


Different types of drifts in two seasonal forecast systems and their dependence on ENSO

L. Hermanson¹  · H.-L. Ren² · M. Vellinga¹ · N. D. Dunstone¹ · P. Hyder¹ · S. Ineson¹ · A. A. Scaife¹ · D. M. Smith¹ · V. Thompson¹ · B. Tian² · K. D. Williams¹

Received: 17 November 2016 / Accepted: 10 October 2017 / Published online: 11 November 2017
© The Author(s) 2017. This article is an open access publication

Abstract Seasonal forecasts using coupled ocean–atmosphere climate models are increasingly employed to provide regional climate predictions. For the quality of forecasts to improve, regional biases in climate models must be diagnosed and reduced. The evolution of biases as initialized forecasts drift away from the observations is poorly understood, making it difficult to diagnose the causes of climate model biases. This study uses two seasonal forecast systems to examine drifts in sea surface temperature (SST) and precipitation, and compares them to the long-term bias in the free-running version of each model. Drifts are considered from daily to multi-annual time scales. We define three types of drift according to their relation with the long-term bias in the free-running model: asymptoting, overshooting and inverse drift. We find that precipitation almost always has an asymptoting drift. SST drifts on the other hand, vary between forecasting systems, where one often overshoots and the other often has an inverse drift. We find that some drifts evolve too slowly to have an impact on seasonal forecasts, even though they are important for climate projections. The bias found over the first few days can be very different from that in the free-running model, so although daily weather predictions can sometimes provide useful information on the causes of climate biases, this is not always the case. We also find that the magnitude of equatorial SST drifts, both in the Pacific and other ocean basins, depends on the El Niño Southern Oscillation (ENSO) phase. Averaging over all hindcast years can therefore hide the details of ENSO state dependent drifts and obscure the underlying

physical causes. Our results highlight the need to consider biases across a range of timescales in order to understand their causes and develop improved climate models.

Keywords General circulation model · Bias · Drift · Seasonal prediction · ENSO

1 Introduction

It is essential to reduce model biases and improve the representation of underlying physical processes to achieve credible regional information in climate projections (Xie et al. 2015). This is also true for seasonal and decadal forecasts, where the models start from an observed state, but quickly drift (Smith et al. 2012). Although biases can be removed from seasonal and decadal forecasts using retrospective forecasts (also known as hindcasts, e.g. Arribas et al. 2011; Kharin et al. 2012; Fučkar et al. 2014), understanding how and why they develop is crucial for improving models and hence improving future climate predictions and projections on all timescales. Despite its importance, there are relatively few studies of the development of model biases, and the World Climate Research Group (WCRP) Working Group on Sub-seasonal to Interdecadal Prediction (WGSIP) has recently initiated a project to focus on this area.¹

Strong drifts away from an observed state have been reported in weather forecasts to result from an imbalance in fluxes between the atmosphere and ocean conditions due to insufficient communication between the two model components whilst creating the initial conditions (Mullholand et al. 2015). This can be mitigated with coupled data

✉ L. Hermanson
leon.hermanson@metoffice.gov.uk

¹ Met Office, Exeter, UK

² China Meteorological Agency, Beijing, China

¹ <https://www.wcrp-climate.org/wgsip-projects/lrftip>.

assimilation where all model components are coherently adjusted towards observations (Zhang 2011). However, at least on seasonal time scales, increasing the balance between components at the cost of moving the initial state further from reality can also increase forecast errors (Balmaseda and Anderson 2009).

The DEMETER project (Palmer et al. 2004) was a European multi-model seasonal and inter-annual prediction exercise. Several papers from this project were published that investigated model biases in initialized models (Lazar et al. 2005; B. Huang and Jha 2007; Jin et al. 2008). Lazar et al. (2005) decompose the bias into two terms: the bias in the first month and the remaining time-evolving bias. For errors above 1 °C they find that the first term is often due to the ocean and atmosphere model component chosen. The second term is mainly controlled by the atmosphere and can be rapidly varying. More recently, Liu et al. (2012) take an alternative approach and start from a resting ocean. They also find an initial bias (in the atmospheric component) which is followed by a later bias that is caused by the ocean response to winds set up by the initial bias.

Meehl et al. (2014) look at biases in decadal climate predictions and how they develop over time. They find a range of sources, including poorly constrained initial conditions, drift from the initial state to the model preferred state, unrealistic simulation of modes of internal variability and unknown future radiative forcings (such as volcanoes and aerosols). It is clear that such a range of sources of biases will produce drifts on many different time scales.

Vannière et al. (2014) present a systematic approach to examining biases using seasonal hindcasts applicable to any tropical SST bias. The approach uses the time scale over which biases develop in initialized coupled experiments to uncover the origin of the bias. Regional restoring experiments and ocean-only experiments are necessary to confirm the ultimate cause of the bias.

Sanchez-Gomez et al. (2015) look at the relationship between drift and large modes of variability. They find that in the tropical Pacific, where the model has a shallow mixed layer bias, it reaches its biased mean state after initialization by increasing meridional heat transport in the Pacific. In the North Atlantic they find an atmospheric model bias of weak winds over the subpolar gyre (SPG). Although the free-running model has a barotropic streamfunction in the SPG close to observations, a multi-decadal integration from observed initial conditions initially drifts away due to the atmospheric bias and then gradually reduces again. This leads the authors to state that drifts can be mostly interpreted as the integration by the ocean of intrinsic atmospheric biases.

These previous studies show that drifts are difficult to understand and can have many sources. In this study, we take a global view of the questions relevant to seasonal to decadal forecasters that arise from the literature. For example, are

there types of drifts that have properties in common? Is there a quick way to diagnose these drifts? Phillips et al. (2004) argue that the biases of the first few hours or days relate to the longer drifts. If so, the development of model biases may be investigated efficiently by using a climate model set up as a weather forecast. If the model state is always close to the observations it was initialized with, it can be argued that the errors from the dynamics are small so that this technique primarily evaluates parameterizations. Since climate models have coarser resolution than numerical weather prediction models, the parameterizations are potentially the most important source of biases (Rodwell and Palmer 2007).

As already mentioned, in seasonal and decadal prediction systems the forecast output is adjusted based on a set of hindcasts. The question that then arises is, are the drifts mentioned above a problem if they can be removed? However, there is an inherent assumption that the biases are stationary. Meehl et al. (2014) find that there is some residual uncertainty on the mean drift due to the time period over which the hindcasts are run not being long enough to sample multi-decadal variability. Here, we investigate whether different initial conditions in the tropical East Pacific can influence the drift locally and remotely on seasonal timescales, and find that the biases in that region are non-stationary.

We investigate the development of model biases relevant to seasonal to decadal forecasts using two operational seasonal forecast systems, those of the Beijing Climate Center and the Met Office. The models, data and methods are described in Sect. 2. In Sect. 3 we first compare drifts in seasonal to decadal forecasts with the long-term biases of the free-running versions of the same models for sea surface temperature (SST) and precipitation over a number of regions where known model problems exist. This is followed by Sect. 3.1 that introduces different types of drifts and then Sect. 3.1.1 explores how they could be categorized.

Fast and slow drifts are considered in Sects. 3.1.2 and 3.1.3, respectively. Our aim at this stage is not to provide a full explanation for how each drift occurs, but to make a broad characterization of drifts and draw general conclusions about their time scales and behaviour. We investigate the dependence of drifts on the state of El Niño in Sect. 3.2. Finally, Sect. 4 summarizes and discusses the results.

2 Models, data and methods

This study uses the hindcasts from two operational seasonal forecast systems to study the evolution of biases as a function of forecast lead time. The Beijing Climate Center-Climate Prediction System (BCC-CPS) is the 2nd generation seasonal prediction system of the Beijing Climate Center (BCC) at the China Meteorological Administration (CMA). It is quasi-operational and has been in use since 2014.

BCC-CPS is based on the BCC Climate System Model version 1.1m (BCC_CSM1.1m) described in Wu et al. (2013). Its atmospheric component is the BCC Atmospheric GCM with a T106 (1.125°) horizontal resolution and 26 vertical hybrid sigma/pressure levels (Wu et al. 2010) and its land component is the BCC Atmosphere and Vegetation Interaction Model version 1.0. The ocean component of BCC-CPS is the Geophysical Fluid Dynamics Laboratory Modular Ocean Model version 4.² The ocean horizontal resolution is 1° × 1° poleward of 30°N and 30°S, incrementally increasing to $\frac{1}{3}$ ° latitude within 30°N and 30°S. This ocean model has 40 vertical layers and includes the thermodynamic and dynamic elastic-viscous-plastic Sea Ice Simulator (Winton 2000). All of these components are coupled without flux adjustment.

The BCC-CPS hindcasts are initialized on 1 May and 1 November from 1991 to 2014 with a 13-month integration (though only the first four months are used here). The atmospheric initial values are initialized from the four-times daily NCEP Reanalysis I and the oceanic initial values are from the 3-D temperatures of the NCEP Global Oceanic Data Assimilation System (GODAS) (Saha et al. 2006), using a nudging scheme. Each hindcast includes 15 ensemble members initialized with a combination of different atmospheric and oceanic initial conditions from preceding days.

The Met Office Unified Model (UM) Global Coupled configuration 2 (GC2) version of HadGEM3 (Williams et al. 2015) is used in both seasonal and decadal forecasts at the Met Office. This is a global coupled model, consisting of dynamical ocean (NEMO), atmosphere (UM), sea ice (Los Alamos Sea Ice model, CICE) and land surface (JULES) components. The horizontal resolution in the atmosphere (longitude by latitude) is 0.83° × 0.55°, also referred to as N216. Vertical resolution is 85 levels in the atmosphere, allowing stratospheric dynamics to be resolved. The horizontal resolution in the ocean is nominally $\frac{1}{4}$ ° and there are 75 levels in the vertical.

The operational seasonal forecast system at the Met Office is Global Seasonal forecast system version 5 (GloSea5) described in MacLachlan et al. (2015). A set of hindcasts were used that are initialized on 1 May and 1 November. The hindcasts were made for 1996–2010 and each date has 8 ensemble members. The atmosphere and land surface is initialized from ERA-Interim and the ocean is initialized with a Met Office ocean and sea-ice re-analysis that uses the NEMOVAR system (Waters et al. 2014). Note that the two seasonal forecast systems have different years in their hindcast set. We have used all years available in order to characterize the systems as well as possible. Using the same

years for both systems does not change the conclusions made from the results.

To complement GloSea5 and provide some information on multi-annual model drifts we use the Met Office Decadal Prediction System version 3 (DePreSys3) described in Dunstone et al. (2016), which is initialized only on 1 November for 26 start dates between 1960 and 2014. Like GloSea5 it uses HadGEM3-GC2 and has the same resolution in its ocean and atmosphere, but the initialization is different. The initialization method of GloSea5 would not work for the sparse observations (particularly in the sub-surface ocean) in the 1960s, when the hindcast period of DePreSys3 begins. Weakly coupled initial conditions for hindcasts are provided by an assimilation integration where both atmosphere and ocean are nudged towards observations in a coupled model setup. The assimilation uses ERA-40/ERA-Interim in the atmosphere and has a 6 h relaxation time scale. In the ocean, a 3-D analysis based on EN4 is used (Smith et al. 2007) and has a 10-day relaxation time scale.

We define “bias” as an average model error and “drift” to be the change of this with lead time during a forecast. Our definition of drift includes changes caused by imperfect initialization, which are sometimes referred to as “shock” (Mullholland et al. 2015). We consider the drift in the first four months of the 1 May and 1 November start dates and also the summer (June–August) and winter (December–February) biases. DePreSys3 only has a 1 November start date and is run for five years. All the available hindcasts have been averaged over the available start years to make a lead-time dependent monthly mean climatology that can be compared with an observed 30 years seasonal cycle. For GloSea5 we also have daily data for the first month.

We have used spun-up, free-running model versions of the BCC and HadGEM3-GC2 models as controls to determine the long-term biases. The model configurations have been run under fixed concentrations of greenhouse gases and aerosols, typical of the 1980s and year 2000, respectively. Years 0–40 are discarded to allow the model to spin up, and the following 60 years used to evaluate the model.

The variables we have looked at are SST (°C) and precipitation (mm/day). We have chosen eight SST regions and six precipitation regions to study. These are either key regions dynamically or have well-known model biases for many climate models. The chosen regions are marked as boxes in Figs. 1, 2, 3 and 4.

Four of the SST regions are in the tropics: the equatorial Atlantic (as studied by Richter and Xie (2008)), the equatorial Indian Ocean (Han et al. 2012), NINO3.4, and the southeast tropical Atlantic (Toniazzi and Woolnough 2014). The Kuroshio extension and North Atlantic Current (NAC) are included due to their importance to atmosphere-ocean interactions (Kwon et al. 2010). Finally, the North Atlantic subpolar gyre (SPG) is included as it is important to decadal

² The MOM4 technical guide is available at http://mom-ocean.org/web/docs/project/MOM4_manual.pdf.

Fig. 1 Biases for JJA (left) and DJF (right) SST in free-running control versions of the BCC (top) and HadGEM3-GC2 (bottom) models. The boxes show areas that are averaged to make the indices discussed in this paper

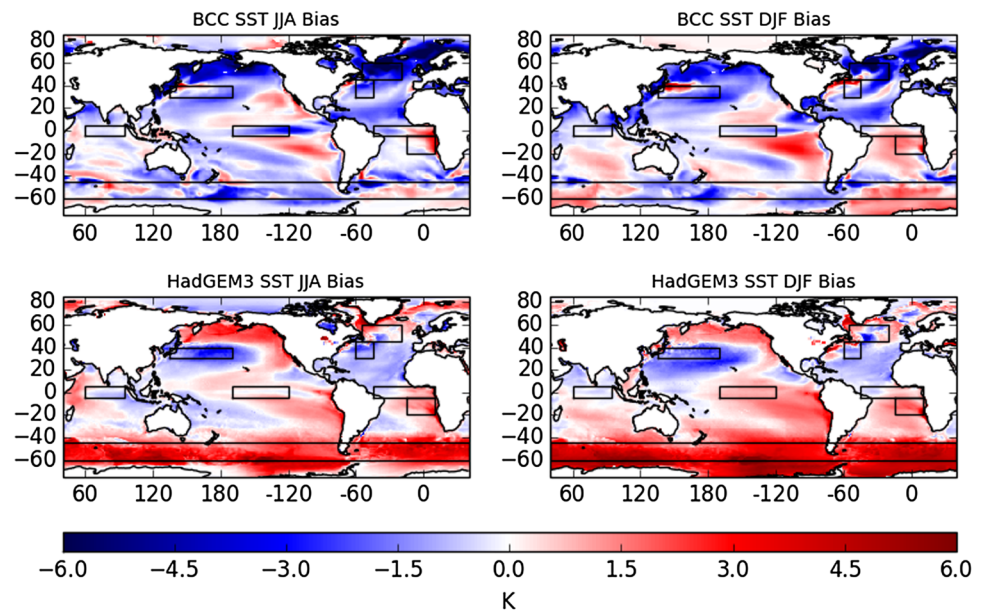
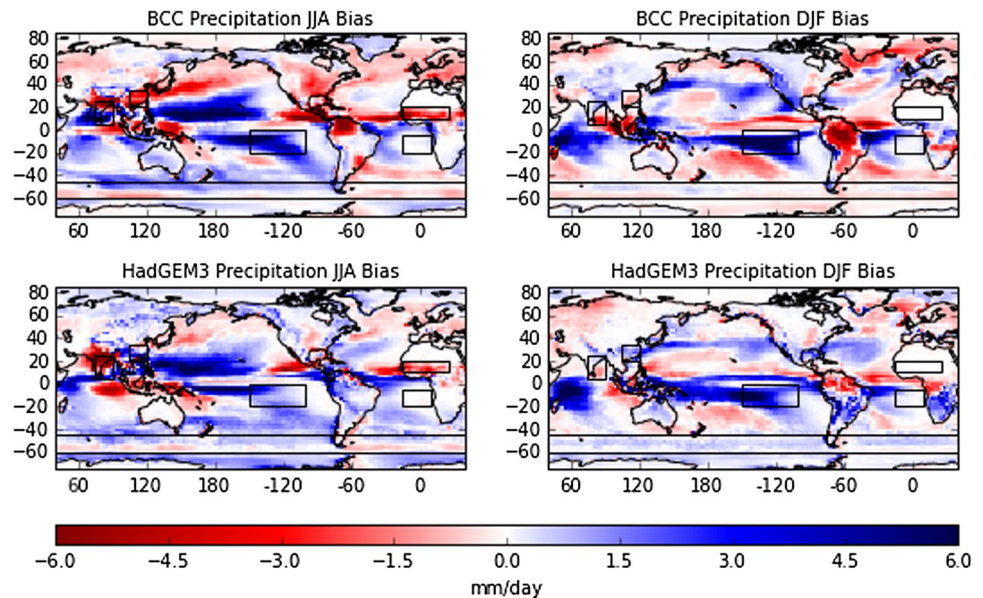


Fig. 2 Same as Fig. 1, but for precipitation. Note that for the indices over India and the Yangtze basin, only land points are used



prediction (Hermanson et al. 2014) and the Southern Ocean (SO) as many models have a bias there (Bodas-Salcedo et al. 2014).

The precipitation regions include three over land: the Sahel, India, and Yangtze river basin. These regions are of economic and agricultural importance. The Pacific double inter-tropical convergence zone (ITCZ) is a common bias in models (Bellucci et al. 2010), so included here. The south-east tropical Atlantic and Southern Ocean are included to complement the SST indices in the same region.

To assess model biases and drifts we use 30 years (1981–2010) of Reynolds et al. (2007) NOAA OI V2 high resolution SST. We used this product as it is on a $\frac{1}{4}^\circ$ grid

that can resolve sharp SST gradients. We also did not want to favour any forecast system by using the SST data set it is initialized with. To evaluate precipitation we use the same 30 years of GPCP V2.2 Combined Precipitation data set (Adler et al. 2003).

3 Results

The results are presented in two parts. In the first part, there is an overview and discussion about the general properties of the drifts, followed by an investigation of fast (time scale of

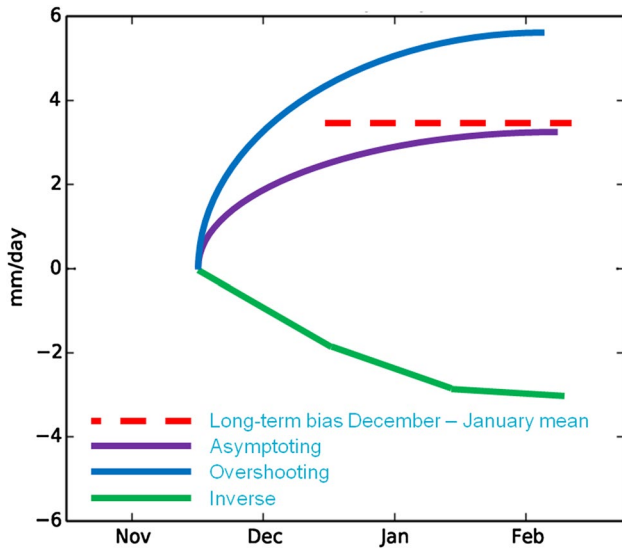


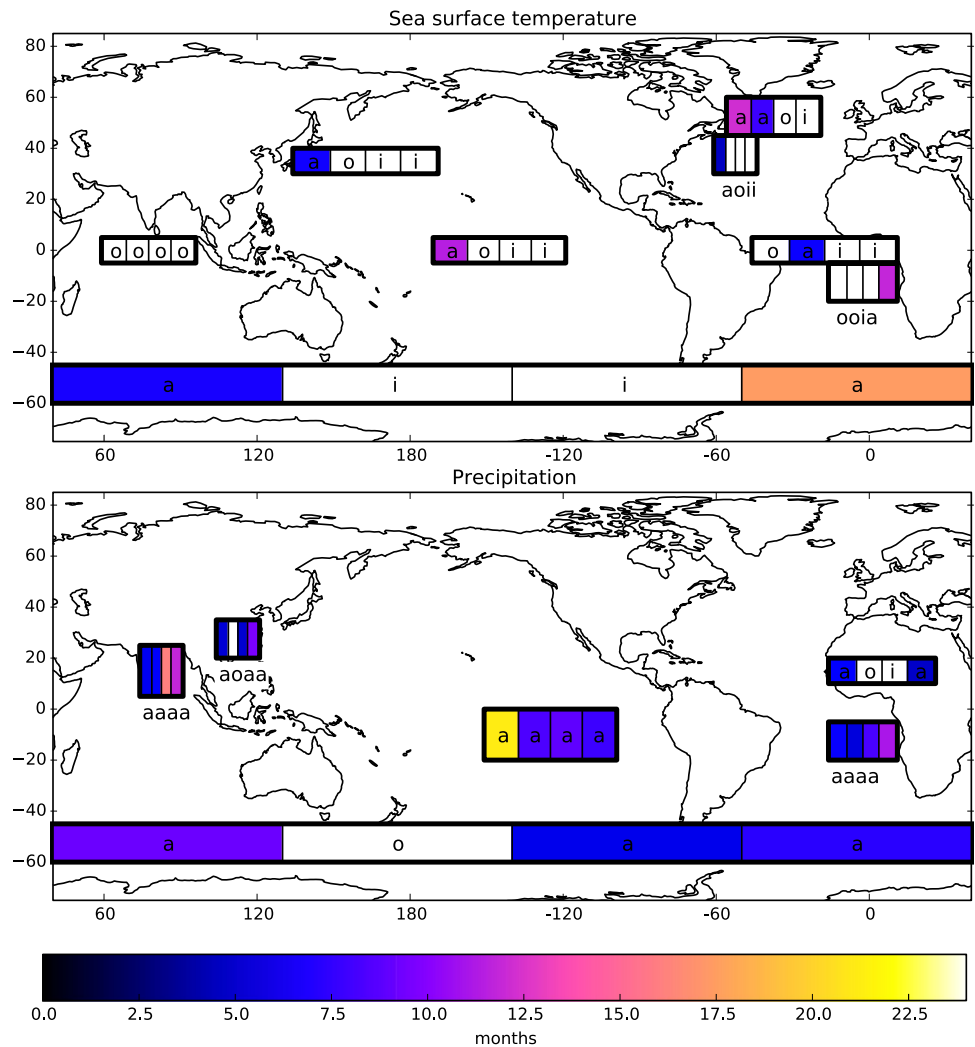
Fig. 3 Schematic showing the types of drifts included in the classification proposed in this paper

days) and slow (time scale of many months or years) drifts. The second part focuses on how drift depends on the initial state in the NINO3.4 region.

We start by looking at the summer and winter biases for the free-running control simulations. The biases for SST are shown in Fig. 1. There are some similarities between the two models. For example, both have biases that are negative in the sub-tropical and tropical oceans (apart from the southern hemisphere where they are positive). Note that the biases are of different sign in the equatorial East Pacific. HadGEM3-GC2 is generally too warm and BCC is generally too cold, although both models are too cold in the northern hemisphere and too warm in the southern hemisphere. For example, in December–January–February (DJF) the global mean SST bias for BCC is -0.4 K and the northern and southern hemispheres have a bias of -1.1 and 0.1 K, respectively. For HadGEM3-GC2 in the same season, the global mean SST bias is 0.9 K and the northern and southern hemispheres have a bias of -0.2 and 1.7 K, respectively.

The HadGEM3-GC2 integration shown here uses forcings from the year 2000, so it should be expected to have

Fig. 4 Drift type (letters) and time scales (colours) for SST (top) and precipitation (bottom). The first (second) letter or colour in a box refers to the May (November) start date for the BCC model hindcasts. The third (fourth) show May (November) start dates for GloSea5. Drift marked as ‘a’ is asymptotic, ‘o’ is overshooting and ‘i’ is inverse drift (away from the long-term bias). Time scales for overshooting and inverse drifts have not been calculated



an overall positive bias compared to the observational period 1981–2010, but only a few tenths of a Kelvin. Early on in this integration the northern hemisphere has a clear cold bias, but this becomes less evident as the global temperatures increase (not shown).

The prominent dipole biases appearing around 45°N in the Pacific and Atlantic basins imply a change in meridional SST gradients. In HadGEM3-GC2, a warm bias in the north and cold bias in the south leads to weaker meridional SST gradients than observations. For BCC, the polarity is reversed leading to strengthened SST gradients in the NW Atlantic and Pacific. The main SST bias for HadGEM3-GC2 is the Southern Ocean (SO) (Williams et al. 2015). The SO has both positive and negative biases for BCC, which when averaged over the box shown in the figure, will cancel to some extent. Similarly, for HadGEM3-GC2 for the box in the North Atlantic subpolar gyre (SPG), a generally warm SPG is partly cancelled in the box average index by a cold region in the south, while BCC is uniformly cold.

The precipitation biases for the control simulations are shown in Fig. 2. The two models have again some biases in common. Prominent are biases in the Sahel in June–July–August (JJA), India in JJA, the South Pacific (double ITCZ) and the SO. Both models have a southward shifted ITCZ in the Atlantic, probably related to the hemispheric biases in temperature mentioned above (Kang et al. 2010; Hermanson et al. 2014). There are uncertainties in both observational data sets, especially the precipitation (GPCP) away from the US and western Europe, where the satellite retrievals are unconstrained when there are no rain gauges. Therefore the exact magnitude of the biases in remote regions such as the SO is unknown.

3.1 Types of drifts

In order to aid the study of drifts we provide a framework that can be used to classify them into different types. The framework is based on the analogue of a damped harmonic oscillator whose equilibrium is the equilibrium state of the climate model. Through initialization the model is displaced from its equilibrium and may also be given an initial trajectory. When the forecast is started and the model is allowed to run freely, the behaviour of the drift is given by the strength of the damping, the restoring timescale of the model in this region and the initial trajectory. Following initialization, the free running model can either evolve towards or away from the long term bias, if it moves towards the bias it can either overshoot and oscillate back or it can gradually asymptote to the long term bias. There are therefore only three possible outcomes:

- Asymptotic drift: the mean forecast bias is of the same sign but smaller than the long-term bias. This is the case of a strongly damped oscillator.
- Overshooting drift: the mean forecast bias is larger than the long-term bias. This corresponds to a weakly damped oscillator that overshoots before equilibrating back to the long term bias.
- Inverse drift: the mean forecast bias is of opposite sign to the long-term bias due to the initial trajectory of the forecast.

Figure 3 shows a schematic of these three types of drift. The red dashed line represents the long-term bias in the model. At the start of the forecasts, the bias is close to zero. The purple forecast asymptotes monotonically towards the long-term bias as might naively be expected. The blue forecast shows a similar situation, but in this case the drift evolves rapidly and overshoots the long-term bias before eventually returning. Finally, the green forecast drifts away from the long-term bias. An example of how an inverse drift might occur is a strong ocean upwelling due to thermocline adjustments, which could change the SST until this has equilibrated after which other processes (perhaps related to biases in the atmospheric component of the model) become dominant and cause an opposite long-term bias.

3.1.1 General drift behaviour

We now categorize the type of drift for each of the boxes shown in Figs. 1 and 2 to be one of asymptoting, overshooting or inverse drift. This is done by comparing months 2–4 of the May and November hindcasts (that is JJA and DJF, respectively) with the same season of their respective free-running control integrations. In the case of an asymptoting drift, a time scale for the magnitude of the drift to become the same as the control integration bias is calculated by fitting an exponential to the hindcast drift and control bias. The results are summarized in Fig. 4.

The BCC-CPS hindcasts most often asymptote or overshoot. GloSea5 is more likely to inverse drift or asymptote. There is a notable difference between the two variables. Precipitation most often has an asymptoting drift, while SST has similar numbers of each drift. Most asymptotic drifts reach the long-term mean in 8 months or less (especially for precipitation), but there are exceptions such as the SO SST in November for GloSea5 and precipitation in the Pacific ITCZ in May for BCC-CPS, which take 18 months or more.

For SST, the only region where both forecast systems have the same type of drift for both start dates is the equatorial Indian Ocean, as shown in Fig. 4, where both systems always overshoot. The equatorial Indian Ocean region is difficult to model and has a large impact on the ability to correctly model the South Asian monsoon due to the strong

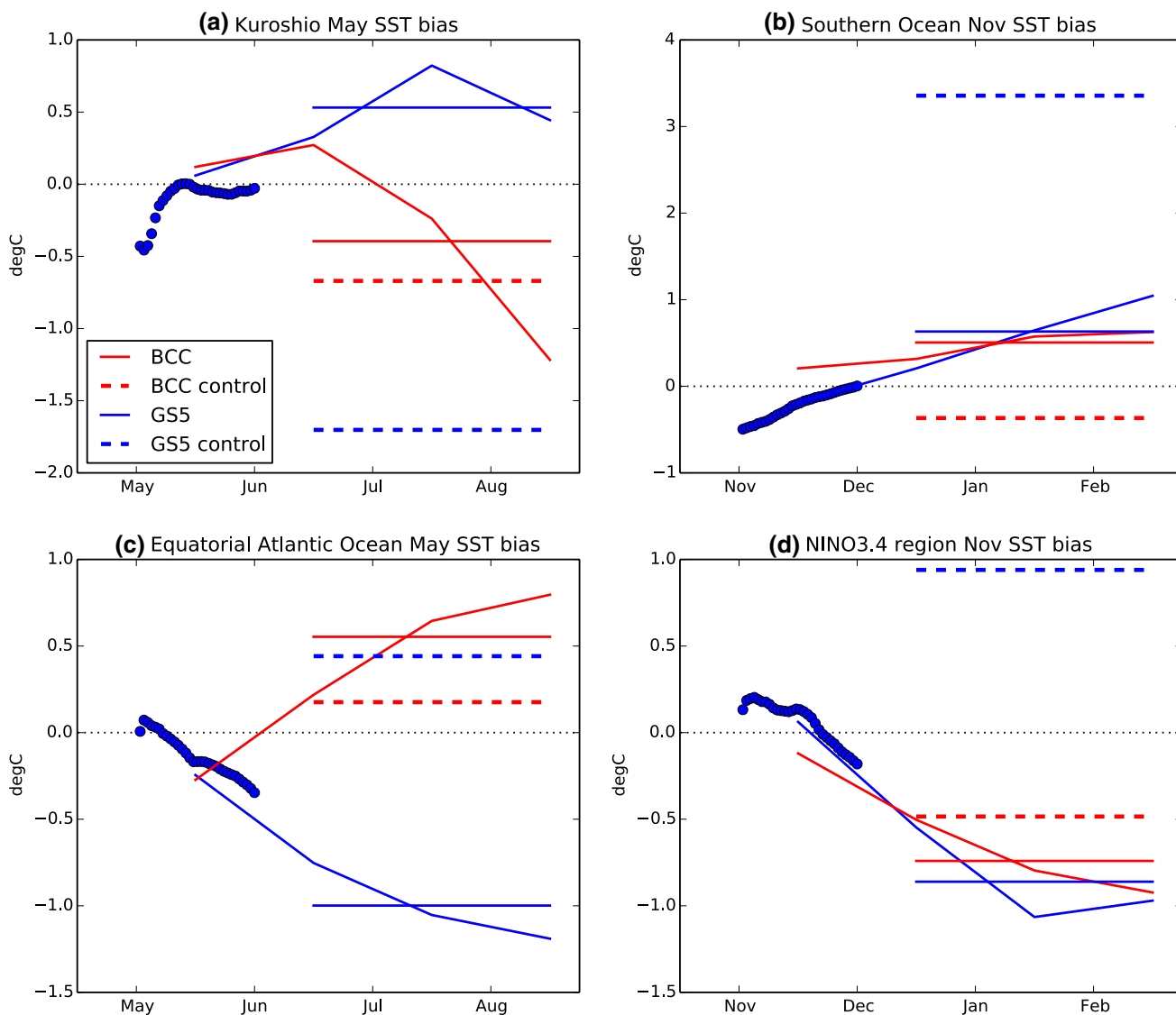


Fig. 5 Examples of SST drifts from initialization as a function of lead time for BCC (red) and GloSea5 (blue) compared to the long-term biases of their free-running controls (dashed of same colour). Four monthly means are shown as well as a horizontal line that indi-

cates the mean for the last three months, that is directly comparable to the control value. Blue dots shows the daily mean bias for GloSea5 relative to linearly interpolated monthly observations

coupling between SST, large-scale atmospheric circulation and local convection in the region (Schott et al. 2009; Bolasina and Ming 2012). This also makes it difficult to disentangle the origin of a drift here. However, the uniform behaviour of the drift between these two models and start dates gives some hope that a future study might find a common mechanism (though the similarity here may be incidental). One potentially important feature is that this basin does not have a subtropical gyre in the northern hemisphere, unlike all the other oceans in the northern hemisphere.

Figure 5 shows the drifts in SST of two extra-tropical regions, the Kuroshio Extension (1 May start) and the SO (1 November start) (a, b) and two tropical regions, the

equatorial Atlantic Ocean (1 May start) and the NINO3.4 region (1 November start) (c, d). BCC-CPS (red) shows all three types of drift: asymptoting in a, an inverse drift in b and overshooting in both c, d. The drifts can be both cold and warm without any dependence on latitude. On the other hand, GloSea5 (blue) has mainly inverse drifts (except b) and drifts cold along the equator and warm in the extra-tropics. This is generally true for GloSea5 and the extra-tropical warm drift is largest for the summer hemisphere (not shown). This points to short wave radiation biases (modulated by clouds) being important for the initial drift in the extra-tropics. Using the same model as GloSea5, Bodas-Salcedo et al. (2014) show that this problem persists in the

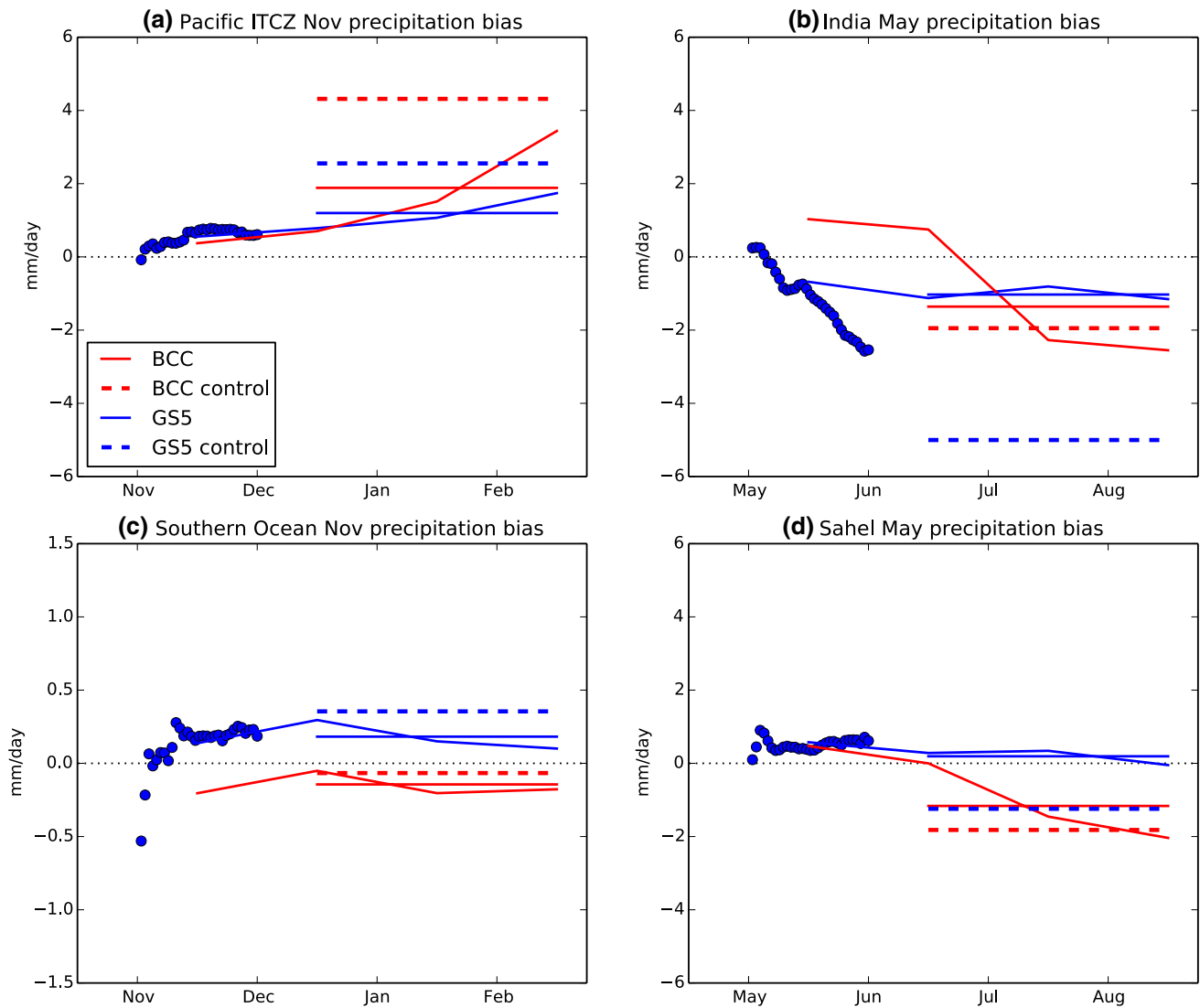


Fig. 6 Examples of precipitation drifts from initialization as a function of lead time for BCC (red) and GloSea5 (blue) compared to the long-term biases of their free-running controls (dashed of same colour). Four monthly means are shown as well as a horizontal line that

indicates the mean for the last three months, that is directly comparable to the control value. Blue dots shows the daily mean bias for GloSea5 relative to linearly interpolated monthly observations

SO, making that an asymptoting drift, but other processes must become dominant in the North Pacific as the long-term bias has the opposite sign.

The initial cooling in equatorial regions may be a response to the extra-tropical SST drifts. In the framework of Lazar et al. (2005), drifts on the time scale of months is likely to have an atmospheric origin. In Atlantic hosing experiments with this model, where one hemisphere is deliberately cooled, there is a shift of the ITCZ in all ocean basins towards the warmer hemisphere (Jackson et al. 2015). Maps of the early precipitation biases show that this occurs here as well (not shown). The equatorial cooling is therefore likely a result of increased cross-equatorial winds due to the off-equatorial ITCZ.

3.1.2 Fast drifts

Here we use daily data from GloSea5 to look at drifts on the time scale of days to weeks. These drifts, as long as the background state is close to the observed state, are mostly due to deficiencies in the model parameterizations (Phillips et al. 2004; Martin et al. 2010). The question we ask here is whether the fast drifts one might find in a weather forecast by the seasonal system are indicative of the seasonal and control biases. This was the approach taken in the transpose AMIP II experiments (Williams et al. 2013).

In Figs. 5 and 6, the bias for SST and precipitation, respectively, is shown by a blue dot for every day in the first month. It appears that SST for the Kuroshio extension and

the SO (Fig. 5a, b) have been poorly initialized in GloSea5 as the first day does not start with a zero bias. As already mentioned, the observed SST data set used (Reynolds) is not the same as that which GloSea5 is initialized with, but these biases are larger than $0.5\text{ }^{\circ}\text{C}$ in both cases for a large region. There is likely a problem with the initial conditions or a fast change in the atmosphere (eg. clouds) in the first to second day. Either way, the model state is initially far from the observed state, which makes a study of model parameterization deficiencies difficult.

For many precipitation indices, there is a big change from the first to the second day. All panels, apart from (b), show this in Fig. 6. From day one to day two the globally averaged precipitation bias increases by 0.5 mm/day and in the tropics it increases by 0.8 mm/day (averaged over both start dates). From day two to day three the change in the global mean bias is 0.1 mm/day . As the change over the first two days is largest in the tropics, it is likely that convective precipitation is causing the change (as this is the most common type of precipitation). It is possible the jump in equatorial SST at the same time, which can be seen in Fig. 5, panels c, d, is related to this fast change in the hydrological cycle.

For India in May, panel b in Fig. 6, just before the monsoon starts, in GloSea5 the magnitude of the precipitation bias is less than 1 mm/day , but after the first week this increases quickly throughout the rest of May. When the monsoon starts in June, the precipitation is biased dry. BCC-CPS shows the same evolution where May and June the precipitation is too strong and then the bias switches sign to being at least -2 mm/day . It is difficult to explain these drifts as the Indian rainfall has complicated teleconnections to both Indian Ocean and Pacific Ocean SST (Clark et al. 2000), but as the drift is about 3 mm/day in the first month there are clearly large changes in the local circulation.

In Fig. 6, panel c shows the SO in November. For GloSea5, in the first two days the precipitation bias is negative as there is less precipitation than in observations (anomalies may seem insignificant, but the region they represent is large). However, the HadGEM3-GC2 long-term bias is positive for precipitation (the model has too much). It appears that the precipitation increases quickly over the first few days after which the bias growth is slower. It is only after ten days that the daily bias becomes close to that for the DJF season. This behaviour is also in the May start date as well (not shown). The HadGEM3-GC2 model has too much precipitation (Fig. 2) over most of the globe, including the SO. This initial growth of the precipitation over the first weeks is consistent with the time scale of increasing moisture transport from the tropics, which implies that the mid-latitude atmosphere is initialized from ERA-Interim with less precipitable water than in the control integration.

In summary, it appears that the initial bias, for certain regions for the first day or week, can be of the opposite sign

to the long-term bias due to initialization error or a lack of precipitation. In these regions with inverse drift, it would not be appropriate to use short forecasts to evaluate long-term model biases. One approach would be to confirm that these biases are asymptoting or overshooting before proceeding with such an analysis.

3.1.3 Slow drifts

In some cases, drifts do not grow appreciably over the length of a seasonal forecast, compared to the long-term bias. A good example is Fig. 6a, where the precipitation biases for both systems are growing slowly. BCC-CPS also shows an asymptoting drift in Fig. 5a for Kuroshio SST in the May start date and the SPG (not shown). Figure 4 indicates that to reach the full bias would take BCC-CPS 12 months and 8 months in the SPG for the May and Nov start dates, respectively. Both time scales are longer than the time scale of a seasonal forecast.

In the case of GloSea5 for these regions a simple calculation cannot be made for the time scale as it mostly has an inverse drift. However, decadal forecasts from DePreSys3, which is based on the same model, can provide an estimate. Figure 7a shows that the SPG might take 18 months and even then the seasonal cycle of the bias is still not the same as the free-running model. It is possible that the slow bias growth is related to an interaction with the large-scale meridional overturning circulation (MOC) in the ocean and anomalous freshening in the northern North Atlantic as seen by Huang et al. (2014). In that case, the biases took several years to grow. Another explanation is that the ocean is integrating atmospheric biases (Sanchez-Gomez et al. 2015) and it seems reasonable that this time scale is of the order of that needed for the large-scale ocean transport and heat content to adjust to the new atmospheric forcing.

For Sahel precipitation in May, shown in Fig. 6d, BCC-CPS has a time scale of six months for the drift, but GloSea5 has a time scale longer than the length of the hindcast. As an indication of how long the drift may take, in DePreSys3, Fig. 7b, the Sahel precipitation takes more than 2 years to halve the difference with the long-term bias for the summer rainy season.

In the SO drifts are slow, as seen in Fig. 5b, partly due to a deep mixed layer and vertical exchanges in the ocean reducing the impact on SST of short wave biases caused by a poor representation of clouds (Bodas-Salcedo et al. 2014). GloSea5 exhibits much smaller SST and precipitation biases in its seasonal forecasts than the free-running model. As mentioned earlier, HadGEM3-GC2 climate integrations have a large bias in the SO, but the GloSea5 forecasts do not. Figure 4 shows that BCC-CPS also has slowly asymptoting drifts in the extra-tropics.

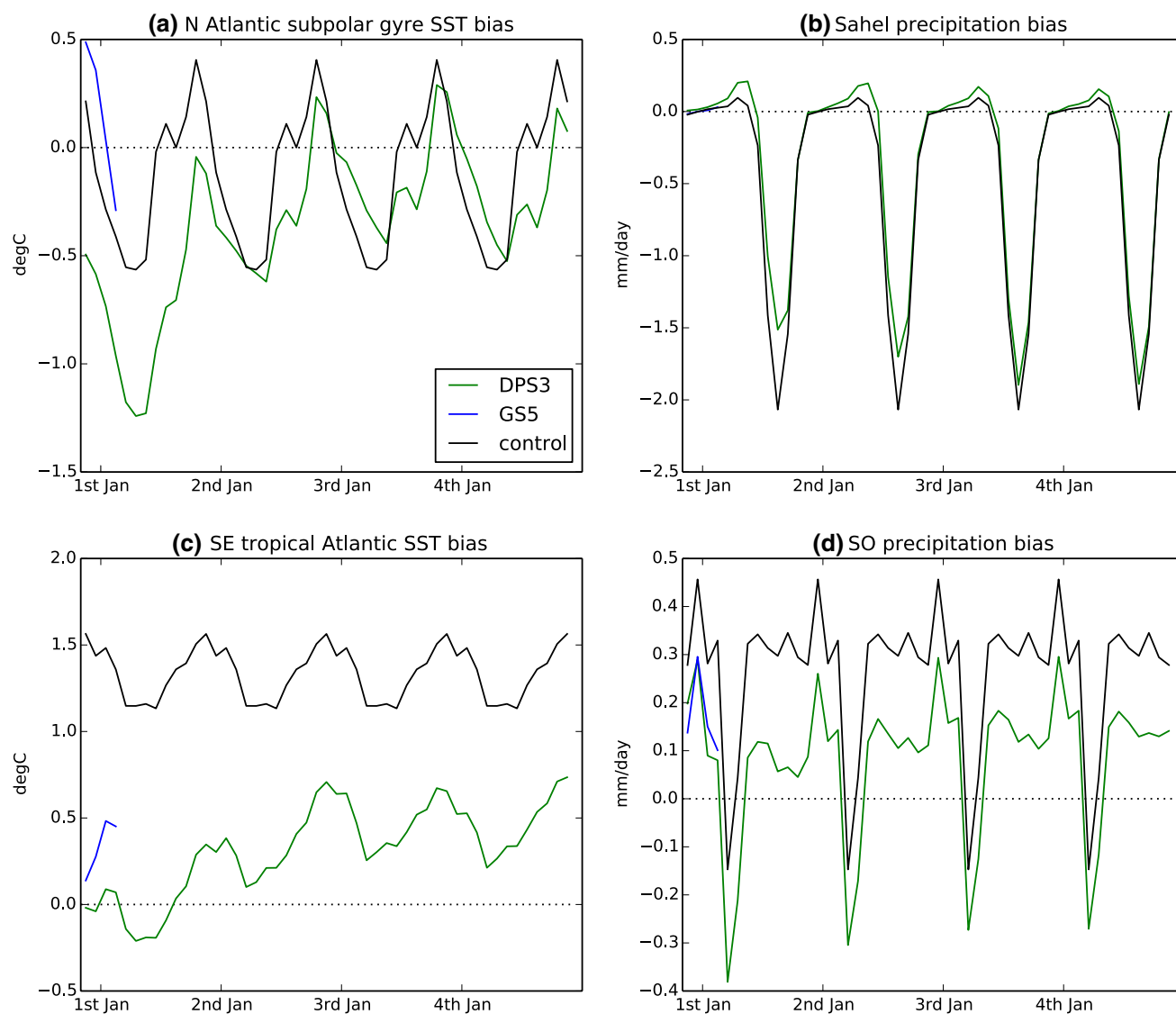


Fig. 7 Forecast system biases from initialization as a function of lead time for DePreSys3 (green) and GloSea5 (blue) compared to the long-term bias of their free-running counterpart (black)

The same can be said for some of the tropical biases in HadGEM3-GC2, such as the Southeast tropical Atlantic shown in Fig. 7c, where even at the end of four years of integration the bias has not reached half the magnitude of control integration. It can be seen from Fig. 4 that for the November start date with GloSea5 both the SST and precipitation have slow drifts. This is expected in this region as convection is linked to local SST changes and SST gradients. It is perhaps more surprising that in the SO this is also the case. The differences in precipitation may seem insignificant, but they represent a very large area (more than 40 million km²). The precipitation in DePreSys3 is shown in Figure 7d and convergence with the long-term bias does not happen within the time of the forecast. As already discussed in relation to Fig. 6c, this could be

related to changes in southward moisture transport as the tropical biases grow.

The time scales of the drifts mentioned in this section mean they do not fully impact seasonal forecasts or inter-annual forecasts, even in the case of precipitation. This is further discussed in Sect. 4.

3.2 ENSO dependent drifts

So far, we have examined drifts averaged over all initial states, but drifts are state dependent, as explored by Goddard and Dilley (2005), Ren (2008) and Fučkar et al. (2014). The ENSO region is of particular interest to seasonal prediction as it is predictable and has many strong teleconnections. This is also a problem, because if a model drifts in this region, the

teleconnections mean that this drift may also be transferred globally. The temperature difference in the NINO3.4 region between El Niño and La Niña can be more than 2°C. This indicates there are two quite different climate states in this region (or three counting the neutral state). Is it possible that a model may be able to simulate one of these states better than the other? Are drifts from different ENSO states also different? We attempt in this section to look closer at these questions by stratifying our drift according to the ENSO state. Note that there may be more than two different climate states in this region (Johnson 2013).

It was shown in Fig. 5d that the NINO3.4 region in November has a typical overshoot for BCC-CPS and inverse drift for GloSea5 in the November start date. Are these drifts independent of the initial conditions in the equatorial Pacific? We choose 10 El Niño years and 10 La Niña years from the NINO3.4 index of ERSSTv4 data (Huang et al. 2015). The criteria are a minimum three month anomaly of 0.5 °C magnitude for at least four consecutive three-month running means. It is assumed, as the lead time is short, that the hindcasts for these years initialized in November will contain the desired ENSO state. This can be seen from absolute values of the NINO3.4 index in the hindcasts averaged over the respective years (not shown). The El Niño years are (year for November): 1982, 1986, 1987, 1991, 1994, 1997, 2002, 2004, 2006, 2009. The La Niña years are: 1983, 1984, 1988, 1995, 1998, 1999, 2000, 2007, 2008, 2010.

The left and right columns of Fig. 8 show the drifts for El Niño and La Niña, respectively. In the NINO3.4 region itself (a, b) the forecast systems have different drift during different ENSO phases. For both seasonal forecast systems the difference in the drift is more than 0.5 °C. Although there are only a few events of each phase, the differences are significant at the 5% level for the monthly difference in bias between phases over December – February. The long-term bias on average (Figs. 5 or 1) show that HadGEM3-GC2 has a warm bias and BCC a cold bias in the NINO3.4 region. In Fig. 8a, b the long-term biases have also been calculated in the control integrations using the same criteria as in the observations. The forecast systems respond in opposite ways to being initialized with one or the other of the ENSO phases. For El Niño in a, GloSea5 has the smallest long-term bias and drifts the least. For La Niña in b, BCC-CPS has the smallest long-term bias and drifts the least. In the NINO3.4 region, it appears that the model with the smallest long-term bias for a particular ENSO phase, also has the smallest drift in a seasonal forecast for that phase.

The drifts for both models also change in the equatorial Indian Ocean with ENSO phase, shown in Fig. 8c, d. Again BCC-CPS has a larger drift for El Niño than La Niña and GloSea5 has a bigger drift for La Niña than El Niño.

A Hovmuller plot for GloSea5 averaged over 5°N–5°S in the Indo-Pacific-Atlantic in Fig. 9 shows in more detail how

the biases grow in GloSea5. The drift in the bias is strongest in the western Indian Ocean, East Pacific and the central Atlantic. This is seen most clearly for La Niña years. There also appears to be some propagation towards the maritime continent. The eastward propagation that starts from about 50°E in November has a speed of roughly 1 m/s, consistent with an equatorial Kelvin wave, which could have been caused by a change in the wind forcing from initialization to the free-running forecast. The westward propagation starting at about 160°W is faster, so is not a Rossby wave and could be mediated by the atmosphere. Another explanation for these drifts is a re-adjustment of the thermocline in the ocean (Vannière et al. 2013).

Precipitation drifts can also change with ENSO phase. India in November sees the last of its monsoon rain, so precipitation December–February is only 1–2mm/day. However, in GloSea5 the difference in bias between ENSO phases for November is about 0.8 mm/day (about 70% of observed average rainfall for that month), negative in El Niño and positive in La Niña. The model appears to give the same rainfall independent of the initial conditions, but in the observations and for BCC-CPS El Niño years have more precipitation in November than La Niña years.

The Yangtze basin has a difference of about 1 mm/day in observations for November between El Niño years (when there is more precipitation) and La Niña years (Xiao et al. 2015). GloSea5 can only reproduce a fraction of this difference between the two phases. This leads to a change in the bias between El Niño (–0.2 mm/day) and La Niña (0.6 mm/day). BCC-CPS has the opposite response to the observed, and to GloSea5, leading to a bigger difference between biases, –1.1 and 0.7 mm/day, for El Niño and La Niña years, respectively. In the Pacific ITCZ region there is a more precipitation in El Niño years than La Niña years in the observations (not shown). Fig. 8e, f show forecast systems have larger drifts in La Niña years, when the precipitation is less. It seems the models are unable to reduce their precipitation enough in La Niña years. This is possibly because the anomalous cross-equatorial moisture transport that causes this bias is also driven by shortwave biases in the subtropics as suggested by Hwang and Frierson (2013).

4 Summary and Discussion

We used two seasonal forecast systems, BCC-CPS and GloSea5, as well as a decadal prediction system, DePreSys3, to examine climate model drift in SST and precipitation on time scales of days to years. We characterized the drift into three types in Sect. 3.1 depending on how the bias approached the long-term bias of the free-running counterpart of the model. These were applied to the two seasonal prediction systems. Asymptoting drifts are most common

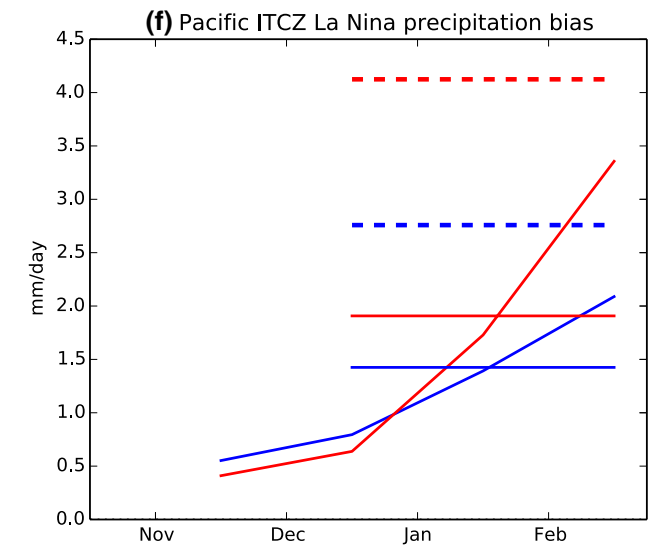
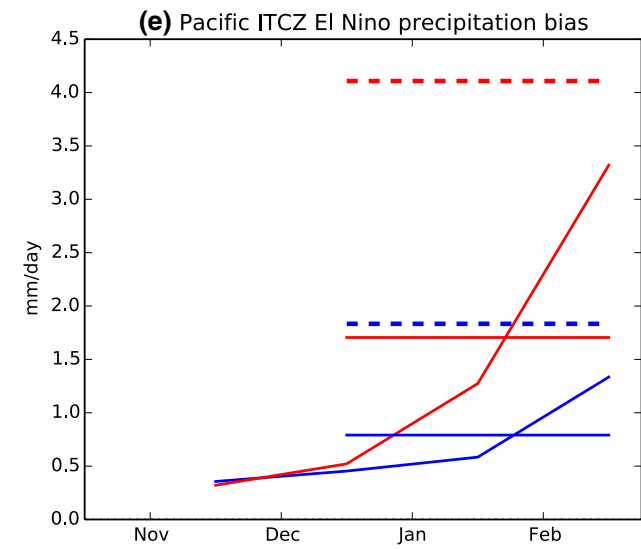
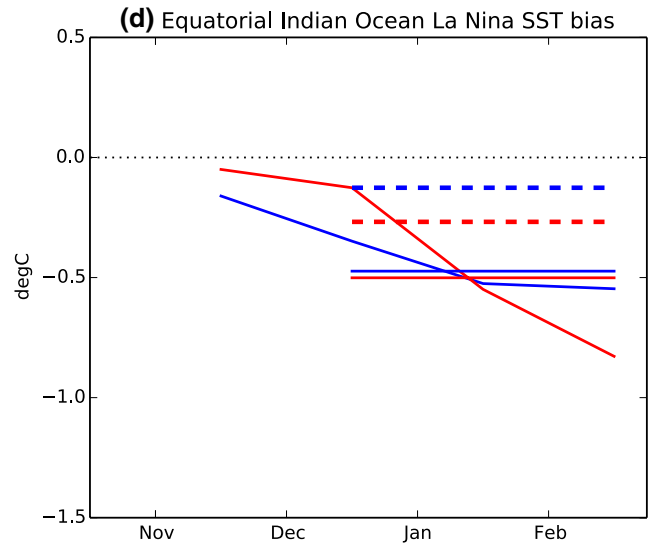
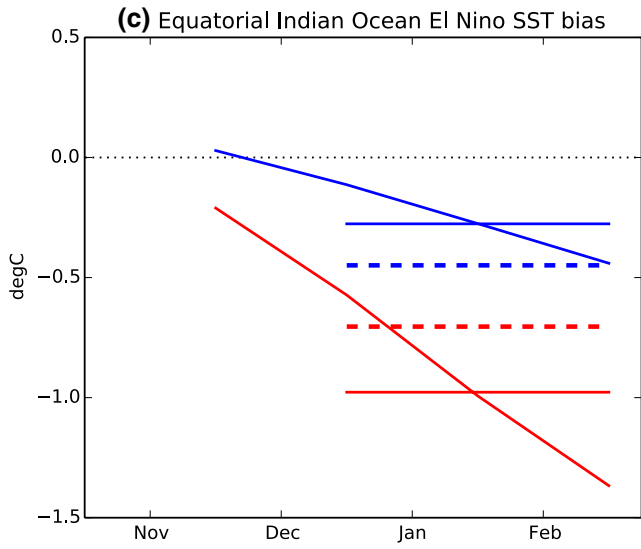
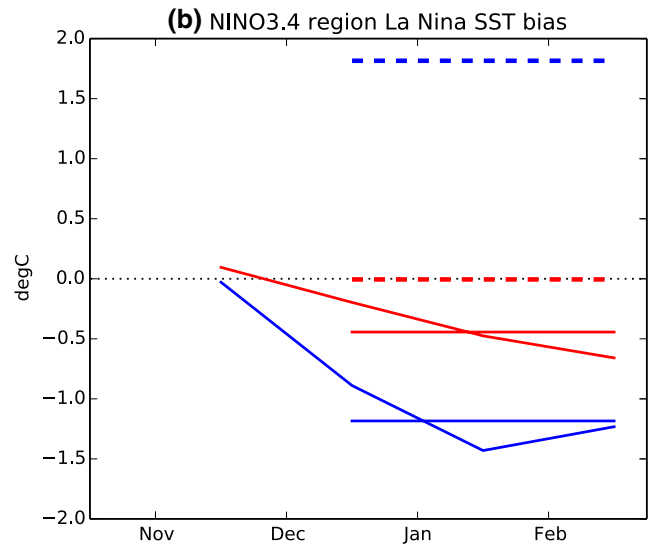
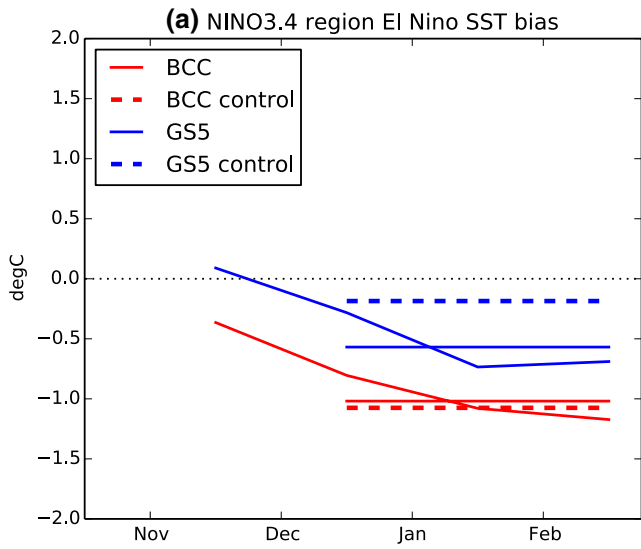


Fig. 8 Forecast system biases from initialization as a function of lead time for BCC (red) and GloSea5 (blue) compared to the long-term biases of their free-running controls (dashed of same colour). Four monthly means are shown as well as a horizontal line that indicates the mean for the last three months, that is directly comparable to the control value. Left hand column for years with El Niño initial conditions and right hand column for La Niña years

for precipitation. However, unexpectedly, asymptoting drifts are not the most common drift for SST. BCC-CPS tends to overshoot, while GloSea5 tends to exhibit inverse drifts, for both tropics and extra-tropics.

Our main conclusions are:

- There are often fast drifts in the first month, sometimes related to initialization problems (Fig. 5a, b). In most cases, the biases of the first few days are not representative of the seasonal or long-term biases due to changes after day 1 (Figs. 5c, d, 6).
- Some drifts are so slow they can be considered much less of a problem for seasonal and even multi-annual forecasts than for climate model simulations (Sect. 3.1.3).
- Both forecast systems show that drifts in the NINO3.4 region depend on the state of ENSO in initial conditions. This is also apparent in other ocean basins and in several precipitation indices (Sect. 3.2).

The fast drifts described here suggest that using a climate model in numerical weather prediction mode, as suggested by Phillips et al. (2004), could give misleading results in diagnosing parameterization errors. This is particularly true of precipitation, which behaves differently in the first few days to the long-term bias. Precipitation can take up to a week to develop a significant bias and it may also initially be of the opposite sign to the seasonal forecast bias. In the case of SST, the initialization errors mean that the model climatology is not close to observations, even at the beginning of the forecast. Care should be taken when applying the methods of Phillips et al. (2004) and transpose AMIP II (Williams et al. 2013) to seasonal prediction.

One prominent example of a slow drift that may not be important for seasonal prediction for GloSea5 is the Southern Ocean (SO) warm SST bias that takes many years to grow. The Pacific double ITCZ problem is probably partly linked to this bias (Hwang and Frierson 2013) and so may also be less important for seasonal forecasts than climate integrations. This is a case where precipitation adjusts slowly, contrary to what one might think considering that the atmosphere can change quickly compared to the ocean. Precipitation over land can also adjust slowly, for example the Sahel precipitation shown in Fig. 7b take several years to reach the full long-term bias in JJA.

The BCC model has a mean state that is biased cold and drifts more when initialized with an El Niño than a La Niña

state. In contrast, HadGEM3-GC2 has a mean state that is biased warm and drifts the more when initialized with a La Niña state. So although the responses are different, the models are behaving consistently and drifting more when the long-term bias is larger in this region. Note that this refers to the long-term bias during ENSO events and not the average bias over all years.

A strong relationship between both circulation biases and total precipitation bias, and the NINO3 index was found by Ren (2008) for BCC-CGCM1.0 seasonal predictions. Our results further support the conclusions of Ren (2008), highlighting the need for more research to understand alternatives to standard forecast bias correction methods based on the linear average hindcast bias. Longer hindcast periods that include more ENSO events could be beneficial, though the additional computational costs need to be balanced against the need for large ensembles. Eade et al. (2014) show that large ensembles are necessary to forecast the North Atlantic region.

BCC-CPS is too cold in most regions and the processes that cool the ocean (or decrease the mixed-layer depth) at the start of the forecast may cause the drift to overshoot the long-term bias. For GloSea5, extra-tropical inverse drifts are mostly warm, despite the long-term bias being cold and the reverse is true for tropical drifts. It appears that cloud and radiation biases cause the warming drift, which is largest in the summer hemisphere. The tropical SST drifts appear to follow from the ITCZ response to the extra-tropical drifts. However, more research is needed to confirm this.

Another likely explanation for these SST drifts is that they are caused by changes between the assimilation and free-running forecast. Mullholland et al. (2015), for example, found that imbalances between the atmosphere and ocean initial conditions could create drifts. Heat fluxes to the ocean component will almost certainly be erroneous in the initial state as during the assimilation phase ocean data increments can compensate for errors in the atmospheric forcing. Therefore, in the forecast the ocean state would drift even given no forcing error.

In summary, our results show that initialized forecasts do not always simply drift monotonically towards the free-running model climatology. Furthermore, the drift evolution may depend on the initial state of the forecast as we found in the NINO3.4 region. This complicates diagnosing the underlying causes of model biases, and highlights the need to consider a range of timescales in order to understand the the causes of model biases in forecasts and improve climate models in the future.

Acknowledgements The authors were supported by the UK-China Research and Innovation Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP) China as part of the Newton Fund. The research leading to these results has also received funding from the European Union Seventh Framework Programme

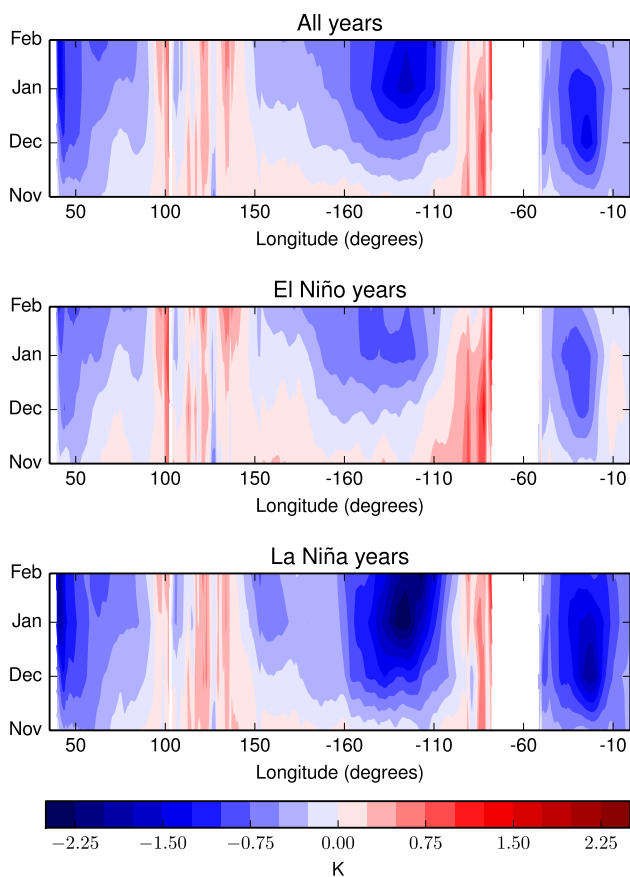


Fig. 9 GloSea5 Indo-Pacific-Atlantic SST biases averaged over 5° N–5°S for all years (top), years with El Niño conditions (middle) and La Niña years (bottom)

(FP7/2007–2013) under SPECS project (Grant agreement 308378). The authors acknowledge the scientific guidance of the World Climate Research Programme in helping to guide this work through the framework of the Working Group on Subseasonal to Interdecadal Prediction (WGSIP). NOAA High Resolution SST data and GPCP Precipitation data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>. The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Adler R, Huffman G, Chang A, Ferraro R, Xie P-P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P, Nelkin E (2003) The version-2 global precipitation climatology project (gpcp) monthly precipitation analysis (1979–present). *J Hydrometeorol* 4:1147–1167.

doi:[10.1175/1525-7541\(2003\)004%3C1147:TVGPCP%3E2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004%3C1147:TVGPCP%3E2.0.CO;2)

- Arribas A, Glover M, Maidens A, Peterson K, Gordon M, MacLachlan C, Graham R, Fereday D, Camp J, Scaife AA, Xavier P, McLean P, Colman A, Cusack S (2011) The GloSea4 ensemble prediction system for seasonal forecasting. *Mon Weather Rev* 139:1891–1910. doi:[10.1175/2010MWR3615.1](https://doi.org/10.1175/2010MWR3615.1)
- Balmaseda M, Anderson D (2009) Impact of initialization strategies and observations on seasonal forecast skill. *Geophys Res Lett*. doi:[10.1029/2008GL035561](https://doi.org/10.1029/2008GL035561)
- Bellucci A, Gualdi S, Navarra A (2010) The double-itcz syndrome in coupled general circulation models: the role of large-scale vertical circulation regimes. *J Clim* 23:1127–1145. doi:[10.1175/2009JCLI3002.1](https://doi.org/10.1175/2009JCLI3002.1)
- Bodas-Salcedo A, Williams KD, Ringer MA, Beau I, Cole JNS, Dufresne J-L, Koshiro T, Stevens B, Wang Z, Yokohata T (2014) Origins of the solar radiation biases over the southern ocean in CFMIP2 models. *J Clim* 27:41–56. doi:[10.1175/JCLI-D-13-00169.1](https://doi.org/10.1175/JCLI-D-13-00169.1)
- Bollasina M, Ming Y (2012) The general circulation model precipitation bias over the southwestern equatorial Indian Ocean and its implications for simulating the South Asian monsoon. *Clim Dyn* 40:823–838. doi:[10.1007/s00382-012-1347-7](https://doi.org/10.1007/s00382-012-1347-7)
- Clark CO, Cole J, Webster P (2000) Indian Ocean SST and indian summer rainfall: predictive relationships and their decadal variability. *J Clim* 13:2503–2519
- Dunstone N, Smith D, Scaife A, Hermanson L, Eade R, Robinson N, Andrews M, Knight J (2016) Skilful predictions of the winter North Atlantic Oscillation one year ahead. (submitted)
- Eade R, Smith D, Scaife A, Wallace E, Dunstone N, Hermanson L, Robinson N (2014) Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys Res Lett* 41:5620–5628. doi:[10.1002/2014GL061146](https://doi.org/10.1002/2014GL061146)
- Fučkar N, Volpi D, Guemas V, Doblas-Reyes F (2014) A posteriori adjustment of near-term climate predictions: accounting for the drift dependence on the initial conditions. *Geophys Res Lett* 41(14):5200–5207. doi:[10.1002/2014GL060815](https://doi.org/10.1002/2014GL060815)
- Goddard L, Dilley M (2005) El nin/ o: catastrophe or opportunity. *J Clim* 18:651–665. doi:[10.1175/JCLI-3277.1](https://doi.org/10.1175/JCLI-3277.1)
- Han Z-Y, Zhou T-J, Zou L-W (2012) Indian Ocean SST biases in a flexible regional ocean atmosphere land system (FROALS) model. *Atmos Ocean Sci Lett* 5:273–279
- Hermanson L, Eade R, Robinson N, Dunstone N, Andrews M, Knight J, Scaife A, Smith D (2014) Forecast cooling of the atlantic subpolar gyre and associated impacts. *Geophys Res Lett* 41:5167–5174. doi:[10.1002/2014GL060420](https://doi.org/10.1002/2014GL060420)
- Huang B, Hu Z-Z, Jha B (2007) Evolution of model systematic errors in the Tropical Atlantic Basin from coupled climate hindcasts. *Clim Dyn* 28:661–682. doi:[10.1007/s00382-006-0223-8](https://doi.org/10.1007/s00382-006-0223-8)
- Huang B, Banzon V, Freeman E, Lawrimore J, Liu W, Peterson T, Smith T, Thorne P, Woodruff S, Zhang H-M (2015) Extended reconstructed sea surface temperature version 4 (ersst.v4). Part i: Upgrades and intercomparisons. *J Clim* 28:911–930. doi:[10.1175/JCLI-D-14-00006.1](https://doi.org/10.1175/JCLI-D-14-00006.1)
- Huang B, Zhu J, Marx L, Wu X, Kumar A, Hu Z-Z, Balmaseda MA, Zhang S, Lu J, Schneider EK, Kinter JL III (2014) Climate drift of amoc, north atlantic salinity and arctic sea ice in CFSV2 decadal predictions. *Clim Dyn* 44(1):559–583. doi:[10.1007/s00382-014-2395-y](https://doi.org/10.1007/s00382-014-2395-y)
- Hwang Y-T, Frierson DMW (2013) Link between the double-inter-tropical convergence zone problem and cloud biases over the southern ocean. *Proc Nat Acad Sci* 110:4935–4940. doi:[10.1073/pnas.1213302110](https://doi.org/10.1073/pnas.1213302110)
- Jackson LC, Kahana R, Graham T, Ringer MA, Woollings T, Mecking JV, Wood RA (2015) Global and European climate impacts of a

- slowdown of the AMOC in a high resolution GCM. *Clim Dyn* 45:3299–3316. doi:[10.1007/s00382-015-2540-2](https://doi.org/10.1007/s00382-015-2540-2)
- Jin E, Kinter J, Wang B, Park C-K, Kang I-S, Kirtman B, Kug J-S, Kumar A, Luo J-J, Schemm J, Shukla J, Yamagata TT (2008) Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Clim Dyn* 31:647–664. doi:[10.1007/s00382-008-0397-3](https://doi.org/10.1007/s00382-008-0397-3)
- Johnson NC (2013) How many enso flavors can we distinguish? *J Clim* 26:4816–4827. doi:[10.1175/JCLI-D-12-00649.1](https://doi.org/10.1175/JCLI-D-12-00649.1)
- Kang S, Held I, Frierson D, Zhao M (2010) The response of the ITCZ to extratropical thermal forcing: idealized slab-ocean experiments with a GCM. *J Clim* 21:3521–3532. doi:[10.1175/2007JCLI2146.1](https://doi.org/10.1175/2007JCLI2146.1)
- Kharin VV, Boer GJ, Merryfield WJ, Scinocca JF, Lee W-S (2012) Statistical adjustment of decadal predictions in a changing climate. *Geophys Res Lett*. doi:[10.1029/2012GL052647](https://doi.org/10.1029/2012GL052647)
- Kwon Y-O, Alexander MA, Bond NA, Frankignoul C, Nakamura H, Qiu B, Thompson L (2010) Role of the gulf stream and Kuroshio–Oyashio systems in large-scale atmosphere–ocean interaction: a review. *J Clim* 23:3249–3281. doi:[10.1175/2010JCLI3343.1](https://doi.org/10.1175/2010JCLI3343.1)
- Lazar A, Vintzileos AA, Doblas-Reyes F, Rogel P, Delecluse P (2005) Seasonal forecast of tropical climate with coupled ocean–atmosphere general circulation models: on the respective role of the atmosphere and the ocean components in the drift of the surface temperature error. *Tellus A* 57(3):387–397. doi:[10.1111/j.1600-0870.2005.00135.x](https://doi.org/10.1111/j.1600-0870.2005.00135.x)
- Liu X, Yang S, Kumar A, Weaver S, Jiang X (2012) Diagnostics of subseasonal prediction biases of the asian summer monsoon by the NCEP climate forecast system. *Clim Dyn* 41:1453–1474. doi:[10.1007/s00382-012-1553-3](https://doi.org/10.1007/s00382-012-1553-3)
- MacLachlan C, Arribas A, Peterson KA, Maidens A, Fereday D, Scaife AA, Gordon M, Vellinga M, Williams A, Comer RE, Camp J, Xavier P, Madec G (2015) Global seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system. *Q J R Meteorol Soc* 141(689):1072–1084. doi:[10.1002/qj.2396](https://doi.org/10.1002/qj.2396)
- Martin GM, Milton SF, Senior CA, Brooks ME, Ineson S, Reichler T, Kim J (2010) Analysis and reduction of systematic errors through a seamless approach to modeling weather and climate. *J Clim* 23:5933–5957. doi:[10.1175/2010JCLI3541.1](https://doi.org/10.1175/2010JCLI3541.1)
- Meehl GA, Goddard L, Boer G, Burgman R, Branstator G, Cassou C, Corti S, Danabasoglu G, Doblas-Reyes F, Hawkins E, Karspeck A, Kimoto M, Kumar A, Matei D, Mignot J, Msadek R, Navarra A, Pohlmann H, Rienecker M, Rosati T, Schneider E, Smith D, Sutton R, Teng H, van Oldenborgh G-J, Vecchi G, Yeager S (2014) Decadal climate prediction: an update from the trenches. *Bull Am Meteorol Soc* 95:243–267. doi:[10.1175/BAMS-D-12-00241.1](https://doi.org/10.1175/BAMS-D-12-00241.1)
- Mullholland D, Laloyaux P, Haines K, Balmaseda M (2015) Origin and impact of initialization shocks in coupled atmosphere–ocean forecasts. *Mon Weather Rev* 143:4631–4644. doi:[10.1175/MWR-D-15-0076.1](https://doi.org/10.1175/MWR-D-15-0076.1)
- Palmer TN, Alessandri A, Andersen U, Cantelaube P, Davey M, Délecluse P, Déqué M, Díez E, Doblas-Reyes FJ, Feddersen H, Graham R, Gualdi S, Gurmy J-F, Hagedorn R, Hoshen M, Keenlyside N, Latif M, Lazar A, Maisonnave E, Marletto V, Morse AP, Orfila B, Rogel P, Terres J-M, Thomson MC (2004) Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER). *Bull Am Meteorol Soc* 85:853–872
- Phillips TJ, Potter G, Williamson D, Cederwall R, Boyle J, Fiorino M, Hnilo J, Olson J, Xie S, Yio JJ (2004) Evaluating parametrizations in general circulation models: climate simulation meets weather prediction. *Bull Am Meteorol Soc* 85:1903–1915
- Ren H (2008) Predictor-based error correction method in short-term climate prediction. *Prog Nat Sci* 18:129–135. doi:[10.1016/j.pnsc.2007.08.003](https://doi.org/10.1016/j.pnsc.2007.08.003)
- Reynolds R, Smith T, Liu C, Chelton D, Casey K, Schlax M (2007) Daily high-resolution-blended analyses for sea surface temperature. *J Clim* 20:5473–5496. doi:[10.1175/2007JCLI1824.1](https://doi.org/10.1175/2007JCLI1824.1)
- Richter I, Xie S-P (2008) On the origin of equatorial atlantic biases in coupled general circulation models. *Clim Dyn* 31:587–598. doi:[10.1007/s00382-008-0364-z](https://doi.org/10.1007/s00382-008-0364-z)
- Rodwell MJ, Palmer TN (2007) Using numerical weather prediction to assess climate models. *Q J R Meteorol Soc* 133:129–146. doi:[10.1002/qj.23](https://doi.org/10.1002/qj.23)
- Saha S, Nadiga S, Thiaw C, Wang J, Wang W, Zhang Q, den Dool HMV, Pan H-L, Moorthi S, Behringer D, Stokes D, Pea M, Lord S, White G, Ebisuzaki W, Peng P, Xie P (2006) The NCEP climate forecast system. *J Clim* 19:3483–3517. doi:[10.1175/JCLI3812.1](https://doi.org/10.1175/JCLI3812.1)
- Sanchez-Gomez E, Cassou C, Ruprich-Robert Y, Fernandez E, Terray L (2015) Drift dynamics in a coupled model initialized for decadal forecasts. *Clim Dyn*. doi:[10.1007/s00382-015-2678-y](https://doi.org/10.1007/s00382-015-2678-y)
- Schott FA, Xie SP, McCreary JP (2009) Indian ocean circulation and climate variability. *Rev Geophys* 47:RG1002. doi:[10.1029/2007RG000245](https://doi.org/10.1029/2007RG000245)
- Smith D, Scaife A, Boer G, Caian M, Doblas-Reyes F, Guemas V, Hawkins E, Hazeleger W, Hermanson L, Ho C, Ishii M, Kharin V, Kimoto M, Kirtman B, Lean J, Matei D, Merryfield WJ, Müller W, Pohlmann H, Rosati A, Wouters B, Wyser K (2012) Real-time multi-model decadal climate predictions. *Clim Dyn* 41:2875–2888. doi:[10.1007/s00382-012-1600-0](https://doi.org/10.1007/s00382-012-1600-0)
- Smith DM, Allan R, Coward A, Eade R, Hyder P, Liu C, Loeb N, Palmer M, Roberts C, Scaife A (2007) Earth’s energy imbalance since 1960 in observations and CMIP5 models. *Geophys Res Lett*. doi:[10.1002/2014GL062669](https://doi.org/10.1002/2014GL062669)
- Toniazzo T, Woolnough S (2014) development of warm sst errors in the southern tropical Atlantic in CMIP5 decadal hindcasts. *Clim Dyn* 43(11):2889–2913. doi:[10.1007/s00382-013-1691-2](https://doi.org/10.1007/s00382-013-1691-2)
- Vannière B, Guilyardi E, Madec G, Doblas-Reyes F, Woolnough S (2013) Using seasonal hindcasts to understand the origin of the equatorial cold tongue bias in CGCMS and its impact on ENSO. *Clim Dyn* 40(3):963–981. doi:[10.1007/s00382-012-1429-6](https://doi.org/10.1007/s00382-012-1429-6)
- Vannière B, Guilyardi E, Toniazzo T, Madec G, Woolnough S (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. doi:[10.1007/s00382-014-2051-6](https://doi.org/10.1007/s00382-014-2051-6)
- Waters J, Lea D, Martin M, Mirouze I, Weaver A, White J (2014) Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. *Q J R Meteorol Soc* 141:333–349. doi:[10.1002/qj.2388](https://doi.org/10.1002/qj.2388)
- Williams KD, Bodas-Salcedo A, Dqu M, Fermepin S, Medeiros B, Watanabe M, Jakob C, Klein SA, Senior CA, Williamson DL (2013) The transpose-ampii experiment and its application to the understanding of southern ocean cloud biases in climate models. *J Clim* 26:3258–3274. doi:[10.1175/JCLI-D-12-00429.1](https://doi.org/10.1175/JCLI-D-12-00429.1)
- Williams KD, Harris CM, Bodas-Salcedo A, Camp J, Comer RE, Copsey D, Fereday D, Graham T, Hill R, Hinton T, Hyder P, Ineson S, Masato G, Milton SF, Roberts MJ, Rowell DP, Sanchez C, Shelly A, Sinha B, Walters DN, West A, Woollings T, Xavier PK (2015) The met office global coupled model 2.0 (GC2) configuration. *Geosci Model Dev* 8(5):1509–1524. doi:[10.5194/gmd-8-1509-2015](https://doi.org/10.5194/gmd-8-1509-2015)
- Winton M (2000) An elastic-viscous-plastic model for sea ice dynamics. *J Atmos Ocean Technol* 17:525–531
- Wu T, Yu R, Zhang F, Wang Z, Dong M, Wang L, Jin X, Chen D, Li L (2010) The Beijing climate center atmospheric general circulation model: description and its performance for the present-day climate. *Clim Dyn* 34:123–147. doi:[10.1007/s00382-008-0487-2](https://doi.org/10.1007/s00382-008-0487-2)
- Wu T, Li W, Ji J, Xin X, Li L, Wang Z, Zhang Y, Li J, Zhang F, Wei M, Shi X, Wu F, Zhang L, Chu M, Jie W, Liu Y, Wang F, Liu X, Li Q, Dong M, Liang X, Gao Y, Zhang J (2013) Global carbon budgets simulated by the Beijing climate center climate system model for the last century. *J Geophys Res Atmos* 118:4326–4347. doi:[10.1002/jgrd.50320](https://doi.org/10.1002/jgrd.50320)

- Xiao M, Zhang Q, Singh V (2015) Influences of ENSO, NAO, IOD and PDO on seasonal precipitation regimes in the Yangtze river basin, China. *Int J Climatol* 35:3556–3567. doi:[10.1002/joc.4228](https://doi.org/10.1002/joc.4228)
- Xie S-P, Deser C, Vecchi GA, Collins M, Delworth TL, Hall A, Hawkins E, Johnson NC, Cassou C, Giannini A, Watanabe M (2015) Towards predictive understanding of regional climate change. *Nat Clim Change* 5:921–930. doi:[10.1038/nclimate2689](https://doi.org/10.1038/nclimate2689)
- Zhang S (2011) A study of impacts of coupled model initial shocks and state-parameter optimization on climate predictions using a simple pycnocline prediction model. *J Clim* 24:6210–6226. doi:[10.1175/JCLI-D-10-05003.1](https://doi.org/10.1175/JCLI-D-10-05003.1)