of SST in the eastern equatorial Pacific (NINO3, NINO4), and that Argo floats are essential observations for SST prediction in the tropical Pacific and Indian Oceans.

Ultimately the impact of the observations in the seasonal forecasts depends strongly on the ability of the assimilation to retain the relevant information. Fujii et al. (2015) compare results from observing system experiments conducted in several systems, and they show, for instance, that the impact of TAO/TRITON in the ocean estate varies a great deal among data assimilation systems.

6. Assessment of initialization strategies

Systematic model error leads to difficulties in the forecasting process. This mostly happens when transferring of information between the observation space and the model space is required, namely the initialization and the issuing of the forecast. At the initialization stage, information needs to be transferred from observations to model space. When issuing the forecast, the model output needs to be calibrated using reliable information about the real world. In numerical weather prediction (NWP), the forecast typically covers 1–15 days, and, because of the relatively short forecast time, the difference between model and observed climatologies can be ignored (i.e., the model error is neglected). At longer lead times (monthly, seasonal, and decadal timescales) the systematic model error cannot be ignored and strategies for accounting for model error are needed. Below is a description of three different strategies (Full Initialization, Anomaly Initialization, and Flux Correction), followed by a summary of the results from Magnusson, Alonso-Balmaseda, Corti et al. (2012), who produced an assessment of the different forecast strategies.

a. Full initialization

The full initialization strategy follows the NWP approach, i.e., the model (in this case the ocean model) is initialized from an ocean analysis performed via data assimilation. The analysis is a combination of the latest observations together with a short-range forecast (typically 10 days). By continuously using the information from the observations, the analysis state is kept close to the attractor of the nature (although in poorly observed areas, a difference could still be present), and often an explicit bias correction is used. During the forecast, the state of the model will diverge from the state of nature both due to the loss of predictability, related to high sensitivity to initial conditions, and development of systematic errors. At long lead times (monthly, seasonal, and decadal timescales), the model bias is often large compared with the random component of the forecast error. In these cases the model bias cannot be neglected and the strategy for accounting for the model systematic error is the a-posteriori removal of it. The bias is corrected by applying a lead-time-dependent bias correction in postprocessing. The bias correction is also made dependent on the seasonal cycle. This is the strategy commonly used in monthly and seasonal forecasts (Stockdale 1997). For example, in an operational seasonal forecast issued every month with a typical lead time of seven months, the estimation of 84 (7×12) bias correction terms is needed to account for all lead times and all starting dates. The robust estimation of this large number of bias fields requires a large data set of hindcasts (retro-perspective forecasts).

This strategy will fail if the bias is nonstationary, and can lead to suboptimal forecast skill. The nonstationarity of the bias may be due to nonstationary errors on the initial conditions (Kumar et al. 2012), or to flow-dependent bias arising from the nonlinear nature of the system (Balmaseda and Anderson 2009). Generally speaking, if the systematic error is large enough, the nonlinear terms will become nonnegligible and therefore a mere linear calibration process will be insufficient.

The full initialization strategy may also be affected by the so-called initialization shock, a term referring to rapid adjustment processes caused by the imbalance between the initial conditions and the forecast model. This can occur if the forecast model is different from the initialization model, or if the initialization does not preserve physical constraints. In the case of the full initialization, the imbalance can be induced by the fact that the forecast model is coupled, while the initialization has taken place in uncoupled mode.

Figure 15 from Kumar et al. (2012) illustrates the nonstationarity of the SST bias in the NCEP CFS model (version 2). It shows that the bias depends on the lead time and the seasonal cycle (accounted for in the a-posteriori removal of the bias), but it also shows nonstationary behavior at interannual timescales (not accounted for in the strategy) leading to degradation of forecast skill. The figure also hints at the presence of initialization shock, since in some cases the bias at eight-month lead time is smaller than the bias at one-month lead time.

Although the full initialization strategy follows the NWP approach, its practical implementation in monthly and seasonal timescales differs from NWP in three main elements: i) the model bias is explicitly corrected during the data assimilation, ii) the forecast error

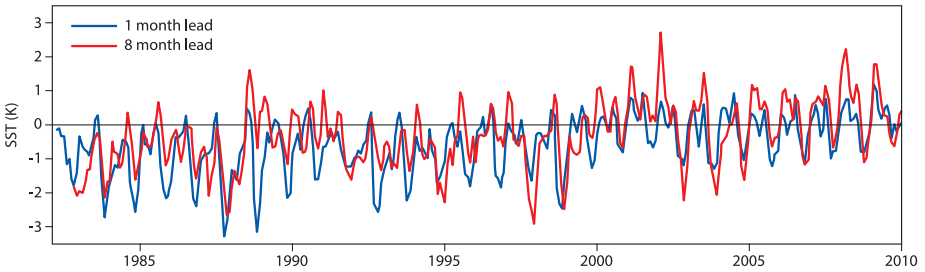

Figure 15. Time evolution of the sea surface temperature (SST) forecast bias in the NCEP CFS version 2. The figure shows the bias at one-month and eight-month lead times, and it illustrates the nonstationarity of the bias (from Kumar et al. 2012).

needs to be removed a-posteriori, and iii) the forecast and initialization models are not the same. The latter can lead to initialization shock, and it prevents using the bias information obtained by the data assimilation in the forecast stage.

b. Anomaly initialization

The idea of using anomaly initialization is to avoid the nonstationary model drift and initialization shock, by initializing the model around its own climatology. The procedure is to calculate anomalies in the observations, with respect to the observations climatology, and add such anomalies to the climate of the model. The method has been very popular for decadal forecasts (Smith et al. 2007, among others).

Eq. (4) describes the anomaly initialization procedure, using the same naming convention as in Eq. (1):

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}[\mathbf{y} - \bar{\mathbf{y}} - \mathbf{H}(\mathbf{x}^f - \bar{\mathbf{x}})] \tag{4}$$

Two new variables ($\bar{\mathbf{y}}$ and $\bar{\mathbf{x}}$) are introduced in Eq. (4) to denote the observed and model multiyear climatology respectively. Therefore, while in Eq. (1) the assimilation increment was proportional to the differences between model and observed *full states*, in Eq. (4) the assimilation increment is proportional to the differences between model and observed *anomalies*. The anomaly initialization is often called bias-blind data assimilation: the presence of bias is acknowledged, but there is no attempt to correct it. In the full initialization (Eqs. 1–2), the bias was also corrected by the data assimilation.

The procedure of the anomaly initialization is not without problems. The estimation of the anomaly requires the knowledge of the observed climatology ($\bar{\mathbf{y}}$). This introduces two kinds of difficulties. On one hand, it is important that the sampling period used for the observed climatology is consistent with that used for the model climatology $\bar{\mathbf{x}}$. For instance, a model climatology estimated for the preindustrial era should not be used for the anomaly initialization of decadal forecast post-1960s, with an observed climatology estimated during the period 1970–2005. The other kind of problem relates to defining the

climatology of new or sporadic observations. For instance, some regions such as southern oceans had not been observed prior to the advent of Argo. Most of the deep ocean has only been observed sporadically with cruise data, and there is not enough information to define a long-term climatology. To avoid this problem, the anomaly initialization strategy of the ocean often uses gridded fields from existing ocean reanalysis. In this way, it turns an initial weakness into a good advantage, since it means that different coupled modeling groups can initialize their decadal forecasts with external ocean reanalysis, without the need of having to develop data assimilation systems for their own models.

The problem with nonlinear interaction between mean state and anomaly is even more acute in the anomaly than in full initialization, since the mean error is fully developed during the coupled model integrations. Although it is often claimed that the anomaly initialization avoids initialization shock, this is by no means guaranteed since it depends how the anomaly is assimilated into the model. The structure of the observed anomaly may not be consistent with the model mean state. For example, when anomalies are associated with displacement of sharp fronts or gradients that are in different locations in model and observations, simply adding the anomaly can lead to rapid adjustment processes. (Examples are anomalies associated with vertical displacements of the equatorial thermocline, Gulf Stream, or sea-ice edge.)

A more interesting advantage of the anomaly initialization, which is often not discussed, is the avoidance of model drift. By avoiding model drift, the a-posteriori correction of the forecast does not require the bias dependence on the forecast lead time (so typically only the 12-month climatology of the bias is required), and the bias estimators can be more robust. This is more relevant for decadal forecast ranges, when it is also more computationally expensive to conduct the calibrating hindcasts. The procedure requires, however, a long integration to estimate the model climatology.

c. Flux correction

It is clear from a variety of studies that strong nonlinear interactions between mean state and anomaly are at play in the coupled model forecasts. Model improvement is the ultimate way of reducing model biases. However, this is a slow process, especially if the systematic errors are related to model resolution (as in the case of the correct Gulf Stream). A temporary solution, until the problems in the model are detected and solved, is to compensate for the systematic errors by applying empirical corrections.

One specific correction is the so-called flux correction, applied only in the coupling between the atmosphere and the ocean. The aim of this strategy is to avoid (or limit) the model drift by adding a correction term to the model during the simulation, to avoid nonlinear interactions between model mean state and variability. In this strategy the empirical correction of the forecast is done during the model integration rather than only in the final calibration phase. Magnusson and colleagues investigated this topic, using both momentum-flux correction and a combination of momentum and heat-flux correction (Magnusson,

Alonso-Balmaseda and Corti et al. 2012; Magnusson, Alonso-Balmaseda and Molteni 2012). The following section provides a brief summary of their findings.

d. Assessment of initialization strategies

Full initialization, anomaly initialization, and flux correction have been implemented in the ECMWF coupled forecasting system. The three strategies have been evaluated at seasonal and decadal timescales. The results are presented along with the practical implications of the different strategies in Magnusson, Alonso-Balmaseda and Molteni (2012) and Magnusson, Alonso-Balmaseda and Corti et al. (2012).

Magnusson, Alonso-Balmaseda and Molteni (2012) investigate the impact of the mean state on the properties of ENSO in a set of coupled decadal integrations, where the mean state and its seasonal cycle have been modified by applying flux correction to the momentum-flux and a combination of heat and momentum fluxes. They show that correcting the mean state and the seasonal cycle improves the amplitude of SST interannual variability and also the penetration of the ENSO signal into the troposphere and the spatial distribution of the ENSO teleconnections. An analysis of a multivariate PDF of ENSO shows clearly that the flux correction affects the mean, variance, skewness, and tails of the distribution. The changes in the tails of the distribution are particularly noticeable in the case of precipitation, showing that without the flux correction the model is unable to reproduce the frequency of large events. For the interannual variability, the momentum-flux correction alone has a large impact, while the additional heat-flux correction is important for the teleconnections.

Magnusson, Alonso-Balmaseda, Corti, and their colleagues (2012) show that full initialization results in a clear model drift towards a colder climate (although for other models the drift could be towards a warmer climate). The anomaly initialization is able to reduce the drift, by initializing around the model mean state. However, the erroneous model mean state results in degraded seasonal forecast skill. The best results on the seasonal timescale are obtained using momentum-flux correction, mainly because it avoids the positive feedback responsible for a strong cold bias in the tropical Pacific. These results are illustrated in Figure 16. It is likely that these results are model dependent: the coupled model used here shows a strong cold bias in the Central Pacific, resulting from a positive coupled feedback between winds and SST. At decadal timescales it is difficult to determine whether any of the initialization strategies is superior to the other. Similar conclusions are reached by Smith et al. (2013), using the MetOffice forecasting system.

7. Conclusions

It has been shown that seasonal forecasting of SST is an initial condition problem. In order to initialize the upper thermal structure of the ocean it is important to assimilate ocean observations.

Assimilation of ocean observations reduces the large uncertainty (error) in the ocean model state due to the uncertainty (error) of forcing fluxes. Using information from

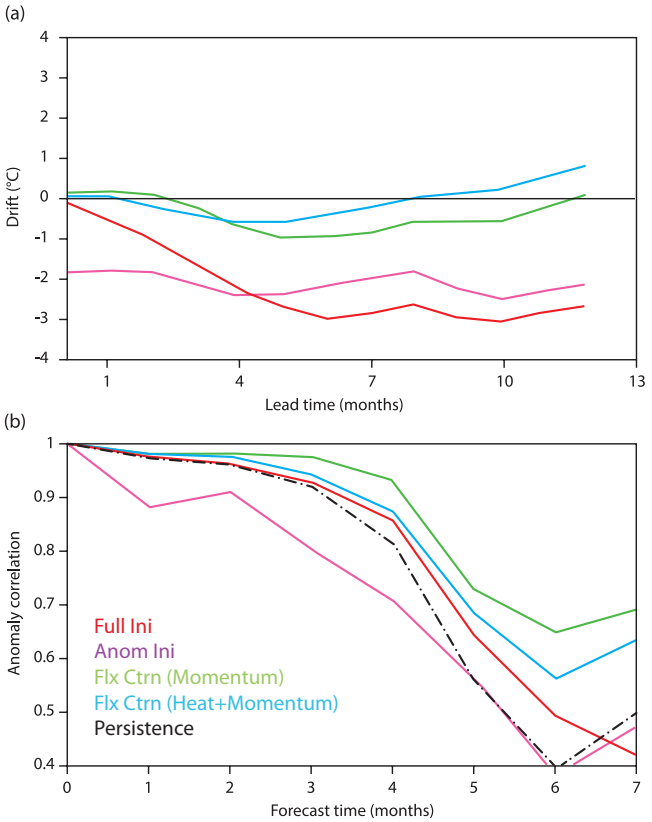

Figure 16. Forecast drift in sea surface temperature (a) and skill in precipitation (b) in the central Pacific from different forecast strategies: full initialization (red), anomaly initialization (purple), momentum flux correction (green) and momentum and heat flux correction (blue). For comparison, the skill of persistence is also shown (dashed black line). The best skill is achieved by the momentum flux correction (from Magnusson, Alonso-Balmaseda and Corti et al. 2012).

SST, surface fluxes from atmospheric reanalyses, subsurface temperature and salinity, and altimeter-derived sea-level anomalies is instrumental in the ocean initialization and may improve forecast skill.

Because seasonal forecasts need a-posteriori calibration, a sample of the model performance over a long enough period is required; this is obtained by performing a series of coupled hindcasts during some historical period. A historical record of hindcasts is also needed for skill assessment. Ocean reanalyses with reliable representation of the interannual variability are then required to initialize these hindcasts.

The most common initialization strategy is the so-called *full initialization*, where the data assimilation corrects the ocean model mean state, as well as the variability. In order

to avoid spurious variability associated with changes in the observing system, consistent ocean reanalysis requires an explicit treatment of the bias during the initialization procedure. The bias estimation obtained during the initialization procedure could in principle be used to correct model error during the forecasts. However, this is not possible when the full initialization is conducted in uncoupled mode, which currently is the most common practice. The separate initialization of the ocean and atmosphere systems can also lead to initialization shock during the forecasts. A more balanced "coupled" initialization is desirable, but it remains challenging.

The *anomaly initialization* is more frequently used in decadal forecasts, but shows weaker performance than the *full initialization*, especially at seasonal timescales. In decadal forecasts the anomaly initialization shows practical advantages regarding the computational cost of the calibration data set.

Systematic model error remains a difficult problem for seasonal forecasting and climate predictions. An error in the mean state could affect the variability of the system. Results indicate that the current forecast practices of removing the forecast bias a-posteriori or anomaly initialization are by no means optimal, since they cannot deal with the strong nonlinear interactions. A consequence of the results presented here is that the predictability on annual time-ranges could be higher than currently achieved. The conclusion from the ECMWF model that the correction of the model mean state by some sort of flux correction leads to better forecasts needs to be assessed in other prediction systems. This may also lead to further model improvements since flux correction may be a powerful tool for diagnosing coupled model errors and predictability studies.

REFERENCES

Alves, O., M. Balmaseda, D. Anderson, and T. Stockdale. 2004. Sensitivity of dynamical seasonal forecasts to ocean initial conditions. Q. J. Roy. Meteorol. Soc., *130*, 647–668.

Balmaseda, M. A. and D. Anderson. 2009. Impact of initialization strategies and observations on seasonal forecast skill. Geophys. Res. Lett., *36*, L01701.

Balmaseda, M. A., D. Dee, A. Vidard, and D. L. T. Anderson. 2007. A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. Q. J. Roy. Meteorol. Soc., *133*, 167–179.

Balmaseda, M. A., Y. Fujii, O. Alves, T. Awaji, D. Behringer, N. Ferry, T. Lee et al. 2010. Initialization for Seasonal and Decadal Forecasts *in* Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21–25 September 2009. J. Hall, D. E. Harrison, and D. Stammer, eds. ESA Publication WPP-306. doi: 10.5270/OceanObs09.cwp.02

Balmaseda, M. A., Y. Fujii, O. Alves, T. Lee, M. Rienecker, T. Rosati, D. Stammer et al. 2010. Role of ocean observations in an end-to-end seasonal forecasting system *in* Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 1), Venice, Italy, 21–25 September 2009. J. Hall, D. E. Harrison, and D. Stammer, eds. ESA Publication WPP-306. doi: 10.5270/OceanObs09.pp.03

Balmaseda, M. A., K. Mogensen, and A. T. Weaver. 2013. Evaluation of the ECMWF ocean reanalysis system ORAS4. Q. J. R. Meteorol. Soc., *139*, 1132–1161. doi: 10.1002/qj.2063

Balmaseda, M. A., A. Vidard, and D. Anderson. 2008. The ECMWF ORA-S3 ocean analysis system. Mon. Weather Rev., *136*, 3018–3034.

Behera, S. K, J.-J. Luo, S. Masson, T. Yamagata, P. Delecluse, S. Gualdi and A. Navarra. 2005. Paramount impact of the Indian Ocean dipole on the east African short rains: A CGCM study. J. Clim., *18*, 4514–4530.

Bell, M. J., M. J. Martin, and N. K. Nichols. 2004. Assimilation of data into an ocean model with systematic errors near the equator. Q. J. Roy. Meteorol. Soc., *130*, 873–893.

Bjerknes, J. 1969. Atmospheric teleconnections from the equatorial Pacific. Mon. Weather Rev., *97*, 163–172.

Cai, W., A. Sullivan, T. Cowan, J. Ribbe, and G. Shi. 2011. Simulation of the Indian Ocean Dipole: A relevant criterion for selecting models for climate projections, Geophys. Res. Lett., *38*, L03704. doi: 10.1029/2010GL046242

Daget, N., A. T. Weaver, and M. A. Balmaseda. 2009. Ensemble estimation of background-error variances in a three dimensional variational data assimilation system for the global ocean. Q. J. Roy. Meteorol. Soc., *135*, 1071–1094.

Dee, D. P. 2005. Bias and data assimilation. Q. J. Roy. Meteorol. Soc., *131*, 3323–3343. doi: 10.1256/qj.05.137

Dee, D. P. and A. M. Da Silva. 1998. Data assimilation in the presence of forecast bias. Q. J. Roy. Meteorol. Soc., *124*, 269–295. doi: 10.1002/qj.49712454512

Dee, D. P. and R. Todling. 2000. Data assimilation in the presence of forecast bias: The GEOS moisture analysis. Mon. Weather Rev., *128*, 3268–3282.

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

ECMWF (European Centre for Medium-Range Weather Forecasts). 2005. SST and wind stress perturbations for seasonal and annual simulations. December 12, 2005; accessed May 18, 2017. https://www.ecmwf.int/sites/default/files/docu_perturbations.pdf

Folland, C. K., A. W. Colman, D. P. Rowell, and M. K. Davey. 2001. Predictability of northeast Brazil rainfall and real-time forecast skill, 1987–98. J. Clim., *14*, 1937–1958.

Freeland, H. J., D. Roemmich, S. L. Garzoli, P-Y LeTraon, M. Ravichandran, S. Riser, V. Thierry et al. 2010. Argo—A decade of progress *in* Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21–25 September 2009. J. Hall, J., D. E. Harrison and D. Stammer, D., eds., ESA Publication WPP-306. doi: 10.5270/OceanObs09.cwp.32

Fujii, Y., J. Cummings, Y. Xue, A. Schiller, T. Lee, M. A. Balmaseda, E. Rémy et al. 2015. Evaluation of the Tropical Pacific Observing System from the ocean data assimilation perspective. Q. J. Roy. Meteorol. Soc., *141*, 2481–2496. doi: 10.1002/qj.2579

Fujii, Y., T. Yasuda, S. Matsumoto, M. Kamachi, and K. Ando. 2008. Observing System Evaluation (OSE) using the El Niño forecasting system in Japan Meteorological Agency. Proceedings of the Oceanographic Society of Japan 2008 fall meeting (in Japanese).

Giannini, A., R. Saravanan, and P. Chang. 2003. Oceanic forcing of Sahel rainfall on interannual to interdecadal timescales, Science, *302*, 1027–1030.

Goddard, L. and M. Dilley. 2005. El Niño: Catastrophe or opportunity. J. Clim., *18*, 651–665.

Goddard, L. and N. E. Graham. 1999. The importance of the Indian Ocean for simulating precipitation anomalies over Eastern and Southern Africa. J. Geophys. Res., *104*, 19099–19116.

Goni, G., D. Roemmich, R. Molinari, G. Meyers, C. Sun, T. Boyer, and M. Baringer. 2010. The ship of opportunity program *in* Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21–25 September 2009, J. Hall, D. E. Harrison and D. Stammer, eds., ESA Publication WPP-306. doi: 10.5270/OceanObs09.cwp.35

Ingleby, B. and M. Huddleston. 2007. Quality control of ocean temperature and salinity profiles—historical and real-time data. J. Mar. Syst., *65*, 158–175.

Ji, M., A. Leetmaa, and J. Derber. 1995. An ocean analysis system for seasonal to interannual climate studies. Mon. Weather Rev., *123*, 460–481.

Jin, F.-F. 1997. An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model. J. Atmos. Sci., *54*, 811–829. http://dx.doi.org/10.1175/1520-0469(1997)054<Assimilation of ocean0811:AEORPF>2.0.CO;2

Keenlyside, N., M. Latif, J. Jungclaus, L. Kornblueh, and E. Roeckner. 2008. Advancing decadal-scale climate prediction in the North Atlantic Sector. Nature, *453*, 84–88.

Klein, S. A., B. J. Soden, and N. C. Lau. 1999. Remote sea surface temperature variations during ENSO: Evidence for a tropical atmospheric bridge. J. Clim., *12*, 917–932.

Kumar, A., M. Chen, W. Wang, Y. Xue, C. Wen, B. Marx, and L. Huang. 2012. An analysis of the non-stationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) Version 2. Mon. Weather Rev., *140*, 3003–3016. http://dx.doi.org/10.1175/MWR-D-11-00335.1

Lau, N.-C. and M. J. Nath. 1996. The role of the "atmospheric bridge" in linking tropical Pacific ENSO events to extratropical SST anomalies. J. Clim., *9*, 2036–2057.

Madec, G. 2008. NEMO Reference Manual, Ocean Dynamics Component. NEMO-OPA. Preliminary version. Note du Pôle de modélisation 27, Institut Pierre-Simon Laplace (IPSL), France.

Magnusson, L., M. Alonso-Balmaseda, S. Corti, F. Molteni, and T. Stockdale. 2012. Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors. Clim. Dynam., *41*, 2393–2409. doi: 10.1007/s00382-012-1599-2

Magnusson, L., M. Alonso-Balmaseda, and F. Molteni. 2012. On the dependence of ENSO simulation on the coupled model mean state. Clim. Dynam., *41*, 1509–1525. doi: 10.1007/s00382-012-1574-y

McPhaden, M. J., K. Ando, B. Bourles, H. P. Freitag, R. Lumpkin, Y. Masumoto, and V. S. N. Murty et al. 2010. The global tropical moored buoy array *in* Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21–25 September 2009, J. Hall, D. E. Harrison, and D. Stammer, eds., ESA Publication WPP-306. doi: 10.5270/OceanObs09.cwp.61

Mogensen, K., M. Alonso Balmaseda, and A. Weaver. 2012. The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4. European Centre for Medium Range Weather Forecasts Technical Memorandum No. 668, 59 pp.

Molteni, F., T. Stockdale, M. Balmaseda, G. Balsamo, R. Buizza, L. Ferranti, L. Magnusson et al. 2011. The new ECMWF seasonal forecast system (System 4). European Centre for Medium Range Weather Forecasts Technical Memorandum 656, 51 pp.

Palmer, T. N. and D. L. T. Anderson. 1994. The prospects of seasonal forecasting: a review paper. Q. J. Roy. Meteorol. Soc. *120*, 755–793.

Rodwell, M. and C. K. Folland. 2002. Atlantic air–sea interaction and seasonal predictability. Q. J. Roy. Meteorol. Soc., *128*, 1413–1443.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata. 1999. A dipole mode in the tropical Indian Ocean. Nature, *401*, 360–363.

Smith, D., A. Cusack, A. Colman, C. Folland, and G. Harris. 2007. Improved surface temperature predictions for the coming decade from a global circulation model. Science, *317*, 796–799.

Smith, D. M., R. Eade, and H. Pohlmann. 2013. A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. Clim. Dynam., *41*, 3325–3338. doi: 10.1007/s00382-013-1683-2

Stockdale, T. 1997. Coupled ocean-atmosphere forecast in the presence of climate drift. Mon. Weather Rev., *125*, 809–818.

Stockdale, T. N., M. A. Balmaseda, and A. Vidard. 2006. Tropical Atlantic SST prediction with coupled ocean–atmosphere GCMs. J. Clim., *19*, 6047–6061. doi: 10.1175/JCLI3947.1

Uppala, S. M., P. W. KÅllberg, A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J. K. Gibson et al. 2005. The ERA-40 re-analysis. Q. J. Roy. Meteorol. Soc., *131*, 2961–3012.

Vialard, J., F. Vitart, M. A. Balmaseda, T. N. Stockdale, and D. L. T. Anderson. 2003. An ensemble generation method for seasonal forecasting with an ocean-atmosphere coupled model. European Centre for Medium Range Weather Forecasts Technical Memorandum No. 417, 20 pp.

Wilson, S., F. Parisot, P. Escudier, J-L Fellous, J. Benveniste, H. Bonekamp, M. Drinkwater et al. 2010. Ocean surface topography constellation: the next 15 years in satellite altimetry *in* Proceedings OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21–25 September 2009, J. Hall, D. E. Harrison, and D. Stammer, eds., ESA Publication WPP-306. doi: 10.5270/OceanObs09.cwp.92

Xue Y., O. Alves, M. A. Balmaseda, N. Ferry, S. Good, I. Ishikawa, T. Lee et al. 2010. Ocean state estimation for global ocean monitoring: ENSO and beyond ENSO *in* Proceedings of OceanObs'09: Conference on Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21–25 September 2009, J. Hall, D. E. Harrison, D. Stammer, eds., ESA publication WPP-306. doi: 10.5270/OceanObs09.cwp.95

Zebiak, S. E. and M. A. Cane. 1987. A model El Nino-Southern Oscillation. Mon. Weather Rev. *115*, 2262–2278.

Zhu, J., B. Huang, R.-H. Zhang, Z.-Z. Hu, A. Kumar, M. A. Balmaseda, L. Marx et al. 2014. Salinity anomaly as a trigger for ENSO events. Sci. Rep., *4*, Article number 6821. doi: 10.1038/srep06821.

Received: *25 October 2015*; revised: *15 March 2017*.

Editor's note: Contributions to *The Sea: The Science of Ocean Prediction* are being published separately in special issues of *Journal of Marine Research* and will be made available in a forthcoming supplement as Volume 17 of the series.