CrossMark

Identifying causes of Western Pacific ITCZ drift in ECMWF System 4 hindcasts

Jonathan K. P. Shonk[1] · Eric Guilyardi[1,2] · Thomas Toniazzo[3] · Steven J. Woolnough[1] ·
Tim Stockdale[4]

Abstract The development of systematic biases in climate models used in operational seasonal forecasting adversely affects the quality of forecasts they produce. In this study, we examine the initial evolution of systematic biases in the ECMWF System 4 forecast model, and isolate aspects of the model simulations that lead to the development of these biases. We focus on the tendency of the simulated intertropical convergence zone in the western equatorial Pacific to drift northwards by between 0.5° and 3° of latitude depending on season. Comparing observations with both fully coupled atmosphere–ocean hindcasts and atmosphere-only hindcasts (driven by observed sea-surface temperatures), we show that the northward drift is caused by a cooling of the sea-surface temperature on the Equator. The cooling is associated with anomalous easterly wind stress and excessive evaporation during the first twenty days of hindcast, both of which occur whether air–sea interactions are permitted or not. The easterly wind bias develops immediately after initialisation throughout the lower troposphere; a westerly bias develops in the upper troposphere after about 10 days of hindcast. At this point, the baroclinic structure of the wind bias suggests coupling with errors in convective heating, although the initial wind bias is barotropic in structure and appears to have an alternative origin.

1 Seasonal and decadal forecasting: scientific background

Since Cane et al. (1986) first experimented with El Niño forecasts, the science of seasonal forecasting has grown steadily (Barnston et al. 2012; Smith et al. 2012). Weather centres around the world have since developed their own seasonal forecasting systems, including the ECMWF (Molteni et al. 2011), the UK Met Office (MacLachlan et al. 2014) and NCEP (Saha et al. 2014), and these are now run operationally. However, there remain many challenges to overcome. The numerical models underlying seasonal forecasting systems are often derived from the general circulation models (GCMs) used for climate prediction and suffer from similar systematic biases. These biases develop over the time scale of a typical seasonal forecast and are of comparable magnitude to the seasonal weather anomalies being predicted (and indeed, often greater), implying a significant effect on the forecast.

In seasonal forecasting, two methods are employed to deal with these systematic biases. The first is to initialise the model with observed conditions and correct the forecast afterwards by removing the diagnosed mean drift (that is, the time-dependent systematic bias; Stockdale 1997). The second method seeks to avoid model drift by expressing the observed conditions as anomalies with respect to the observed climatology and initialising the model from its own long-term mean climatology, to which the observed

✉ Jonathan K. P. Shonk
  j.k.p.shonk@reading.ac.uk

[1] National Centre for Atmospheric Science, University of Reading, Reading, UK

[2] LOCEAN, UPMC, Paris, France

[3] Uni Research, Bjerknes Centre for Climate Research, Bergen, Norway

[4] European Centre for Medium-Range Weather Forecasts, Reading, UK

anomalies are added. The forecast anomalies are then used to predict anomalies with respect to the observed climatology. On account of regional differences in specific meteorological regimes, the first of these bias correction method was found to be the most appropriate for seasonal forecasting (Magnusson et al. 2013) and is currently used to account for model biases in the ECMWF System 4 model (Molteni et al. 2011), the GloSea5 model of the UK Met Office (MacLachlan et al. 2014) and NCEP's CFSv2 model (Saha et al. 2014). However, given the comparable size of the biases and of seasonal anomalies, the underlying assumption of linearity in this method is questionable. The ultimate goal towards achieving better predictions is to reduce the systematic biases generated by the numerical models.

Many of these biases have been present for a long time. A review by Neelin et al. (1992) assessed the performance of seventeen GCMs and concluded that in particular the representation of the "warm pool" (a large body of water in the western Pacific with temperatures in excess of 27.5 °C) and of the Pacific "cold tongue" (a region of colder water along the equator; a surface expression of the equatorial ocean circulation) presented a "challenging test". Mechoso et al. (1995) found that all of the eleven coupled GCMs examined in their study misrepresented the Pacific cold tongue as too narrow, too cold and extending too far to the west. Plus, in the Pacific, all of the models showed an unrealistic geographic distribution of rainfall in the inter-tropical convergence zone (ITCZ; a band of low-level wind convergence and precipitation near the Equator). They produced an erroneous second branch of the ITCZ located south of the Equator, either concurrent or alternating with the (observed) northern hemisphere ITCZ. This is often referred to as the "double ITCZ" bias.

Despite much scientific development and technical progress in numerical modelling since these early studies, these same biases are still present in modern GCMs (IPCC AR5; Flato et al. 2013). The cold tongue and double ITCZ biases have persisted through several generations of model intercomparison project (AchutaRao and Sperber 2006; Zheng et al. 2012; Li and Xie 2014; Zhang et al. 2013). A double ITCZ is still found in the climatological structure and the annual cycle of rainfall in many models (Lin 2007; de; Szoeke and Xie 2008). Associated errors include poor representation of variability in sea-surface temperature, most notably of the El Niño Southern Oscillation (Latif et al. 2001; Guilyardi 2006; Misra et al. 2008; Bellenger et al. 2014).

Before these biases can be reduced, or at least their impact on the forecast quality understood, their causes need to be identified. In diagnostic studies following phase five of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012), several possible causes have been proposed. Hwang and Frierson (2013) suggest that extra-tropical cloud biases may be responsible—a result echoed by Li and Xie (2014) and Mechoso et al. (2016), who highlight the impact of cloud biases in the southern hemisphere on the hemispheric energy budgets. Other studies focus on specific processes, such as entrainment characteristics of cumulus cloud in the model parameterisations for deep convection (Hirota et al. 2011). Information about bias origin can also be obtained from sensitivity experiments. Air–sea coupling frequency (Misra et al. 2008; Bernie et al. 2008) and the parameterisations of low-level cloud (Dai et al. 2005) or of moist convection (Song and Zhang 2009; Chikira 2010) have been shown to have an impact on the double ITCZ problem in different models.

In this study we take the alternative approach of focussing on the model drift rather than the model bias—that is, examining how the simulated climate evolves from the observed state prescribed in the initialisation towards its own "preferred" state represented by the biased model climatology. Whether as a result of physical constraints, realistic physical mechanisms or model tuning, many different processes interact and feed back on one another to contribute to the development of a certain systematic bias. However, the first few days of a suitably initialised climate simulation can provide specific information on the events that lead to the biases before feedbacks take over. Phillips et al's (2004) proposal to evaluate initialised climate integrations as a model validation tool has gained acceptance as a useful test-bed for development work on physics parameterisations in atmospheric models (Williams and Brooks 2008; Boyle et al. 2008; Martin et al. 2010) and for the evaluation of rapidly developing model errors (the Transpose Atmospheric Model Intercomparison Project; Williams et al. 2013).

A similar approach has been employed to investigate the evolution of coupled biases in a number of GCMs. Huang et al. (2007) and Toniazzo and Woolnough (2014) explored the evolution of tropical Atlantic biases in two versions of the NCEP CFS system and in two more CMIP5 GCMs. Vannière et al. (2012) examined the evolution of the Pacific cold tongue bias in the ENSEMBLES database (Weisheimer et al. 2009), comprising five operational seasonal forecasting models. In their study, they hypothesised that the cold tongue bias was generated by biases in zonal wind and surface ocean mixing. They attempted to test this hypothesis by running forced ocean-only model simulations, with winds prescribed from coupled hindcast experiments. Both Vannière et al. (2012) and Toniazzo and Woolnough (2014) found that similar biases in different models could be the result of different initial drifts.

In this study, we analyse the initial bias drift in operational hindcasts made using the ECMWF System 4 model to examine the evolution of systematic biases in the tropical

Pacific. We focus on biases in the western tropical Pacific, where a tendency is found for the model ITCZ to drift to the north. By comparing the fully coupled operational hindcasts and a corresponding set of atmosphere-only hindcasts (where the atmosphere component alone is run and forced by sea-surface temperatures from reanalysis), we can build up a chronology of bias development in different meteorological fields and build a hypothesis about the possible source of the systematic bias. Section 2 introduces our method and describe the data that we use in more detail. Section 3 presents the application of the drift analysis method to the data, as a tool to identify processes linked to the drift of the ITCZ. In Sect. 4, we explore the short-term development of the atmospheric fields in more detail; in Sect. 5, we provides a summary of our results.

2 Sources of data

Throughout this investigation, we use data from hindcasts performed using the ECMWF System 4 seasonal forecasting model, which couples version 3.0 of the Nucleus for European Modelling of the Ocean (NEMO; Madec 2008) to cycle 36R4 of the Integrated Forecast System (IFS; Molteni et al. 2011). NEMO is run using the ORCA1 grid with a horizontal resolution of about 1°, a 0.3° refinement at the Equator and 42 vertical levels. The IFS is an atmosphere model run on a spectral decomposition with a triangular truncation at wavenumber 255 and on 91 vertical levels. Grid point calculations in the IFS are performed on an N128 Gaussian grid, which gives a resolution of about 0.7°. A detailed description of the configuration of System 4 is given by Molteni et al. (2011).

Operational hindcasts from System 4 are available initialised on the first of every month from 1981 to the present day and running out to 7 months ahead. For the rest of this study, we refer to these as "coupled" hindcasts. In addition, we analyse a set of "uncoupled" hindcasts, where the atmosphere component is forced by prescribed sea-surface temperature (SST) from version 2 of the Optimal Interpolation Sea-Surface Temperature dataset (OISSTv2; Reynolds et al. 2002). Uncoupled hindcasts are available over the same set of years as the coupled hindcasts, but with initialisation dates limited to 1 February, 1 May, 1 August and 1 November of each year, and running out to just 4 months ahead.

The full set of available simulations is extensive, and we have investigated ways to make the volume of data for the present analysis more manageable. First, we select a 14-year subset of these hindcasts, spanning years 1996–2009, and exclude initialisation dates that are not common to the coupled and uncoupled models. We explored the sensitivity of the diagnosed ensemble-mean biases to the number of ensemble members considered (up to 51 being available, depending on initialisation date) and found that eight was sufficient to provide a reliable estimate. The systematic biases discussed in this paper are estimated from an average over eight ensemble members and over all 14 years, giving 112 separate model integrations for each initialisation date of the year.

In the period 1996–2009, the operational hindcasts are initialised from the atmospheric fields of the ERA-Interim dataset (Dee et al. 2011) and, for the coupled hindcasts, from ocean fields taken from the ORA-S4 reanalysis dataset (Balmaseda et al. 2013) and SSTs derived from the OISSTv2 product (Reynolds et al. 2002). We focus on this period for our analysis on account of the consistency of the initialisation product, which is derived from the fixed version of IFS used for ERA-Interim (cycle 31R2). As this is not the same version of the IFS used in System 4, there is the chance of an incompatibility of the data leading to a "shock" (a rapid adjustment of the model caused by an imbalance at initialisation). A comparison between the operational analysis of 1 February 2011 (which used the same IFS cycle as System 4) and the corresponding reanalysis from ERA-Interim shows that the differences are much smaller than the systematic biases after a few (1–2) days into the hindcasts. For consistency, however, we exclude hindcasts from 2010 onwards from this analysis, as these were initialised with operational atmospheric analyses using a variety of IFS versions as the operational analysis system was upgraded.

System 4 biases are diagnosed as differences between model output and various observational and reanalysis datasets. For brevity throughout, we will collectively refer to these as "observations". Where possible, the observation data sources are chosen to be representative of the products used to initialise the System 4 hindcasts, in order to facilitate the interpretation of model drift as the evolution of a bias away from the "observed" state. Atmospheric winds and mean-sea-level pressure are obtained from ERA-Interim, SSTs from OISSTv2 (Reynolds et al. 2007), and wind stress from the TropFlux product (Praveen Kumar et al. 2012). We use the Objective Analysis Flux (OAFlux) dataset for sensible and latent heat fluxes (Yu and Weller 2007), and data from ISCCP for short-wave and long-wave fluxes (Zhang et al. 2004). The OAFlux and ISCCP datasets are not directly associated with ERA-Interim data, and for these variables model and observations may differ at initialisation time. To estimate the uncertainty in these observational fluxes, we repeat some of the analysis with surface heat fluxes from TropFlux. Observational rainfall data in this study comes from the CPC Merged Analysis of Precipitation (CMAP; Xie and Arkin 1997) and the Tropical Rainfall Measuring Mission (TRMM; Kummerow et al. 2000). The TRMM product is only available from 1998

onward, but it provides rainfall data at higher spatial and temporal resolution than CMAP—0.25° rather than 2.5°, and daily rather than pentadal accumulations. In this study, we compare our model rainfall with both datasets.

Model biases are defined as the difference between the ensemble mean of the hindcasts and the observed climatology over the same period. As an estimate of the uncertainty in the climatology of the model (and the biases) we use two measures. The first estimates the uncertainty in the ensemble mean by the interannual standard deviation of the ensemble mean. The second measure estimates the uncertainty in the bias associated with comparing a single realisation of the observed climatology with many realisations of the forecast climatology. This is achieved by calculating the standard deviation of a forecast climatology created by sampling one randomly chosen ensemble member for each year. This second approach is used only in Fig. 8.

3 Systematic biases at the ocean surface

The northward shift of the ITCZ in the Pacific can be seen in Fig. 1, which shows rainfall biases in the coupled hindcasts at a lead time of 7 months. Lead time is hereafter

defined as the timespan between hindcast initialisation time and validity time. The northward ITCZ bias spans the Pacific, but it is most marked in the western and central Pacific. For the rest of this study, we focus our attention on the western Pacific.

We define a region labelled "P", also shown in Fig. 1, extending from 125°E to 175°E and from the Equator to 20°N, and define the ITCZ location in terms of the zonal mean rainfall rate across this region. We define a threshold of 50% of the peak zonal-mean rainfall intensity. Within the range of latitudes where rainfall exceeds this threshold, we take the latitude of the rainfall centroid. In contrast to other definitions (such as simply taking the location of the peak zonal-mean rainfall, for example), this definition returns a smoothly varying quantity that has limited sensitivity to the spatial resolution of the rainfall dataset.

Figure 2 shows a comparison of the location of the ITCZ in the P box between observations (TRMM) and both coupled and uncoupled hindcasts as a function of validity time (that is, the month for which the hindcast is valid). The initialisation date is colour-coded for the hindcasts, and the TRMM climatology from 1998 to 2009 is shown in black. The observed ITCZ in the P box moves from about 4°N in northern winter to 10°N in northern summer. The
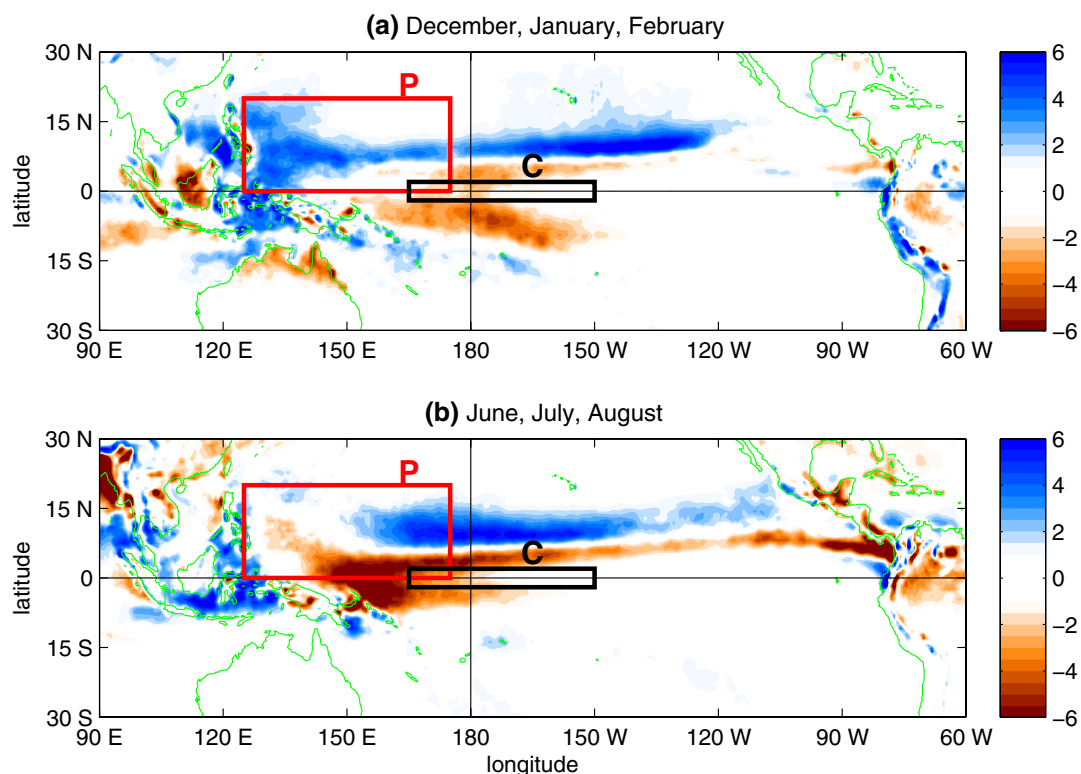

Fig. 1 Map of biases in rainfall (in mm d^{-1}) in the ECMWF System 4 coupled model after 7 months of hindcast. Biases are calculated as monthly means with respect to the TRMM dataset and then averaged over **a** December to February and **b** June to August of years 1998–2009. The *boxes* marked P and C are referred to at various points in this study
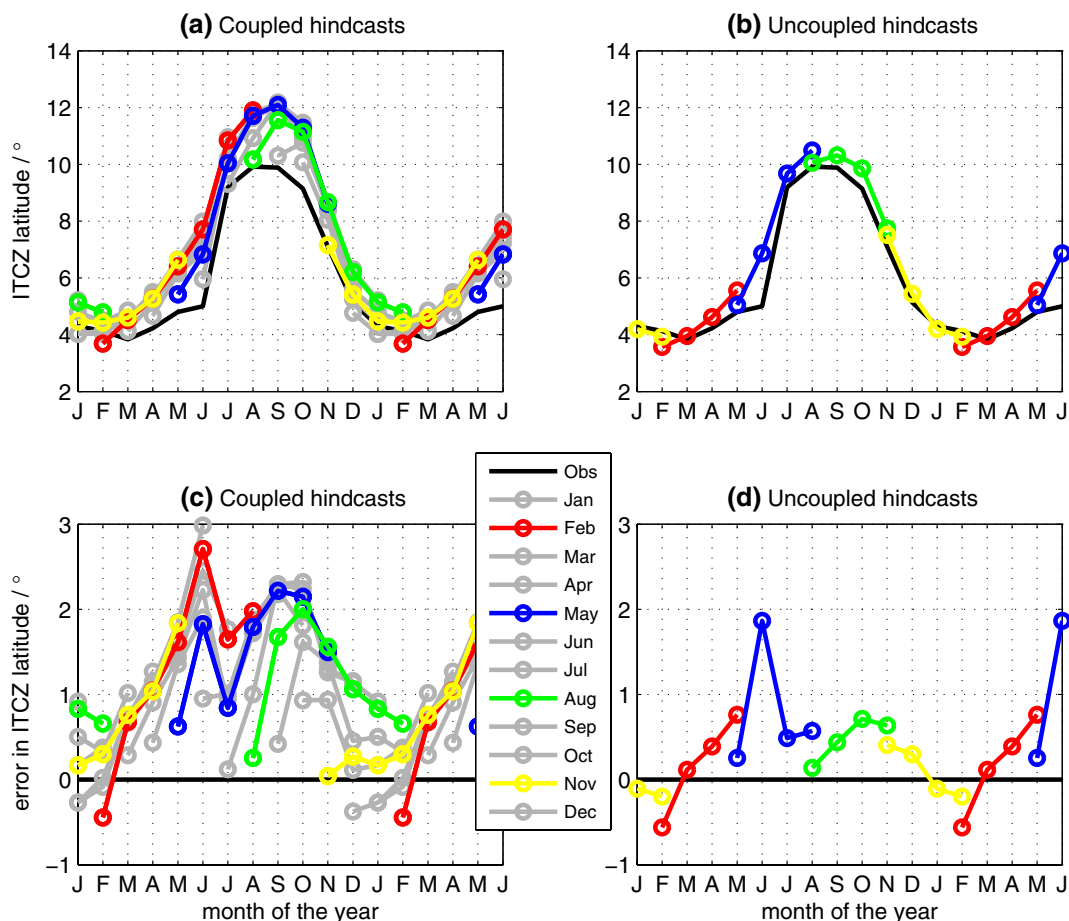
Fig. 2 Evolution of the location of the ITCZ throughout the year in the "P" box, determined for observation data from TRMM (*black line*) and model hindcast data from ECMWF System 4 (*coloured lines*). The *top row of panels* shows the absolute position of the ITCZ; the *bottom row* shows the bias in location with respect to the observa-

tion, averaged over years 1998–2009. The *panels* on the *left* show the position for the coupled model; those on the *right* show the position in the uncoupled model. Common available start months have been highlighted in *colour*; the other months in the coupled hindcasts are shown in *grey*

simulated ITCZ has an additional northward drift, resulting in a monthly-mean northward bias of between 0.5° and 3° of latitude, depending on season (Fig. 2c, d), with a maximum in the summer months. The drift occurs during the first 3 months of hindcast and, by month 4, the ITCZ has settled to its position in the model climatology. A similar but smaller northward drift is seen in the uncoupled hindcasts, suggesting an amplification of the atmospheric model bias through air–sea coupling. As only four-month-long hindcasts are available from the uncoupled model, however, the climatology of the uncoupled ITCZ position is less clear. Use of the CMAP dataset in place of TRMM produces no qualitative difference for the diagnosed ITCZ drift in either the rate or magnitude (figure not shown), despite the differences between the algorithmic methods used for the two datasets (Kummerow et al. 2000).

We proceed by focussing on hindcasts starting on 1 February and 1 August. There are two main reasons

for this choice. First, the observed seasonal march of the ITCZ is strongest approaching the equinoxes, moving northwards in February and southwards in August. Second, analysis of the hindcasts initialised in May and November reveals similar bias characteristics to those in February, while the behaviour of August is somewhat different.

The differences between coupled and uncoupled forecasts suggest that the drift is influenced by developing biases in SST. The bottom panels of Fig. 3 compare the observed zonal-mean SST in the P box with the simulated SST at a lead time of 1 and 4 months. A cold bias can be seen to develop within a few degrees of the Equator, reaching between 0.2 and 0.5 °C on the Equator by the fourth month of hindcast. This bias develops, in fact, quite rapidly, with a cooling of a few tenths of a degree Celsius within 1 month after initialisation, giving an initial cooling rate of about 0.01 °C d^{-1}.
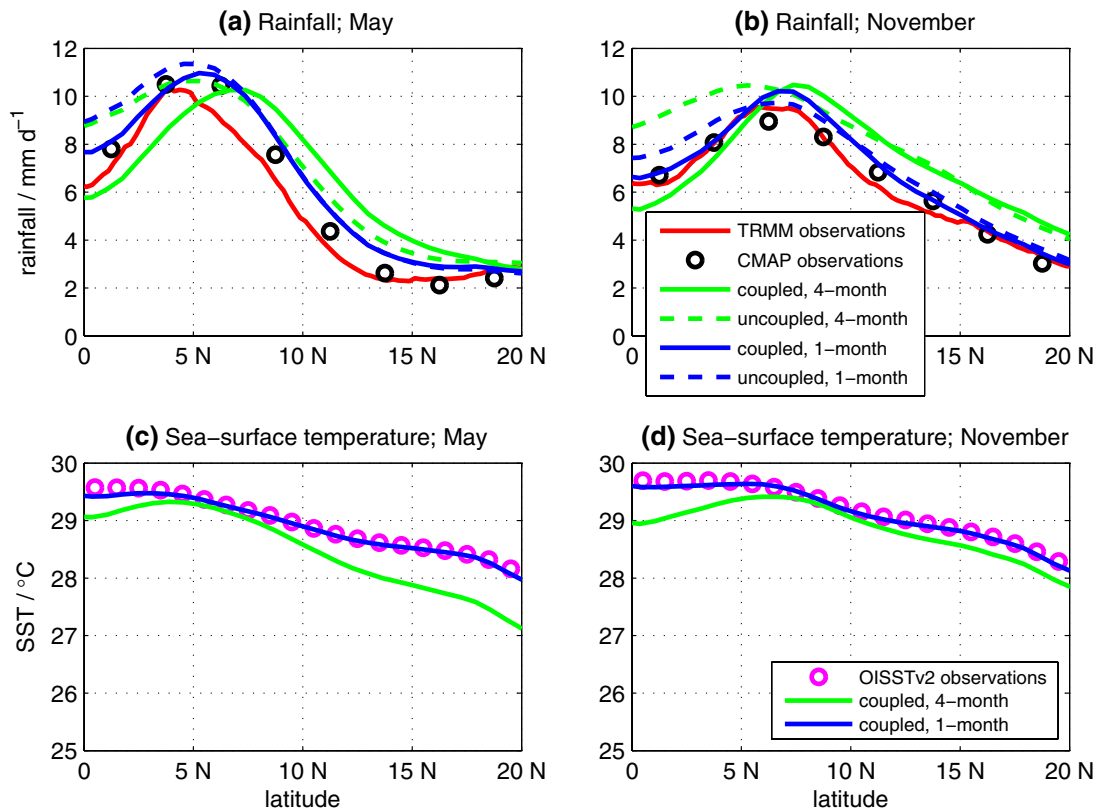
Fig. 3 Zonal mean precipitation (*top row*) and sea-surface temperature (*bottom row*), averaged across the width of the "P" box. Hindcasts for May (*left*) and November (*right*) are shown. The rainfall panels show observed values of rainfall from CMAP (*black circles*) and TRMM (*red line*), averaged over years 1998–2009; the sea-surface temperature panels show observations from OISSTv2 (*purple circles*) from 1996–2009. The *blue* and *green lines* show the values from the hindcasts, averaged over the same sets of years. The *green lines* are the hindcast of that month at a lead time of 4 months; the *blue lines* are the hindcast of that month at a lead time of 1 month. *Solid lines* indicate the coupled model; *dashed lines* indicate the uncoupled model

As a result, the simulated meridional gradient in SST near the Equator, where the maximum SST occurs, also changes rapidly. Depending on season, the cooling either reduces or reverses the gradient such that the warmest temperatures move away from the Equator, which suppresses rainfall on the equatorward side of the ITCZ and drives the ITCZ northwards. The rate at which this gradient reversal takes place is indicated in Fig. 4a, b, which shows the zonal mean SST across the width of the P box at various lead times within the first 50 days for hindcasts starting in February and August. Figure 4c, d shows the evolution in the observations from OISSTv2 over the same period. By about 20 days into the hindcasts, a spurious local minimum in SST appears over the Equator in the simulations.

Figure 5 shows longitude–time plots of the SST bias averaged between 2°S and 2°N. Contours are filled within the region where the bias is outside the 95% statistical envelope of interannual variability. The cold bias in the equatorial Pacific develops in two separate regions. Between 160°W and 110°W, SST cools within the first 5 days of hindcast in a pattern that is reminiscent of the classic cold tongue bias associated with errors in equatorial ocean upwelling and thermocline tilt (for example, Vannière et al. 2012). In the west, the bias develops more gradually and becomes significant only after about 10 days. It develops between 150°E and 170°E in February, and between 180° and 160°W in August. In both seasons, it reaches about −0.5 °C in 20 days, comparable with the bias identified in Fig. 4. Both spatially and temporally, these biases appear separate from the stronger cool bias developing in the eastern Pacific.

As a first step towards identifying potential causes of this SST bias, we correlate it with biases in rainfall rate, zonal and meridional wind stress, and downward heat fluxes at the ocean surface. Of these variables, the zonal wind stress displayed the most evident drift, shown in Figs. 6a and 7a for the equatorial band between 2°S and 2°N. In February, a "pulse" of easterly bias occupies much of the western equatorial Pacific, appearing near the Maritime Continent and peaking after 10 days of hindcast. By day 50, it has largely faded. The location and timing of the SST bias tendency and the zonal wind stress bias approximately match,
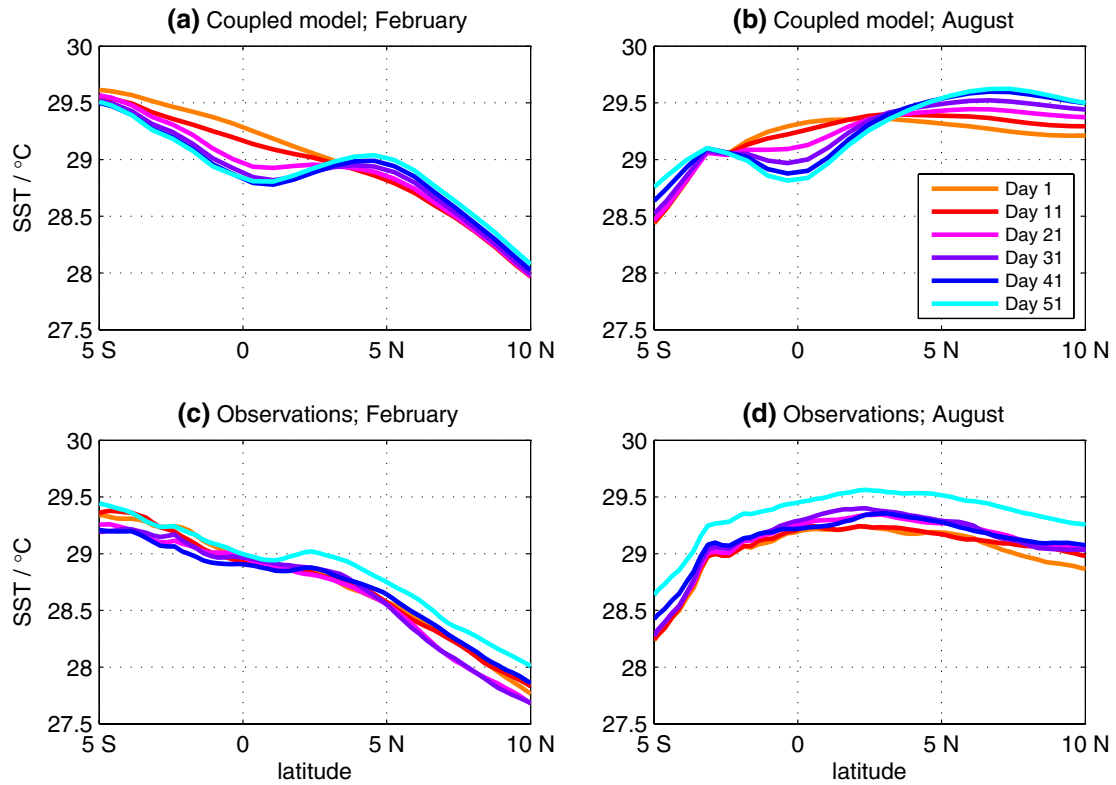
Fig. 4 Zonally averaged sea-surface temperature across the width of the P box at various times in the hindcast (see *legend*), from data starting on 1 February (*left*) and 1 August (*right*). The *top panels* show the sea-surface temperature from the coupled model; the *bottom panels* show the same, but from the observations (OISSTv2)
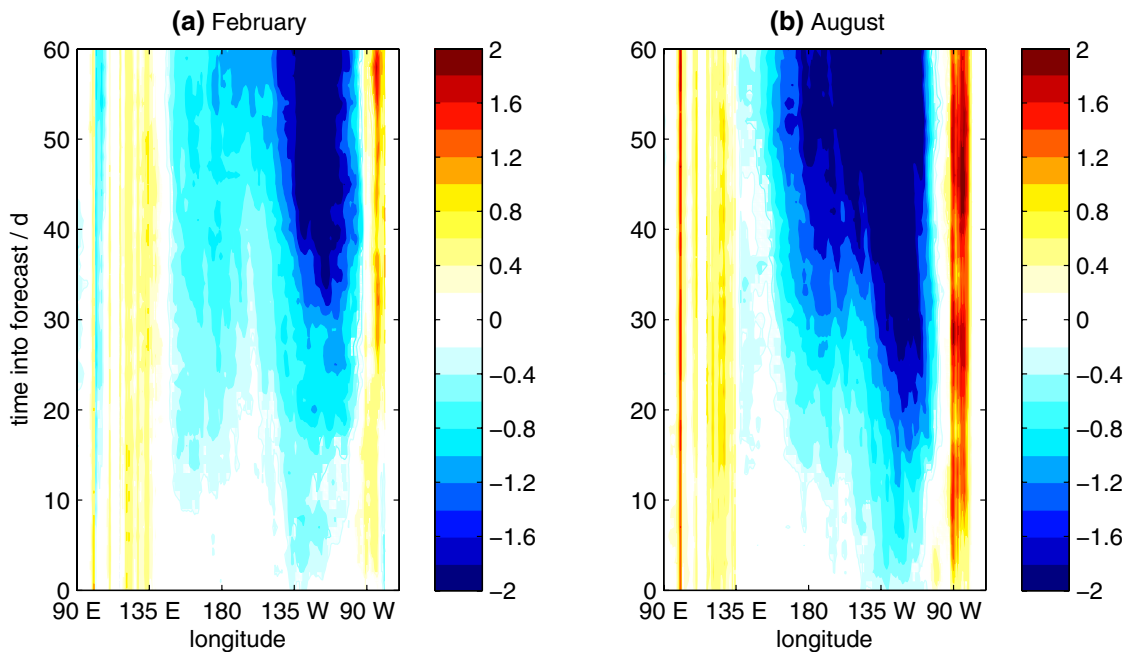

Fig. 5 Longitude—time plots of sea-surface temperature biases, averaged over latitudes 2° S to 2° N and years 1996–2009, for hindcasts initialised on **a** 1 February and **b** 1 August. Biases are in °C and calculated with respect to OISSTv2. *Filled contours* show the developing cold bias in regions of statistical significance; insignificant regions are indicated by *open contours*
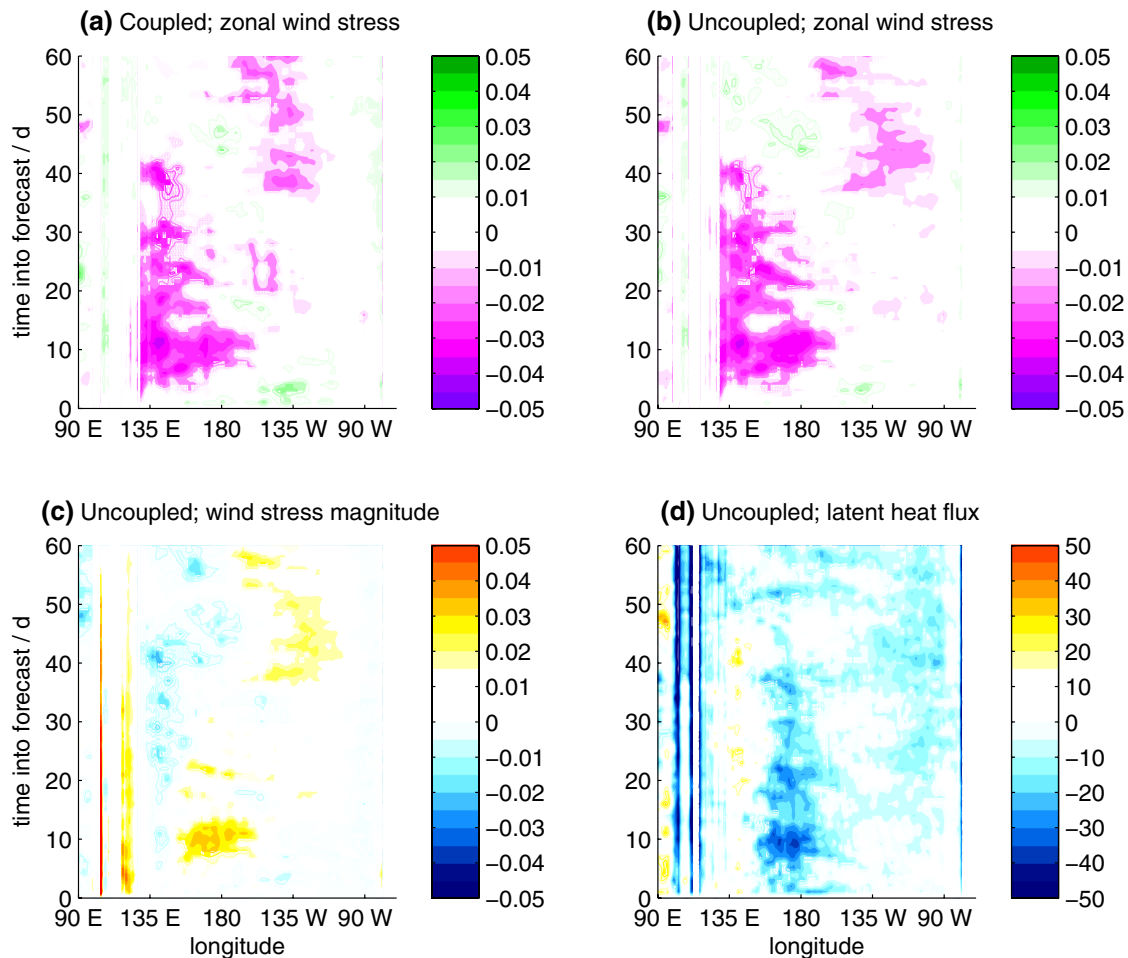
Fig. 6 Longitude—time plots of developing biases, as Fig. 5, averaged over the same years and latitude range, but for **a** zonal wind stress from the coupled model and **b** zonal wind stress, **c** wind stress magnitude and **d** surface latent heat flux from the uncoupled model. Wind stress biases are displayed in N m^{-2} and with respect to Trop-Flux; latent heat biases are displayed in W m^{-2} and with respect to OAFlux. Significant biases are indicated by *filled contours*; insignificant biases are indicated by *open contours*. Biases in hindcasts starting on 1 February are shown here

suggesting a link between the two fields. The same is true in August, where both biases are located further east, in the central Pacific, reaching their peak after 10 days but still persisting after 60 days.

The uncoupled hindcasts give some insight as to whether the SST biases are driving biases in the wind stress, or whether the wind stress biases are cooling the SST. Figures 6b and 7b show the zonal wind stress bias in the uncoupled model. The similarity with the coupled model biases in the first 20 days or so is remarkable, and suggests that the cooling is a response to the wind stress bias. After 20 days, the bias patterns in the coupled and uncoupled models start to differ—likely a result of the feedback of the SST biases on the atmosphere.

We now probe the link between the biases in SST and zonal wind stress. In the area of the cold tongue, tendencies in SST are dominated by the effects of surface heat fluxes and equatorial upwelling (Vialard et al. 2001), both of which depend on the prevailing easterly surface winds. An easterly wind bias tends to enhance both upwelling-driven and evaporative cooling. We find that the latent heat flux bias appears to correlate well with the easterly wind stress biases and the SST drift. Figures 6d and 7d show the latent heat flux biases in the uncoupled model—in the first 20 days, a cooling bias develops over the central Pacific at the same time as the wind stress bias.

However, the zonal wind-stress bias is not co-located with the cooling trend. This is related to the background state, which consists of easterly winds prevailing in the central equatorial Pacific, but lighter winds of variable direction further west. Figures 6c and 7c show the bias in wind stress magnitude in the uncoupled model—to the east, the easterly bias implies higher winds, which enhance surface turbulent fluxes; to the west, instead, the
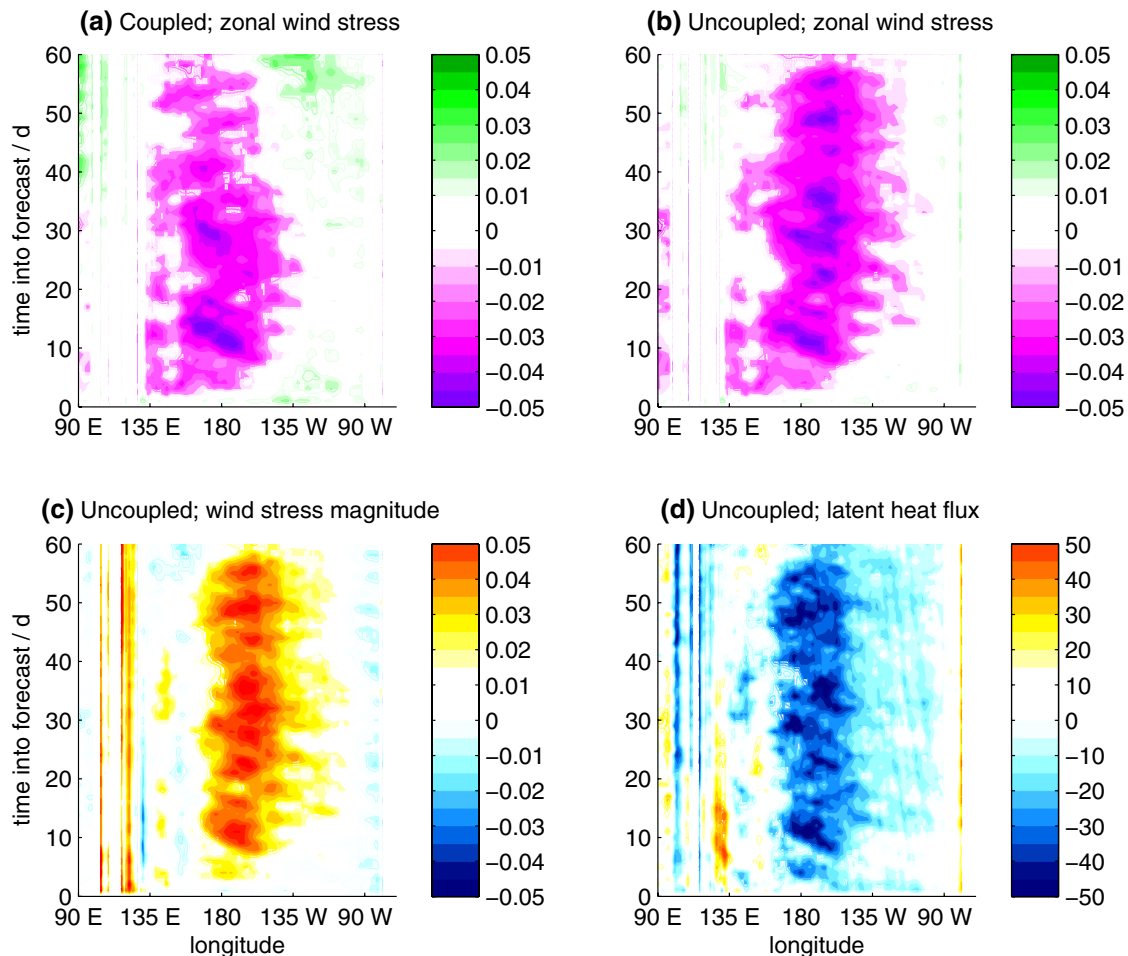
Fig. 7 Longitude—time plots as in Fig. 6, but for hindcasts starting on 1 August

wind stress magnitude is hardly affected by the easterly bias.

We quantify these results in terms of bias time-series in the "C" box, marked on the map in Fig. 1 between 165°E to 150°W and 2°S to 2°N, and spanning the area where easterly trade winds occur. Figure 8 shows the evolution of the mean biases in the C box for both coupled and uncoupled hindcasts. In February, the correlation of the biases in the C box can clearly be seen, both for the coupled and uncoupled models. The zonal wind stress bias peaks after 10 days (Fig. 8c), which is reflected in the bias in the wind stress magnitude (Fig. 8e). The latent heat flux bias also peaks at the same time scale (Fig. 8g). The SST starts cooling after about 10 days, when the wind stress bias is at its greatest, and continues cooling steadily at a rate of about 0.018 °C d^{-1} (Fig. 8a). The biases in August show a similar correlation between these variables, with an initial peak in bias in zonal wind stress (Fig. 8d), wind stress magnitude (Fig. 8f) and latent heat (Fig. 8h). Biases however persist through to day 60 instead of fading, and the cooling is

therefore stronger (Fig. 8b), with a rate between day 10 and day 60 of about −0.031 °C d^{-1}. It should also be noted that the bias in the short-wave component of the surface flux provides additional cooling of comparable magnitude to that of the latent heat flux (Fig. 8i, j). The location and timing of the short-wave bias development is, however, very different.

The error bars in Fig. 8 indicate an estimate of the 10th and 90th percentiles of the bias spread across ensemble members. After 10 days of hindcast, whether coupled or not, the bias is consistent across the realisations, and well separated from zero, indicating that it is unlikely that a climatology based on any individual realisation of the hindcasts would reproduce the observed climatology through this period.

In Fig. 8, we also consider the uncertainty arising from the available observations of surface fluxes. The surface energy budget is challenging to determine and differences exist between different products (Zhang et al. 2004; Praveen Kumar et al. 2012). As a rough quantification of
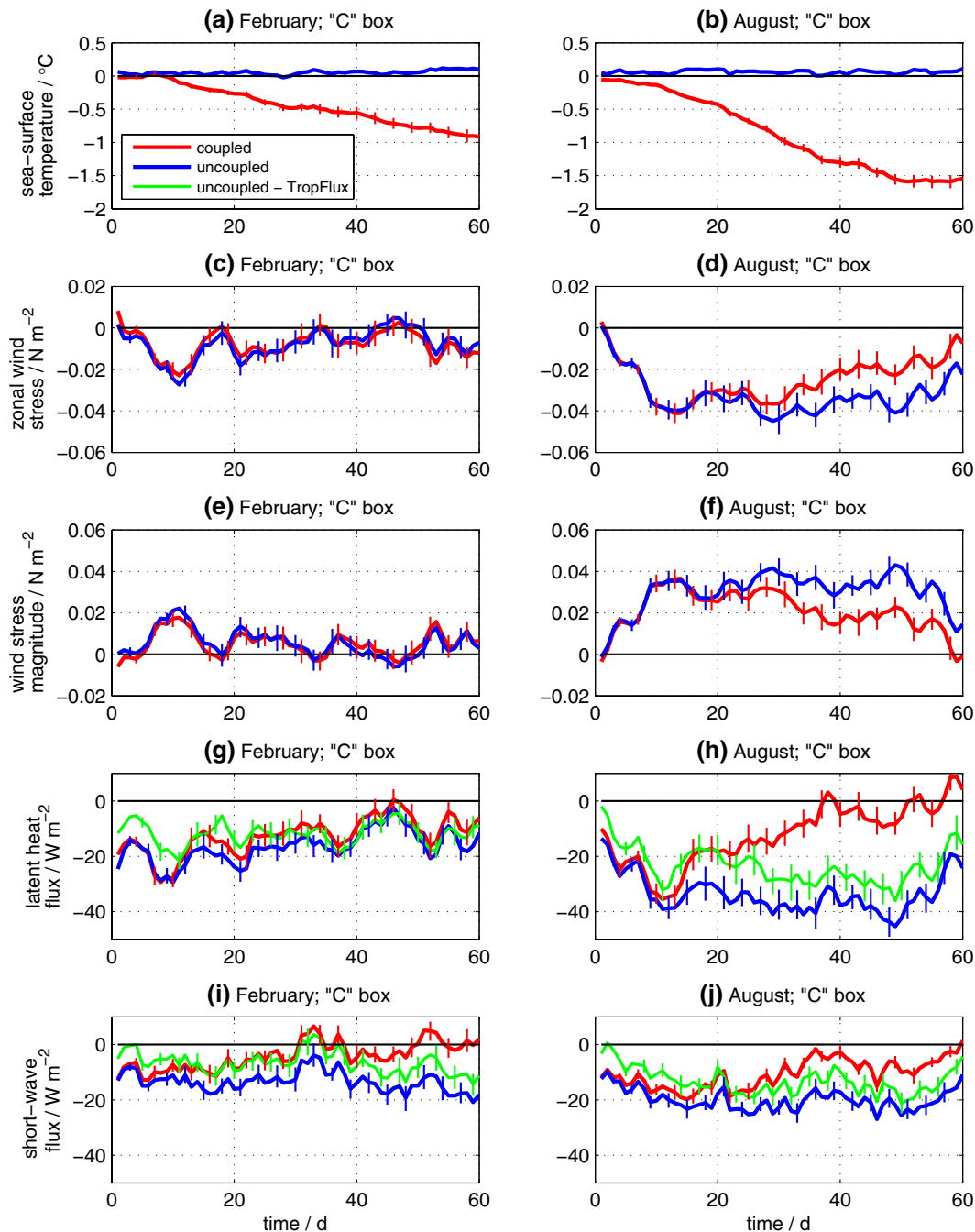
Fig. 8 Summary plots of developing biases in the coupled (*red*) and uncoupled (*blue*) models, averaged over the "C" box (see Fig. 1) and years 1996–2009. Biases in **a, b** sea-surface temperature are with respect to OISSTv2; biases in **c, d** zonal wind stress and **e, f** wind stress magnitude are with respect to TropFlux; biases in **g, h** latent heat flux are with respect to OAFlux; biases in **i, j** short-wave flux are with respect to ISCCP. Extra *green lines* in **g–j** show the uncoupled model flux biases with respect to TropFlux. Biases are shown for hindcasts starting on 1 February (*left column*) and 1 August (*right column*). The *error bars* give an indication of the size of the spread in biases across ensemble members

this uncertainty, we evaluate the biases in the C box using TropFlux fluxes. Compared with ISCCP's short-wave flux product and OAFlux's latent heat product, the downward heat fluxes are larger by about 10 W m^{-2} and 20 W m^{-2} respectively. The simulated flux biases in the uncoupled model with respect to TropFlux are shown as green dashed lines in Fig. 8g–j. Near the beginning of the forecasts, the estimated differences between observation and model are reduced. This is most likely due to the fact that the Trop-Flux merges direct observations with ERA-Interim.

We quantify the effect of flux biases (and their uncertainties) on the SST using a simple heat budget model of the ocean mixed layer. We extract the depth of the mixed layer from the coupled model hindcasts and compare the cooling found in the C box with the cooling that would result from the contribution of the latent heat flux bias. In February, the latent heat flux bias implies a cooling over the first 60 days of between 0.65 °C (OAFlux) and 0.45 °C (TropFlux), compared with a diagnosed coupled SST bias of about −1.0 °C. In August, the situation is more complicated as the cold tongue bias extends into the C box during the 60-day hindcast period. In this region, additional processes govern the temperature of the mixed layer (Vialard et al. 2001), and a budget based on surface heat fluxes alone is inadequate. After 30 days of hindcast, the SST bias is about −1 °C. Latent heat flux biases alone imply a cooling of between 0.4 °C (OAFlux) and 0.1 °C (TropFlux). A more detailed analysis of the mixed-layer heat budget is not possible with the data available to us.

4 Examining drifts in atmospheric fields

A picture of the chain of events leading to the northward drift of the ITCZ in the ECMWF System 4 model is emerging from our analysis. A zonal wind stress bias develops over the western equatorial Pacific, which leads to cooling of the equatorial SST and the subsequent northward drift of the ITCZ. Here we extend the analysis to the depth of the troposphere. Given the similarity of the early bias evolution in the coupled and uncoupled hindcasts, we restrict our analysis to the latter.

Wind stress biases may be induced by biases in the free troposphere or by biases in boundary layer properties. Figure 9 shows longitude–height plots of the bias in zonal wind in the uncoupled model after 1–3, 5 and 10 days of hindcast. In both seasons, a rapidly developing easterly bias is initially confined below 800 hPa before extending upwards to a pressure height of about 600 hPa after 5 days. This represents an acceleration of the wind and can be related to the model easterly wind-stress bias. By 10 days into the hindcasts (Fig. 9i, j), the low-level easterly bias has a magnitude of about 5 m s^{-1}, and is accompanied by upper tropospheric westerly biases of similar magnitude, resulting in a bias pattern reminiscent of an enhanced Walker circulation. The westerly acceleration of tropospheric air suggests additional westerly drag associated with the surface easterly bias, excluding local boundary layer processes (which would be force-balanced locally) as its source. The pattern dominates the western and central Pacific in both seasons, with a westward extension of the low-level easterly bias extending out across the Indian Ocean in the August hindcasts.

The baroclinic structure of the wind biases at day 10 and beyond further suggests that the circulation biases at this time may be coupled to vertical motion associated with excessive rainfall and convective heating over the Maritime Continent and western Pacific. However, the evidence indicates that the initial development of the low-level wind bias is independent of such coupling, even though diabatic heating may substantially strengthen the initial bias and ultimately control the large-scale bias pattern in the model climatology.

Figures 10 and 11 show maps of the biases in rainfall, mean-sea-level pressure and 850 hPa winds. For the February hindcasts, the low-level zonal wind bias is clearly visible in the western and central Pacific in the first 10 days (Fig. 10a) but the rainfall biases are small and there is little evidence of a large-scale heating anomaly. In the second 10 days, there are larger rainfall biases of order 5–10 mm d^{-1} over a large region of the Maritime Continent (Fig. 10b), with low-level wind and mean-sea-level pressure bias patterns that are reminiscent of the Gill (1980) response to equatorial heating.

In the August hindcasts, the picture is further complicated by the circulation associated with the Asian summer monsoon. As in the February hindcasts, the low-level easterly bias in the equatorial western Pacific in the first 10 days does not appear to be related to biases in the rainfall over the Maritime Continent region (Fig. 11a), but there are large rainfall biases to the south of India and in the southern Bay of Bengal. This heating bias is associated with low-level westerly wind biases in the west and southeasterly biases in the east, the latter of which appear to be an extension of the easterly wind stress bias over the equatorial western Pacific. The biases in days 11–20 have similar structure to those in the first 10 days, but with increased magnitude (Fig. 11b).

From this analysis it is difficult to conclusively establish a link between the initial equatorial western Pacific easterly wind bias evolution and errors in the convective heating, or indeed to rule it out. In the first 10 days of the February hindcasts, rainfall biases over the Maritime Continent are small compared to the biases that develop afterwards, but the wind anomalies are of a similar magnitude. It thus appears that this initial wind (and wind stress) bias is not caused by errors in the convective heating, but that excess rainfall over the Maritime Continent region may develop in response to the additional moisture convergence and perhaps subsequently reinforces the wind bias. In August, rainfall errors develop at an early stage over the monsoon region and it is possible that the western Pacific equatorial wind biases are related to these. Further analysis of the relationship between the diabatic heating and circulations and their vertical structure is limited by the lack of appropriate diagnostics: for example, profiles of vertical velocity
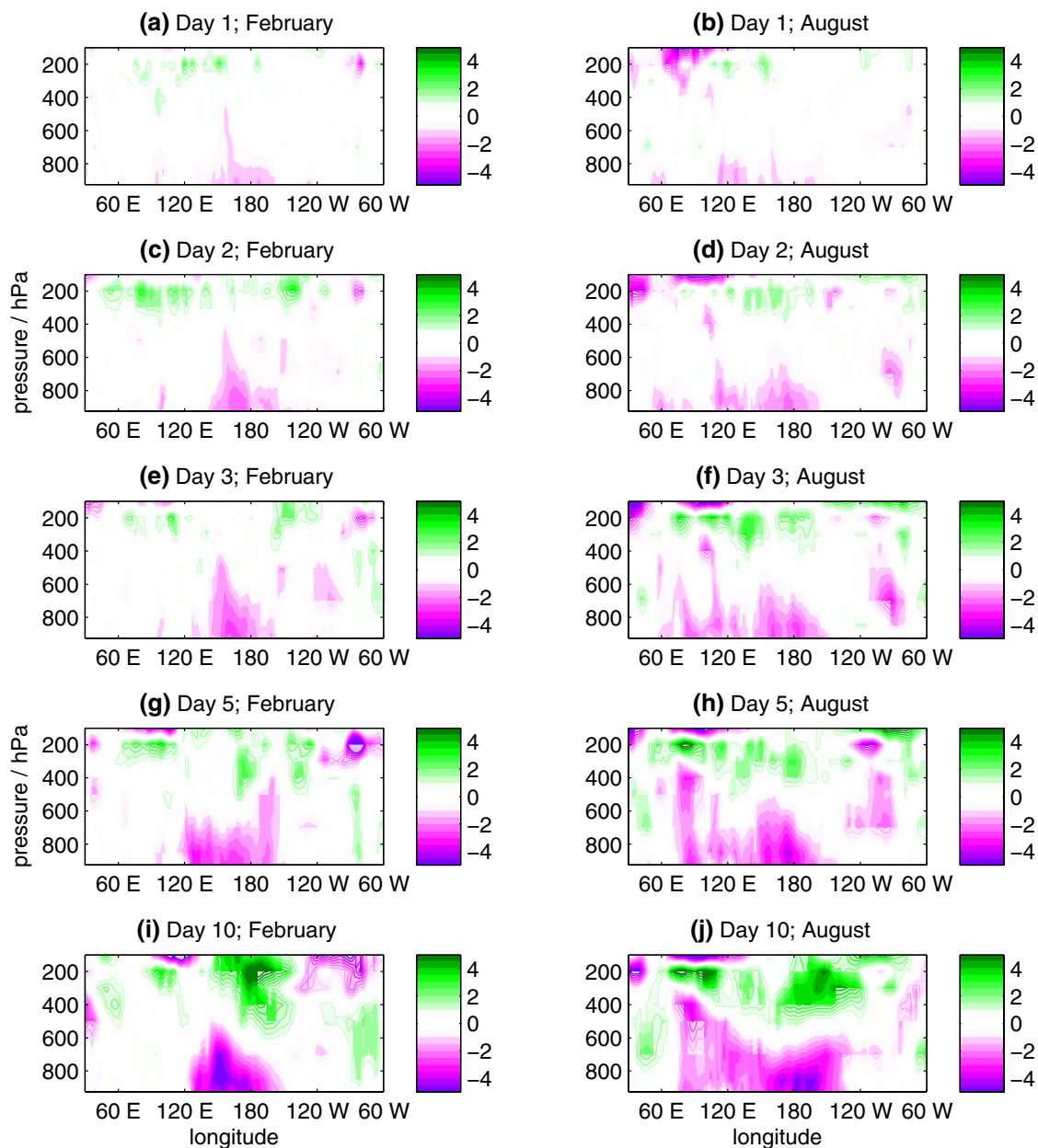
Fig. 9 Longitude—height plots of bias in zonal wind, averaged over latitudes 5° S to 5° N and years 1996–2009, for hindcasts initialised on 1 February (*left column*) and 1 August (*right column*). Biases are in m s^{-1} and calculated with respect to ERA-Interim for five different hindcast lead times. Significant biases are denoted by *filled contours*; insignificant regions are indicated by *open contours*

are not available from the hindcast datasets. Dedicated work with the ECMWF model developers, likely involving sensitivity simulations, is needed to go further.

5 Summary and conclusions

Systematic biases present a challenge to seasonal forecasting and, in this study, we analyse the evolution of biases in the western tropical Pacific in the ECMWF System 4 model, which is used operationally for seasonal forecasting. In this region, the model ITCZ is spuriously displaced north of its observed location by between 0.5° and 3° of latitude, depending on the season. Over the first few months of hindcast, the model ITCZ drifts away from the observed location and reaches this preferred position by about the fourth month of hindcast. The magnitude of the drift is found to be dependent on coupling: in atmosphere-only ("uncoupled") hindcasts, forced by observed sea-surface temperature (SST), the ITCZ drifts to the north to
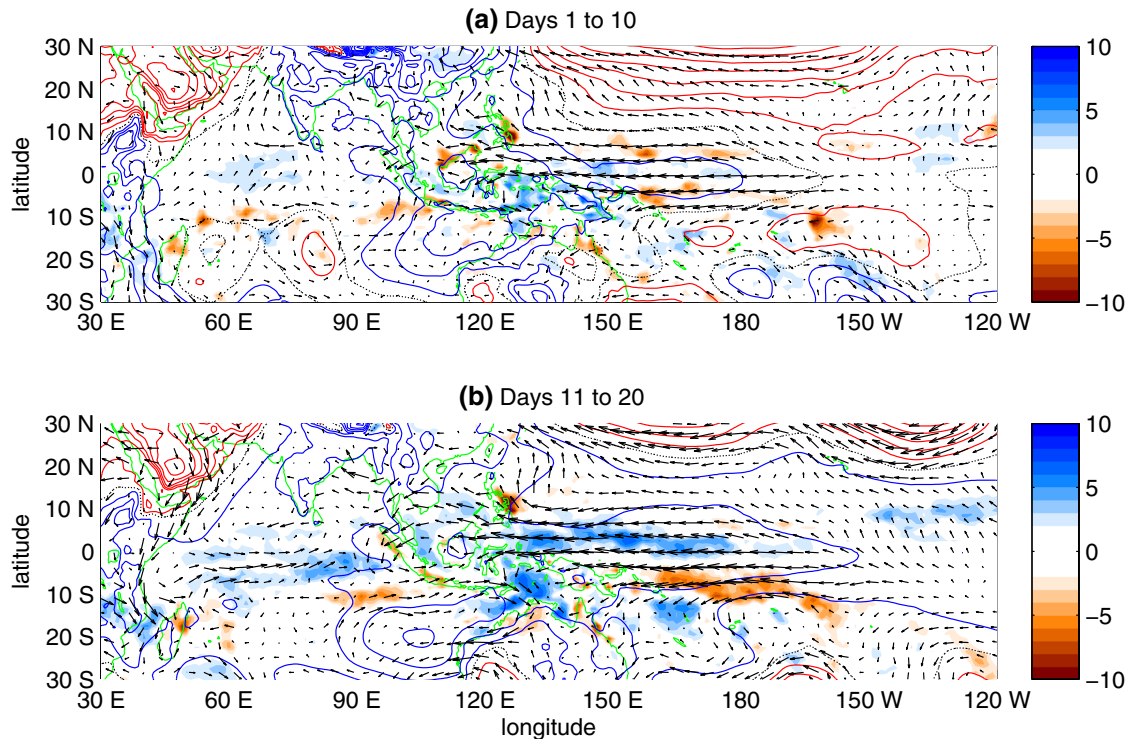
Fig. 10 Maps of biases in rainfall (*filled contours*), mean-sea-level pressure (*open contours*) and wind vector at 850 hPa (*arrows*), averaged over **a** the first and **b** the second 10 days of hindcasts from the uncoupled model. Rainfall biases are in mm d^{-1} and with respect to TRMM. Pressure biases are in hPa, with respect to ERA-Interim and depicted as *contours* with a 0.25 hPa interval—*red contours* indicate a positive (high) pressure bias; *blue contours* indicate a negative (low) pressure bias; a *dotted black contour* indicates the zero bias contour. Wind vector biases are also with respect to ERA-Interim. Biases here are averaged over years 1998–2009 for hindcasts starting on 1 February

a lesser extent—between 25 and 50% of that in hindcasts from a fully coupled version of the model. This difference between the bias evolution in the coupled and uncoupled hindcasts points to the development of biases in the SST being important for the ultimate development of the ITCZ bias in the coupled model.

The SST at the Equator in the western Pacific is found to cool erroneously. This cooling is not an extension of the classic cold tongue bias in the eastern Pacific but a patch of cooling that needs to be considered separately. We find biases in zonal wind stress in the area to be the cause of this spurious cooling, partly via biases in wind-driven latent heat flux. By comparing the developing biases in coupled and uncoupled hindcasts we further show that the easterly bias in wind stress is not affected by the presence of coupling to the ocean, implying that the source of the bias is in the atmosphere component of the model.

Analysis of three-dimensional atmospheric wind fields reveals an equatorial easterly wind bias that develops in the boundary layer above the region of wind stress bias. In the hindcasts initialised in February, the low-level easterly and upper-level westerly wind biases that develop begin to resemble an enhanced Pacific branch of the Walker

circulation after 10 days; in August, these biases extend across both the Indian and Pacific Oceans. The rate at which these wind biases develop and the baroclinic structure at day 10 is suggestive of errors in the convective heating being the cause of this wind bias, but the absence of this baroclinic structure in the early initial development raises doubts. Analysis of the wind fields alongside mean-sea-level pressure and rainfall fields does not reveal any large biases in rainfall over the Maritime Continent region that could be linked to these wind stress biases in the first 10 days. In the August hindcasts, there are large biases in the rainfall of the Asian summer monsoon, and the easterly flow in the equatorial western Pacific does extend off the Equator into this heating region. From this analysis it has not been possible to identify conclusively the initial cause of these wind stress biases or to rule out errors in the convective heating.

The fast rate at which these wind stress biases develop means that short, targeted sensitivity experiments using the uncoupled model with appropriate corrections or nudging should be able to shed some further light on the bias development. For example, simulations in which the lower tropospheric winds are constrained to observations,
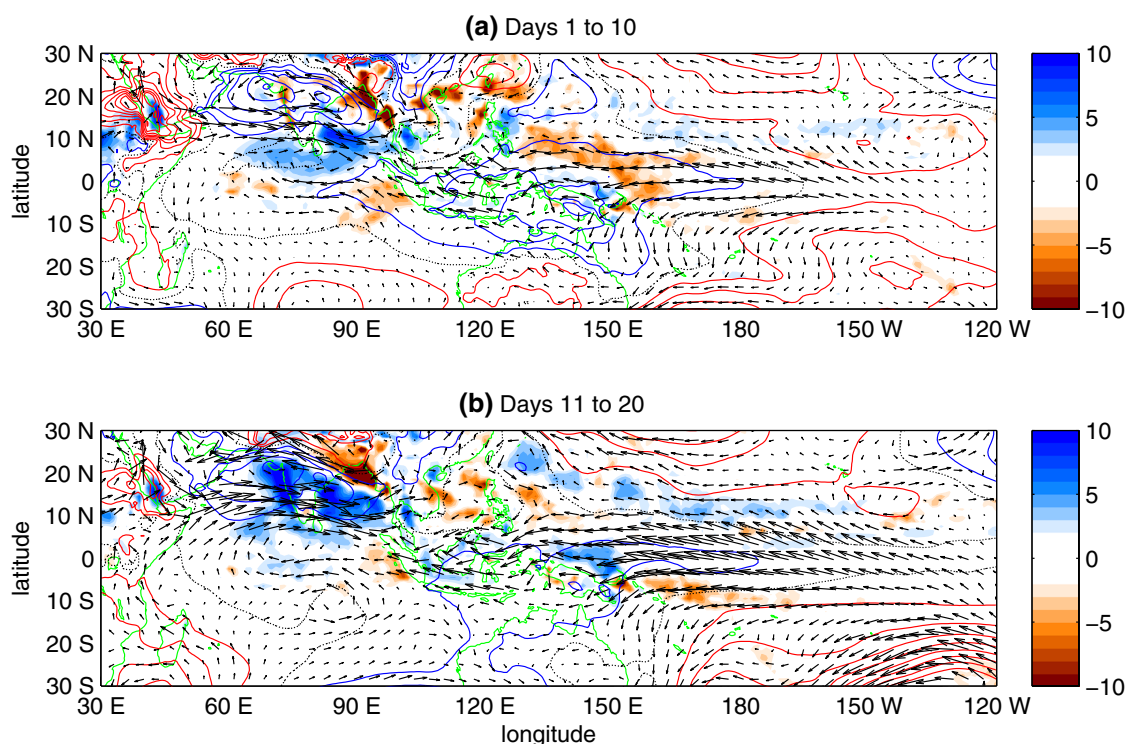
(a) Days 1 to 10

(b) Days 11 to 20

Fig. 11 Maps of biases as in Fig. 10, but for hindcasts starting on 1 August

but surface fluxes and SST are allowed to vary, could yield further clues. In particular, examining the development of the rainfall biases with and without this constraint could provide information on the role of convective heating. Alternatively, analysing the increments from the model data assimilation scheme may provide further information as to the cause of the wind biases. Some of the unanswered questions in this study could be answered by further simulations: for example, the use of a set of ocean-only simulations (similar to those suggested by Vannière et al. 2014) could provide more detailed information about the behaviour of the ocean, allowing us to determine the heat budget of the C box in more detail.

References

AchutaRao K, Sperber K (2006) ENSO simulations in coupled ocean—atmosphere models: are the current models better? Clim Dyn 27:1–16

Balmaseda MA, Mogensen K, Weaver AT (2013) Evaluation of the ECMWF ocean reanalysis system ORAS4. Quarterly J R Meteorol Soc 139:1132–1161

Barnston AG, Tippett MK, L'Heureux ML, Li S, DeWitt DG (2012) Skill of real-time seasonal ENSO model predictions during 2002–11: is our capability increasing? Bull Am Meteorol Soc 93:631–651

Bellenger H, Guilyardi E, Leloup J, Lengaigne M and Vialard J (2014). ENSO representation in climate models: from CMIP3 to CMIP5. Clim Dyn 42:1999–2018

Bernie DJ, Guilyardi E, Madec G, Slingo JM, Woolnough SJ, Cole J (2008) Impact of resolving the diurnal cycle in an ocean–atmosphere GCM. Part 2: a diurnally coupled CGCM. Clim Dyn 31:909–925

Boyle J, Klein S, Zhang G, Xie S, Wei X (2008) Climate model forecast experiments for TOGA COARE. Mon Weather Rev 136:808–832

Cane MA, Zebiak SE, Dolan SC (1986) Experimental forecasts of El Niño. Nature 321:827–832

Chikira M (2010) A cumulus parameterisation with state-dependent entrainment rate. Part II: impact on climatology in a general circulation model. J Atmos Sci 67:2194–2211

Dai F, Yu R, Zhang X, Yu Y, Li J (2005) Impacts of an improved low-level cloud scheme on the Eastern Pacific ITCZ–cold tongue complex. Adv Atmos Phys 22:559–574

De Szoeke SP, Xie SP (2008) The tropical Eastern Pacific seasonal cycle: assessment of errors and mechanisms in IPCC AR4 coupled ocean—atmosphere general circulation models. J Clim 21:2573–2590

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorolog Soc 137:553–597

Flato G, Marotzke J, Abiodun B, Braconnot P, Chou SC, Collins W, Cox P, Driouech F, Emori S, Eyring V, Forest C, Gleckler P, Guilyardi E, Jakob C, Kattsov V, Reason C, Rummukainen M (2013) Evaluation of climate models. In: Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR5). Cambridge University, Cambridge

Gill AE (1980) Some simple solutions for heat-induced tropical circulation. Q J R Meteorolog Soc 106:447–462

Guilyardi E (2006) El Niño–mean state–seasonal cycle interactions in a multi-model ensemble. Clim Dyn 26:329–348

Hirota N, Takayabu YN, Watanabe M, Kimoto M (2011) Precipitation reproducability over tropical oceans and its relationship to the double ITCZ problem in CMIP3 and MIROC5 climate models. J Clim 24:4859–4873

Huang B, Hu ZZ, Jha B (2007) Evolution of model systematic errors in the tropical Atlantic Basin from coupled climate hindcasts. Clim Dyn 28:668–682

Hwang YT, Frierson DMM (2013) Link between the double-inter-tropical-convergence-zone problem and cloud biases over the Southern Ocean. Proc Natl Acad Sci USA 110:4935–4940

Kummerow C, Simpson J, Thiele O, Barnes W, Chang ATC, Stocker E, Adler RF, Hou A et al (2000) The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. J Appl Meteorol 39:1965–1982

Latif M, Sperber K, Arblaster J, Braconnot P, Chen D, Colman A, Cubasch U, Cooper C et al (2001) ENSIP: the El Niño Simulation Intercomparison Project. Clim Dyn 18:255–276

Li G, Xie SP (2014) Tropical biases in CMIP5 multimodel ensemble: the excessive equatorial Pacific cold tongue and double ITCZ problems. J Clim 27:1765–1780

Lin JL (2007). The double-ITCZ problem in IPCC AR4 coupled GCMs: ocean–atmosphere feedback analysis. J Clim 20:4497–4525

MacLachlan C, Arribas A, Peterson KA, Maidens A, Fereday D, Scaife AA, Gordon M, Vellinga M et al (2014) Global Seasonal Forecast System Version 5 (GloSea5): a high-resolution seasonal forecast system. Q J R Meteorol Soc 141:1072–1084

Madec G (2008) NEMO Reference Manual, Ocean Dynamics Component: NEMO-OPA. Note du Pole de Modelisation 27, Institut Pierre-Simon Laplace

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

Martin GM, Milton SF, Senior CA, Brooks ME, Ineson S (2010) Analysis and reduction of systematic errors through a seamless approach to modelling weather and climate. J Clim 23:5933–5957

Mechoso CR, Robertson AW, Barth N, Davey MK, Delecluse P, Gent PR, Ineson S, Kirtman B et al (1995) The seasonal cycle over the tropical Pacific in coupled ocean—atmosphere general circulation models. Mon Weather Rev 123:2825–2838

Mechoso CR, Losada T, Koseki S, Mohino-Harris E, Keenlyside N, Castaño-Tierno A, Myers TA, Rodriguez-Fonseca B, Toniazzo T (2016) Can reducing the incoming energy flux over the southern ocean in a CGCM improve its simulation of tropical climate? Geophys Res Lett 43:11057–11063

Misra V, Marx L, Brunke M, Zeng X (2008) The equatorial Pacific cold tongue bias in a coupled climate model. J Clim 21:5852–5869

Molteni F, Stockdale T, Balmaseda M, Balsamo G, Buizza R, Ferranti L, Magnusson L, Mogensen K, Palmer T and Vitart F (2011). The new ECMWF seasonal forecast system (System 4). ECMWF Technical Memorandum #656

Neelin JD, Latif M, Allaart MAF, Cane MA, Cubasch U, Gates WL, Gent PR, Ghil M et al (1992) Tropical air-sea interaction in general circulation models. Clim Dyn 7:73–104

Phillips TJ, Potter GL, Williamson DL, Cederwall RT, Boyle JS, Fiorino M, Hnilo JJ, Olson G, Xie S and Yio JJ (2004). Evaluating parameterisations in general circulation models: climate simulation meets weather prediction. Bull Am Meteorol Soc 85:1903–1915

Praveen Kumar B, Vialard J, Lengaigne M, Murty VSN, McPhaden MJ (2012) TropFlux: air-sea fluxes for the global tropical oceans—description and evaluation. Clim Dyn 38:1521–1543

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved in-situ and satellite SST analysis for climate. J Clim 15:1609–1625

Reynolds RW, Smith TM, Liu C, Chelton DB, Casey KS and Schlax, MG (2007). Daily high-resolution-blended analyses for sea-surface temperature. J Clim 20:5473–5496

Saha S, Moorthi S, Wu X, Wang J, Nadiga S, Tripp P, Behringer D, Hou YT et al (2014) The NCEP climate forecast system version 2. J Clim 27:2185–2208

Smith DM, Scaife AA, Kirtman BP (2012). What is the current state of scientific knowledge with regard to seasonal and decadal forecasting? Environ Res Lett 7:015602. doi:10.1088/1748-9326/7/1/015602

Song X, Zhang G (2009) Convection parameterisation, tropical Pacific double ITCZ, and upper-ocean biases in the NCAR CCSM3. Part I: climatology and atmospheric feedback. J Clim 22:4299–4315

Stockdale TN (1997) Coupled ocean–atmosphere forecasts in the presence of climate drift. Mon Weather Rev 125:809–818

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93:485–498

Toniazzo T, Woolnough SJ (2014) Development of warm SST errors in the Southern Tropical Atlantic in CMIP5 decadal hindcasts. Clim Dyn 43:2889–2913

Vannière B, Guilyardi E, Madec G, Doblas-Reyes FJ, Woolnough S (2012) Using seasonal hindcasts to understand the origin of the equatorial cold tongue bias in CGCMs and its impact on ENSO. Clim Dyn 40:963–981

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

Vialard J, Menkes C, Boulanger J, Delecluse P, Guilyardi E, McPhaden MJ, Madec G (2001) A model study of oceanic mechanisms affecting equatorial Pacific sea surface temperature during the 1997–1998 El Niño. J Phys Oceanogr 31:1649–1675

Weisheimer A, Doblas-Reyes FJ, Palmer TN, Alessandri A, Arribas A, Déqué M, Keenlyside N, MacVean M, Navarra A, Rogel P (2009) ENSEMBLES: a new multi-model ensemble for seasonal-to-annual predictions—skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. Geophys Res Lett. doi: 10.1029/2009GL040896

Williams KD, Brooks ME (2008) Initial tendencies of cloud regimes in the Met Office Unified Model. J Clim 21:833–840

Williams KD, Bodas-Salcedo A, Déqué M, Fermepin S, Medeiros B, Watanabe M, Jakob C, Klein SA, Senior CA, Williamson DL (2013) The Transpose-AMIP II experiment and its application to the understanding of southern ocean cloud biases in climate models. J Clim 26:3258–3274

Xie P, Arkin PA (1997) Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. Bull Am Meteorol Soc 78:2539–2558

Yu L, Weller RA (2007) Objectively analysed air-sea heat fluxes for the global ice-free oceans (1981–2005). Bull Am Meteorol Soc 88:527–539

Zhang Y, Rossow WB, Lacis AA, Oinas V, Mishchenko MI (2004) Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: refinements of the radiative transfer model and the input data. J Geophys Res. doi:10.1029/2003JD004457

Zhang X, Liu H, Zhang M (2013) Double ITCZ in coupled ocean–atmosphere models: from CMIP3 to CMIP5. Geophys Res Lett 42:8651–8659

Zheng Y, Lin JL, Shinoda T (2012) The equatorial Pacific cold tongue simulated by IPCC AR4 coupled GCMs: upper ocean heat budget and feedback analysis. J Geophys Res. doi:10.1029/2011JC007746