Understanding the development of systematic errors in the Asian Summer Monsoon

Gill M. Martin[1], Richard C. Levine[1], José M. Rodriguez[1] and Michael Vellinga[1]

[1]Met Office, Exeter, Devon, EX1 3PB, UK

5 *Correspondence to*: Gill M Martin (gill.martin@metoffice.gov.uk)

Abstract. Despite the importance of monsoon rainfall to over half of the world's population, many of the current generation of climate models struggle to capture some of the major features of the various monsoon systems. Studies of the development of errors in several tropical regions have shown that they start to develop very quickly, within the first few days of a model simulation, and can then persist to climate timescales. Understanding the sources of such errors requires the

10 combination of various modelling techniques and sensitivity experiments of varying complexity. Here, we demonstrate how such analysis can shed light on the way in which monsoon errors develop, their local and remote drivers and feedbacks. We make use of the seamless modelling approach adopted by the Met Office, whereby different applications of the Met Office Unified Model (MetUM) use essentially the same model configuration (dynamical core and physical parametrisations) across a range of spatial and temporal scales. Using the Asian Summer Monsoon as an example, we show that error patterns

15 in circulation and rainfall over the East Asia Summer Monsoon (EASM) region in the MetUM are similar between multi-decadal climate simulations and seasonal hindcasts initialised in spring. Analysis of the development of these errors on both short-range and seasonal timescales following model initialisation suggests that both the Maritime Continent and the oceans around the Philippines play a role in the development of EASM errors, with the Indian summer monsoon region providing an additional contribution. Regional modelling with various lateral boundary locations helps to separate local and remote

20 contributions to the errors, while regional relaxation experiments shed light on the influence of errors developing within particular areas on the region as a whole.

Copyright statement

1 Introduction

30 Despite many advances in weather and climate modelling over the past decades, systematic errors remain prevalent in key regions such as the Asian Summer Monsoon. Such systematic errors have been shown in past studies to develop rapidly, often within the first few days of simulation, and can persist to climate timescales. This has important implications for forecasting on a wide range of timescales, and for climate projections, in regions where millions of people rely on the seasonal rainfall for their water resources and livelihoods. Several modelling studies have investigated the initial error

35 growth using short-range forecasts from numerical weather prediction models (e.g. Keane et al., 2019; Martin et al., 2010; Rodwell and Palmer, 2007; Phillips et al, 2004). Such studies allow the immediate influence of atmosphere model physical parametrisations to be identified without the complex feedbacks from circulation errors which develop over longer timescales. This approach can be particularly useful where similar model configurations are used for both timescales (Martin et al., 2010; Hurrell et al., 2009).

40 The advent of coupled ocean-atmosphere numerical weather prediction offers further challenges in the development of additional systematic errors through feedbacks between atmosphere and ocean. For extended-range and seasonal predictions, tracking the development of systematic errors through the coupled atmosphere-ocean-cryosphere system and across timescales ranging from individual weather events through intra-seasonal to seasonal variations is also a challenge. Several previous studies have used initialised seasonal hindcasts to shed light on the origin of coupled model errors in tropical

45 regions (e.g. Lazar et al, (2005); Huang et al. (2007); Liu et al. (2012); Vannière et al. (2014); Siongco et al. (2020)). Lazar et al. (2005) demonstrated that both the atmosphere and ocean components of coupled models contribute to the development of errors, on different timescales and in different regions, and with the balance of atmosphere/ocean contribution being model-dependent. Vannière et al. (2013) used a multi-model seasonal hindcast dataset to identify the order in which errors appeared in the tropical Pacific, and Vannière et al. (2014) developed this into a systematic approach that allowed them to

50 identify a range of drivers and timescales for tropical Pacific SST errors in the IPSLCM5A-LR coupled model. Similarly, Siongco et al. (2020) identified different drivers for the fast-developing cold phase and slow-developing warm phase of the equatorial Pacific SST errors in the Community Earth System Model, version 1 (CESM1). Voldoire et al. (2019) used a multi-model ensemble of seasonal hindcasts made by climate models to confirm that easterly wind stress errors drive warm SST errors in the tropical Atlantic from the first month onwards. In a global study analysing daily to multi-annual timescales

55 in two different coupled seasonal prediction models, Hermanson et al. (2018) showed a range of SST drift evolution and timescales among different regions and different times of year, with some regions being affected by poor initialization.

On sub-seasonal to seasonal timescales, there will be contributions to systematic errors both from local processes and from remote teleconnections. Separating these contributions, and identifying their interaction, requires a range of bespoke modelling tools that constrain parts of the climate system while allowing others to develop freely. Examples include:

60 atmosphere-only, land-only or ocean-only model simulations where observed or modelled fields can be used to force one coupled model component at a time; replacing surface fluxes in a coupled model with daily observed or modelled fields;

regional climate modelling with a range of lateral boundary locations (Levine and Martin, 2018; Karmacharya et al., 2015); global or regional relaxation experiments (Klinker et al., 1990; Rodriguez et al., 2017, 2019); and "pacemaker" experiments (where a climate model is forced by observed sea surface temperature variations in a specific region, but allowed to evolve

65   freely everywhere else; e.g., Deser et al. 2017; Zhang et al., 2018; Amaya et al., 2019).

Rodriguez and Milton (2019) describe analysis of the spin-up of the errors over the Asian monsoon region in initialized 15-day atmosphere-only hindcasts. These showed the gradual emergence, over the 15 days, of the key systematic errors seen in the moisture transport/divergence from free-running simulations of the same model. Some of the errors are large even at day 1, supporting previous results from e.g. Keane et al (2019) that errors in local parameterised physics are the key drivers of

70   monsoon errors rather than remote forcing errors of the circulation. Rodriguez and Milton (2019) further investigated which errors were driven from the Maritime Continent (MC) region by using regional relaxation experiments where the winds and temperatures over the MC region were relaxed back to reanalyses. This revealed that deficiencies in tropical convection over the MC region start to contribute to errors in the Asian monsoon circulation within the first 15 days of the hindcasts. Levine and Martin (2018) used a regional climate model centred over India and forced by reanalyses at the lateral boundaries to

75   show that remote errors (particularly, excessive convection over the equatorial Indian Ocean and poor representation of precursor disturbances transmitted from the Western Pacific) contribute significantly to the poor simulation of monsoon low pressure systems in the Met Office model.

In the present study, we illustrate how a combination of many of the techniques outlined above can be used to analyse the development of monsoon errors, their local and remote drivers and feedbacks. We take advantage of the range of Met Office

80   model configurations covering timescales from days through seasons to decades. These share a common dynamical core and similar physical parametrisations as part of the Met Office's seamless approach to modelling weather and climate. We extend and develop the previous work by including analysis of the development of errors in medium-range coupled and atmosphere-only model hindcasts during the first 7-15 days, and in a coupled seasonal hindcast ensemble during the first few pentads following initialization, and by investigating the individual and interacting roles of various remote regions in the

85   development of errors in both atmosphere-only and coupled model configurations. While our study focusses on systematic errors in the Asian summer monsoon, similar methods could be applied to other monsoon, and non-monsoon, regions. Section 2 describes the data and methods used, while Sect. 3 documents the results of the various experiments. In Sect. 4 we discuss the results and their implications for targeted model development.

2 Data and methods

90   Free-running climate simulations using Met Office coupled atmosphere-ocean configuration Global Coupled version 2 (GC2.0; Williams et al, 2015), forced by present-day greenhouse gases and aerosols and covering several decades, are used initially in order to illustrate the model errors of interest to this study. The atmosphere component of GC2.0 is Met Office Unified Model (MetUM) Global Atmosphere 6.0 (GA6.0; Walters et al., 2017), which is coupled to the Joint UK Land

Environment Simulator (JULES; Best et al., 2011) and the NEMO (Nucleus for European Modelling of the Ocean;
95 Madec, 2008) ocean and CICE (Hunke and Lipscomb, 2004) sea-ice models. The model is configured at N216 resolution
(0.833° in latitude and 0.556° in longitude, which is approximately 80 km at the equator, in the horizontal) for the
atmosphere and the ORCA0.25 tripolar grid (0.25°) for the ocean. The vertical resolution is 85 levels for the atmosphere and
75 levels for the ocean. Comparison is made against ERA-interim (ERA-I; Dee et al., 2011) and ERA5 (Copernicus Climate
Change Service, 2017) reanalyses for winds, the Global Precipitation Climatology Project pentad dataset version 2.2 (GPCP
100 v2.2; Xie et al., 2003; Adler et al., 2003) and the Tropical Rainfall Measuring Mission 3B42 product, version 7-7A (TRMM;
Kummerow et al., 1998; Huffman et al., 2010; Huffman and Bolvin, 2013) for precipitation, and NOAA daily Optimum
Interpolation Sea Surface Temperature v2 (OISSTv2; Reynolds et al. 2007) and HadISST1.1 monthly sea surface
temperatures (Rayner et al., 2003).

In order to study the development of errors after initialisation, we make use of a hindcast ensemble from the GloSea5
105 operational long-range forecast system (MacLachlan et al., 2015; Williams et al., 2015). GloSea5 also uses the MetUM
GC2.0 configuration at the same horizontal and vertical resolution as in the free-running simulations described above. The
standard operational hindcast set includes seven members per start date, four start dates (1, 9, 17, 25) per month, and runs
from 1993-2016. A Stochastic Kinetic Energy Backscatter scheme (SKEB2; Bowler et al. 2009) is used to introduce small
grid-level perturbations throughout the integrations to create ensemble spread. The atmosphere and land components are
110 initialized from daily ERA-I reanalyses at 0.75° ×0.75° resolution, while the ocean and sea ice models are initialized from
the GloSea5 ocean and sea ice analysis using GloSea5 Global Ocean 3.0, which is driven by ERA-I and uses the
NEMOVAR data assimilation scheme (Blockley et al. 2014).

In order to separate the influence of local and remote sources of error, we make use of a regional climate model (RCM)
configuration based on GA7.0 (Walters et al., 2019), forced by observed sea surface temperatures (SSTs) from OISSTv2,
115 and run at 0.44° × 0.44° resolution, approximately 50km) which is similar to (but slightly higher than) that of the N216
global models. GA7.0 includes changes, from GA6.0, to both model physics and dynamics that are both incremental
developments and targeted improvements to address critical errors that included a persistent dry bias over the Indian
subcontinent. While some progress was made in GA7.0 towards reducing those errors, the overall pattern of ASM errors
investigated in the present study remains, justifying our use of this RCM configuration. For the RCM domains centred over
120 China, a rotated north pole is used at 61°N, 296.3°E. The RCM is constrained by 6-hourly ERA-I at the lateral boundaries,
but within the domain the model runs freely after initialisation and is therefore able to develop errors due to local processes
and feedbacks, despite the constraint from the boundaries. By adjusting the locations of the lateral boundaries and comparing
between RCM simulations, and against a corresponding 20-year global atmosphere-only model simulation using GA7.0, the
contribution to the systematic errors from different regions can be ascertained. This technique was applied by Karmacharya
125 et al. (2015) and Levine and Martin (2018) for understanding sources of error in intraseasonal variability and monsoon low
pressure systems in the South Asian Summer Monsoon.

We study the evolution of errors after initialisation in the GloSea5 season hindcast ensembles with different start dates and, in addition, initialised 7 to 15-day numerical weather prediction (NWP) hindcasts using atmosphere-only and coupled configurations of GA6.1[1]/GC2 at N768 resolution (0.234° x 0.156°, approximately 26 km at the equator). These are

130 initialised every day between 1st May and 19 September 2016 and each run for 15 days (Vellinga et al 2020). The day-1, day-2, etc hindcasts can be combined to provide a seasonal climatology for each lead time, and the use of the same atmosphere model configuration in the coupled and atmosphere-only hindcasts allows the role of coupling to be ascertained.

To shed light on the drivers of systematic errors, we make use of the "nudging" technique described by Rodriguez and Milton (2019). A 20-year, free-running, atmosphere-only, model simulation using GA7.0 at N96 horizontal resolution (1.25°

135 in latitude and 1.88° in longitude, approximately 200 km at the equator) is relaxed back to analyses over regions from where we consider significant systematic errors may originate and affect other regions through remote teleconnection. Model winds and potential temperatures are nudged back to ERA-I with a 6-hourly relaxation time scale at all model levels. A 10° buffer zone around the relaxation subdomain is applied in which the nudging increments are exponentially damped to zero, in order to ensure a smooth transition between the nudged and free-running parts of the simulation. Similar nudging experiments are

140 also carried out in 15-day hindcasts initialised once per day through JJA of 2016 using the NWP GA6.1 atmosphere-only configuration at N216 resolution (NWP-N216), in order to examine how the influence from the nudged region is manifest in the development of the errors.

3 Results

3.1 Climatological errors in the Asian summer monsoon

145 Figure 1a shows June to August (JJA) mean climatological errors in rainfall and 850 hPa winds in the 30-year, free-running, present-day, GC2 simulation compared with ERA-I and GPCPv2.2. Similar to previous studies of the Asian monsoon system in MetUM configurations (e.g. Keane et al., 2019; Johnson et al., 2016, 2017; Bush et al., 2014; Martin et al., 2010; Ringer et al., 2006) the model exhibits a deficit in rainfall over the Indian peninsula, the eastern Indian Ocean south of the Equator and the Maritime Continent, an excess over the Indian Ocean to the north of the Equator, in the eastern South China Sea

150 (SCS) and the western Pacific, and excess precipitation over the mountains bordering the Tibetan plateau. These are accompanied by a weak Somali jet that diverges into an anticyclonic anomaly over India, excessive westerly flow over southeast Asia, the SCS and across the Philippines into the western Pacific, and a cyclonic error and deficit in precipitation over southeastern China, southern Japan, Korea and the East China Sea.

Johnson et al. (2017) analysed the climatological June to August (JJA) seasonal mean errors in a hindcast ensemble from

155 GloSea5 and showed that they are similar to those seen both in climate models including the MetUM (Sperber et al., 2013) and in other state-of-the-art seasonal forecast systems. Figure 1b shows the JJA climatological errors from the current

---

[1] This operational configuration includes a small number of scientific differences from GA6.0; see Walters et al. (2017).

GloSea5 23-year operational hindcast ensemble initialised each year on the four start dates in April. This confirms once again that, despite the initialisation and the relatively short lead time, the hindcast JJA errors are very similar in pattern, and (with the exception of the Indian region) in magnitude, to those from the 30-year free-running simulation (Fig 1a).

160 Johnson et al. (2017) commented that the seasonal mean errors over the Indian region are largely due to a climatologically late onset of the monsoon in the model, which reduces the precipitation over and around India in May and June. Figure 1(c to h) shows errors in the June, July and August climatologies from GC2 and from the GloSea5 23-year operational hindcast ensemble (at ~1 month lead time for each month, i.e. using the 4 start dates in May, June and July respectively). This shows that, while the rainfall errors over the Indian region as a whole in both GC2 and GloSea5 are indeed largest in June, both the

165 magnitude of the monthly errors and the differences between the three months are noticeably smaller in GloSea5.

The differences between GloSea5 and GC2 in the Indian region particularly in June (GC2 shows a weakened Somali Jet and much larger rainfall deficit than GloSea5) are consistent with the differences between GloSea5 and CMIP5 models commented by Johnson et al. (2017), who considered these attributable in part to a smaller Arabian Sea cold SST error in GloSea5. Figure 2 shows the errors in SST against HadISST observations (1993 to 2015) in JJA and for June, July and

170 August for GC2 and GloSea5 as in Fig. 1. Cold errors in the Arabian Sea are seen in both simulations, particularly in June, but they are considerably larger in GC2. Marathayil et al. (2013) showed that, in CMIP3 models, such errors develop in winter due to anomalously strong north-easterly winter monsoon winds advecting cold, dry air from the Eurasian land mass over the Arabian Sea. Their analysis suggested that excessive rainfall in the equatorial Indian Ocean and anomalously cold winter continental surface temperatures in the CMIP3 models both contribute to the Arabian Sea cold SST error. Levine et

175 al. (2013) showed that these errors persist into spring and early summer and are associated with a weaker monsoon circulation and reduced monsoon precipitation. Initialisation of the GloSea5 hindcasts in spring prevents the growth of a large SST error, thereby reducing the circulation and rainfall errors over the Indian region (Levine and Turner 2012; Levine et al. 2013, personal communication R. Levine).

The free-running and initialised models are consistent in developing cold SST errors around the Maritime Continent, the

180 South China Sea and the central and eastern Indian Ocean, even just over a month after initialisation. An SST error dipole pattern resembling that of the positive Indian Ocean Dipole (IOD; Saji et al., 1999) is apparent in the seasonal hindcasts but is much stronger in the free-running simulation. This is consistent with the circulation anomaly pattern shown in Fig. 1 which strongly resembles the atmospheric component of the IOD teleconnection: south-easterly anomalies along the Sumatran coast and easterly anomalies along the Equator. Previous work (e.g. Marathayil et al., 2013; Johnson et al., 2017)

185 has shown that this SST error pattern is associated with a coupled interaction between excessive rainfall in the central equatorial Indian Ocean, excessive easterly low-level winds and increased upwelling that shoals the thermocline in the east. The additional northeasterly wind anomalies in the western Indian Ocean in GC2 exacerbate this error pattern. Johnson et al. (2017) showed that this coupled mean state error results in errors in the representation of the IOD as a mode of variability in the model, reducing its ability to predict the Indian monsoon circulation.

190 Rodriguez and Milton (2019) showed that local errors in moisture convergence/divergence over the Maritime Continent region also contribute to the development of circulation and rainfall errors in the eastern Indian Ocean, the South China Sea, western Pacific and southeast China in atmosphere-only simulations. It is likely that, in the coupled system, these atmosphere errors drive a cooling response in the SSTs which further contributes to decreases in rainfall and anomalous moisture divergence through coupled feedbacks.

195 The largest differences between the free-running and initialised simulations are seen in the central North Pacific, where the cold SST errors in the free-running GC2 simulation are much larger than in GloSea5. Such errors are common among CMIP5 models: Wang et al. (2018) showed that they are associated with overly strong surface winds driving excessive evaporation, combined (in summer) with a deficit in downward solar radiation at the surface. While Wang et al. (2018) showed that, in CMIP5 models the cold errors are present throughout the year (but largest during JJA), the initialisation of

200 GloSea5 in spring limits the extent to which these can develop during the summer months.

Despite the differences related to errors which develop in the winter in GC2, there are many areas where the similarity between the monthly error patterns at ~1-month lead time and the seasonal mean error pattern demonstrates that the errors develop quickly and then persist to longer timescales in this coupled model. In the following sub-sections we demonstrate how a range of configurations within the seamless modelling system can be used to shed light on various aspects and drivers

205 of these errors.

3.2 Regional climate modelling

To investigate first the local and remote sources of the errors identified in Sect. 3.1, we use regional climate model (RCM) simulations with different domain sizes, centred over China and forced at the lateral boundaries with ERA-I 6-hourly re-analyses, and using time-varying, observed SSTs instead of an ocean model. Such experiments isolate the effects of any

210 remote errors in an atmosphere-only global model (AGCM) that are located outside the RCM domain from those developing within the domain. The RCM simulations are performed at 0.44° x 0.44° resolution, similar to that of the N216 GA7.0 AGCM simulation (0.833° × 0.556°), so that the comparison between RCM and AGCM isolates the local and remote forcing of errors over these two regions during the EASM. Karmacharya et al. (2015) used this approach to investigate local and remote sources of MetUM errors in the Indian monsoon region. They showed that the equatorial Indian Ocean is a key driver

215 of Indian rainfall errors, although errors over the Himalayan foothills also played a role and there was evidence of locally-driven errors that were thought to be related to the model's inherent difficulties in reproducing the diurnal cycle of rainfall over land. Levine and Martin (2018) used similar methods to show that remote errors contribute significantly to the poor simulation of Indian monsoon lows and depressions.

The RCM domains used in the present study are shown in Fig. 3. Figure 4 shows the climatological errors in JJA from the

220 AGCM and the RCM China1 and China1SE (which includes the China1, China1E and China1S regions) domains. Although the magnitude of the error differs in places, the error pattern for JJA in the AGCM (top left) is very similar to that seen in the coupled simulations (Fig. 1). This suggests that neither the change between GA6 and GA7.0 nor the atmosphere-ocean

coupling has a major impact on the overall error pattern. The local RCM simulation, China1, favours southerly and south-westerly anomalies over southeast China, the South China Sea and the Philippine Sea, and a more widespread wet error than the AGCM. This includes much of the steep orography along the southern and south-eastern edges of the Tibetan Plateau, and most of southern China. The circulation anomalies over southeast China and the surrounding seas in China1 are opposite to those in the AGCM, suggesting that the characteristic error of a weakened southwesterly flow and rainfall deficit over southeast China and the East China Sea are not locally driven.

Extending the domain to the south and east (China1SE) contributes north-easterly anomalies and a dry error over the middle/lower Yangtze River Basin (Fig. 4 lower right panel), although this does not recover the full AGCM error (Fig. 4, lower left panel). Neither of the experiments using the China1E and China1S domains individually contribute northeasterly anomalies (Fig. 5 lower middle and centre right panels), indicating that the source of these anomalies is the south-eastern sector of this domain. However, the eastward extension in China1E does produce an anomalous low-level easterly wind component towards eastern China as part of an anomalous cyclonic flow over the western Pacific, representing a weakening and northward displacement of the Western North Pacific Subtropical High (WNPSH). There is associated drying over land and increased rainfall offshore. The southward extension in China1S contributes weakly both to the westerly anomalies across the South China Sea (not shown) and to the easterly anomalies over the middle/lower Yangtze River Basin, but both of these are strengthened when the domain is extended both to the south and the east, thereby including in addition the whole of northern Indonesia. Extending the China1 domain to the north and west has little impact compared with the anomalies contributed locally by China1 itself (Fig. 5), although China1W does contribute additional dry rainfall anomalies over southern China. However, preliminary experiments with much larger RCM domains (not shown) suggest a role for even more remote influence, perhaps through the circum-global teleconnection (Wu et al., 2016).

A limited number of RCM experiments was carried out in which the RCM was initialised each year on 1 May, in order to determine how quickly the influence of local processes and remote teleconnections became apparent. For all domains, the differences between the re-initialised and free-running experiments were minimal (not shown), indicating a rapid and robust evolution of the atmosphere model towards these systematic errors. The development of errors in the first few weeks after initialisation is explored further in the next section.

This analysis illustrates how an RCM with different lateral boundary locations can be used to shed light on the local and remote sources of systematic error in a climate model. For the EASM, we find that much of the circulation and rainfall error pattern seen in the full GCM is not driven locally but is related to errors arising mainly to the south and east of the region, i.e. over the Maritime Continent, South China Sea and the Western Pacific.

3.3 Development of errors in initialised hindcasts

Having identified that in much of the ASM region the errors appear to develop rapidly and persist thereafter to long timescales, we now make use of initialised hindcasts to examine their development and evolution during the first few weeks after initialisation. We first make use of the GloSea5 seasonal hindcast ensemble, which consists of 7 members per start date

for four start dates per month and covers a 23-year period from 1993-2015. In order to reduce the effects of internal variability, we average the ensemble mean precipitation, winds and SSTs into pentads and average both the model and observational fields over the hindcast period.

Figures 6 and 7 show the climatological development of rainfall, wind and SSTs errors in the seasonal hindcast ensemble,
260   pentad by pentad, following initialisation on 25th May. Anomalously warm SSTs and excess precipitation occur in the SCS and western Pacific soon after initialisation, with cold anomalies and deficient precipitation around the Maritime Continent. The circulation anomalies initially show divergence over the Maritime Continent and southerly anomalies into the SCS. As the hindcasts progress, the cold/dry anomalies around the Maritime Continent expand westwards and northwards and the southerly anomalies develop into westerly anomalies that form the southern flank of an anomalous cyclonic pattern over the
265   western Pacific which represents the weakening and displacement of the WNPSH. The westerly anomalies intensify as the South Asian summer monsoon onsets at the start of June and westerly winds across southeast Asia promote further rainfall, creating a positive feedback that develops a westerly wind error (extension of the westerly jet) across the SCS and the Philippines into the western Pacific. A twin cyclonic error develops over the SCS and western Pacific in early June, developing northwards and causing the northeasterly anomalies over southeast China that were highlighted in previous
270   sections. Positive rainfall errors also appear over the Indian Ocean to the south of the Indian peninsula in mid/late May, associated with anomalous northerly winds from the peninsula (and an increasing dry error there) converging with the anomalous easterly winds from the eastern equatorial Indian Ocean (EEIO). By mid-June, the patterns of rainfall and wind errors closely resemble the long-term June climatological errors in the free-running simulations (Fig. 1 and Fig. 4). This analysis confirms the suggestion from the RCM simulations, in Sect. 3.2, that the errors in the EASM over China are driven
275   largely by errors arising to the south and east of the region.

Figure 7 shows similar analysis of the SST errors. Warm anomalies develop immediately in the SCS and western Pacific, associated with the positive rainfall errors seen in Fig. 6. Warm anomalies also develop in the Bay of Bengal soon after initialization, while cold anomalies develop around the Indonesian islands. Over the first 15 days of the hindcast, the south-eastern Bay of Bengal warms while the cold anomalies around the Indonesian islands spread westwards and northwards,
280   temporarily creating a north-south dipole over the EEIO. However, in subsequent pentads the warm anomalies are replaced with colder SSTs, in association with increasing south-easterly anomalies along the Sumatran coast and diverging 850 hPa wind anomalies and a negative rainfall error across the whole of the Maritime Continent. Warm SST errors develop in the western equatorial Indian Ocean by mid-June, creating the dipole error pattern seen in the long-term climatological errors.

We find that the error patterns develop in a similar way when using start dates of 25th June, 25th July and 25th August (not
285   shown). This indicates that they are a robust feature of the model's behaviour during the monsoon season, consistent with their similarity to those in the free-running coupled simulation. However, similar analysis using start dates in late April and early May shows a slightly different development in the first ~15 days of the hindcast (see Fig.s 8 and 9): the anomalous divergence and rainfall deficit over the Indonesian islands is much more localised and takes longer to spread westwards and northwards. There is greater and more widespread warming of the southern Bay of Bengal, while the cold anomalies south of

the Equator off the Sumatran coast do not start to develop until around 20th May. This is thought to be related to the seasonal transition that takes place around mid-May and marks the start of the Asian monsoon season (Wang et al. (2004); Figure 10). Prior to this transition, the mean state low-level winds over the equatorial Indian Ocean are westerly and the mean flow over the Indonesian islands is weak. As noted by Ding and Chan (2005), the onset of the South China Sea summer monsoon is very abrupt, with a rapid switch from easterlies to westerlies over the South China Sea and a rapid expansion north-eastwards of the south-westerlies from the EEIO across the Indochina Peninsula. In hindcast ensembles initialised after this transition (Fig. 6), when the easterly low-level flow over the EEIO south of the Equator is stronger, there is a more widespread anomalous divergence over the Maritime Continent and more rapid cooling of the SSTs to the west of Sumatra. This analysis illustrates that the monsoon error development in initialised hindcasts is dependent on the stage of the monsoon season, as well as on the lead time of the hindcast. Once the broadscale seasonal transition has occurred, the error patterns develop in a similar manner regardless of the initialisation date.

Figure 11 shows pentad rainfall, winds and SSTs averaged over various different regions, from hindcast ensembles initialised on different start dates between 9th April (0409) and 25th August (0825), along with similar timeseries for GPCP and TRMM rainfall observations, ERA-I winds and OISSTv2 SSTs. For start dates in April, the SST in the SCS initially warms excessively, before cooling systematically into a cold error through the JJA season (Fig. 11a). For start dates in late May onward, the SST appears to be initialised systematically warmer than the observations but to cool thereafter. The peak warm SST error coincides with the broadscale seasonal transition that is heralded by the SCSSM onset, as determined by the reversal of the 850 hPa winds over the SCS (Fig. 11b) in the criterion suggested by Wang et al. (2004). The SST cooling after this transition coincides with an acceleration of the westerly winds into a positive error for all start dates. In response both to this and to the additional convergence of moisture into the SCS from the Maritime Continent, the rainfall over the "Philippines" region (Fig. 11c) starts with a positive error and increases thereafter, particularly in the hindcasts initialised in May and June. The East Asian monsoon Index (EASMI: see Wang et al., 2008) decreases rapidly after initialisation in all hindcasts (Fig 11d), indicating the weakening and displacement of the WNPSH. Separation of this index into its two components reveals that this is driven mainly by the increasingly excessive westerly flow in the southernmost box, which extends from southeast Asia across the SCS and the Philippines into the western Pacific, including the SCS box in Fig. 11b, and largely coinciding with the "Philippines" region in which the rainfall also increases (Fig. 11c). However, the hindcasts also rapidly develop an easterly error in the northernmost box, which extends from southern China across the East China Sea and to the south of Japan.

As discussed above, the SSTs in the EEIO to the south of the Equator (Fig. 11f) cool systematically for all start dates. In contrast, the SSTs to the north of the Equator (Fig. 11e) warm substantially over the first few pentads for the April and early May start dates before cooling and ultimately developing a cold error, while for later start dates there is only a short period (2 or 3 pentads) of initial slight warming before a similar cooling begins and persists for the rest of the season. Examination of the 850 hPa winds in Fig.s 8 and 10 shows that this initial warming for the late April start date is associated with a developing rainfall deficit and northeasterly anomalies from southeast Asia, opposing the mean flow. After pentad 28 (when

the broadscale seasonal transition occurs), this is replaced by westerly/southwesterly anomalies (accelerating the mean flow)

325 and a developing positive rainfall error. For start dates after this seasonal transition (Fig. 6) the wind anomalies are persistently westerly/southwesterly with an increasing positive rainfall error.

We examine this behaviour in more detail in a 2016 case study using hindcasts from the Met Office's NWP model in both atmosphere-only and coupled configurations. The atmosphere-only runs are 7-day operational forecasts while the coupled model hindcasts were run for 15 days. In both cases, there was one ensemble member per day run in near-real time since 1

330 May 2016. Results shown in Sect. 3.2 indicated that the errors developing in atmosphere-only configurations closely resemble those in the coupled atmosphere-ocean models. However, SST errors are also identified, so it is important to understand the extent to which these are driven by, and feed back upon, the atmospheric errors.

As discussed above, the development of SST errors in the EEIO over the first ~15 days of hindcast differs according to whether the hindcasts are initialised before or after the broadscale seasonal transition. This is further illustrated by composite

335 analysis of SST and 10m winds at forecast lead times of 1, 5 and 15 days of the coupled NWP-N768 hindcasts, over a period of 10 to 15 days on either side the broadscale seasonal transition (Fig. 12). For 2016, the validity dates chosen are 10 to 19 May ("before") and 10 to 23 July ("after"). Figure 12(a-c) shows the emergence of SST and surface wind errors in the EEIO before the transition. At day 1 the biases are small, showing in part the discrepancies between the OISSTv2 SSTs and the analysis SSTs (FOAM, Waters et al., 2014) used to initialise the hindcasts. At longer lead times, a large warm bias develops

340 in the EEIO and the Maritime Continent, which is associated with a weakening of the equatorial westerly flow and the southerly wind in the Bay of Bengal. The emergence of errors after the transition (Fig. 12d-f) show a different pattern. A cool bias starts to develop in the Maritime Continent and the adjacent ocean in the Bay of Bengal. The error is associated with a zonal wind anomaly extending from the Bay of Bengal to the tropical Western Pacific and southeasterly anomalies along the Sumatran coast and in the central equatorial Indian Ocean. On the other hand, a warm bias develops in the

345 equatorial western Indian Ocean, connected with a weakening of the surface wind in the region. These results are consistent with the analysis of GloSea5 (Fig. 5), despite the greater atmospheric horizontal resolution used in the NWP-N768 hindcasts, confirming that the error patterns emerging in the first 15 days, both before and after the broadscale transition, are robust.

The change in evolution of the SST errors in the northern EEIO box (as used in Fig. 11e) over the first 15 days of the coupled NWP-N768 hindcasts initialised between May and early August (Fig. 13a) is also similar to that seen in GloSea5. In

350 forecasts initialised in May, SST in the northern EEIO box develops a warm error of around 0.5°C relative to the ocean analyses used to initialise the coupled forecasts (FOAM, Waters et al., 2014). This warm error manifests itself as a tendency to under-predict the cooling of SSTs in the second half of May. Forecasts initialised in June and July do not have this problem and develop a much weaker warm SST error within the first 15 days, mostly around 0.1°C, again consistent with the results for GloSea5. In fact, SSTs follow the observed cooling and levelling-off during June and July reasonably well. The

355 warming of SSTs relative to the ocean analyses during the second half of May stems, at least partly, from under-representing the cooling that is seen in ERA5. That cooling is related to increased surface heat loss during that period (Fig. 13d). However, in CPLDNWP, excessive downward solar radiation (not shown) and under-estimated turbulent (i.e. latent and

sensible) heat fluxes (e.g. Fig. 13e) contribute to a reduction in the net surface flux out of the ocean during this period. The error in turbulent fluxes can be partly traced to a weak surface wind error (Fig. 13c). Errors in ocean processes likely also

360 contribute to SST errors. These may be surface-driven (due to the weak surface wind bias, Fig. 13c) or caused by deficiencies in ocean processes (e.g. vertical mixing). Shallow errors in ocean mixed layer depth would exacerbate warming of SST caused by surface flux errors. Figure 13b confirms a lack of deepening of the mixed layer in the model early in the period, consistent with the weak-wind bias during that period. In future work we will examine the contribution from ocean processes in more detail.

365 By comparing surface heat flux errors from coupled and atmosphere-only forecasts in this period we can determine the importance of air-sea coupling in the development of surface flux errors (Fig. 13d,e). For most of the time, the evolution of surface flux errors is very similar between coupled and uncoupled configurations. This suggests that coupled feedbacks are of limited importance here in the development of surface flux errors. The main exception is during the second half of May, when the strongest warm SST error develops. In this period, the differences in surface latent heat flux error between

370 CPLDNWP and UNCPLD are unusually large, differing by 50-100 W m^{-2} (Fig. 13e). Coupled feedbacks cause reduced latent heat loss in CPLDNWP, compared to ERA5 (positive values in Fig. 13e), while UNCPLD shows excessive cooling from surface latent heat flux (negative values in Fig. 13e), consistent with a positive 10m wind bias in UNCPLD (Fig. 13c). Further work is needed to clarify how this coupled feedback operates. This example illustrates how coupled and uncoupled initialised forecasts can be used to home in on some of the long-standing errors seen the Indian Ocean.

3.4 Regional nudging experiments to assess sources of error

375 From the analysis shown in Sect. 3.2 and 3.3, we hypothesise that the reduced rainfall and anomalous outflow from the Maritime Continent and Indian regions play a role in the development of the circulation errors in the ASM at the start of the monsoon season. In order to test this hypothesis, we conduct a series of atmosphere-only sensitivity experiments using the nudging/relaxation methodology described in Rodriguez et al. (2017). This involves relaxing the temperatures and winds

380 back to analyses with a 6-hourly relaxation time scale at all model levels. Assuming a linear response, the difference between the Control and the "Nudged" simulations then gives an indication of the role played by the nudged region in the errors that occur in the Control in other locations (Klinker 1990).

3.4.1 Free-running simulations

We apply this methodology first to climate simulations, using the atmosphere-only configuration GA7. We use four different

385 nudging regions (see Fig. 14), referred to as the "Philippines", "Indonesia", "South Asian Summer Monsoon" (SASM) and "Maritime Continent" (MC) regions. For these experiments, the horizontal winds and potential temperature at all model levels are relaxed back to ERA-Interim reanalyses and the simulations are run for around 20 years, from 1/9/1988 to 1/1/2009.

Figure 15(a,b) shows the climatological differences in 850 hPa winds and precipitation between the Control and Nudged
390   experiments during JJA, for the "Philippines" and "Indonesia" regions. These results suggest that the "Indonesia" region
promotes westerly anomalies extending from the South Asian monsoon westerly jet across the Philippines into the western
Pacific, while the "Philippines" region promotes additional acceleration of these westerly winds as part of an anomalous
cyclonic circulation that includes north-easterly anomalies over southern China. Both regions promote excess rainfall over
the eastern SCS and the western Pacific. Figure 15(c,d) shows the results for the SASM and MC regions. These suggest that
395   errors arising locally over the SASM region are directly responsible for the anticyclonic anomaly and deficit in rainfall over
India and for much of the error pattern in rainfall over the equatorial Indian Ocean. The SASM region also promotes the
acceleration of the westerly winds across the SCS into the western Pacific and the positive error in rainfall in those regions.
The Maritime Continent region as a whole promotes acceleration of the westerly winds and increased rainfall across the SCS
and the western Pacific, and an anticyclonic anomaly that represents weakening and eastward displacement of the WNPSH
400   region. The influence of the Maritime Continent region, and particularly the Indonesian islands, in promoting the
southeasterly(easterly) wind anomalies in the eastern(central) Indian Ocean, as suggested in Sect. 3.2 and 3.3, is also
confirmed by these results.

This analysis suggests that there are both local and remote contributors to the ASM errors seen in the MetUM model
simulations. The experiments indicate that Indonesia and the oceans around the Philippines play a separate, but interacting,
405   role in the development of these errors during the seasonal transition towards the Asian summer monsoon. The SASM region
helps to reinforce those errors while also developing the majority of its circulation and rainfall errors locally.

3.4.2 Initialised simulations

The "nudging" methodology can also be applied in initialised simulations and used to track the influence of a particular
region on the development of errors elsewhere. We show here, as an example of this methodology, the influence of the
410   "Philippines" (PHL) region (used in Sect. 3.4.1 and displayed in Fig. 14) on the growth of remotely forced model systematic
errors over China, the western Pacific and the Maritime Continent, over the first 15-days of NWP-N216 atmosphere-only
simulations conducted during June–August 2016 (Fig. 16). Consistent with the analysis of GloSea5 coupled model hindcasts
in Fig. 5, the total mean error (forecast minus analysis) in the surface wind for forecast days 1, 5, and 15 (see Fig. 16(a–c))
shows the gradual emergence of the systematic errors. This includes erroneous equatorial easterlies west of Sumatra,
415   extending to 80°E, and a large error in the western Pacific, east of the Philippines that extends north to the sub-tropics in an
erroneous cyclonic pattern that reflects the weakening of the WNPSH. Other surface-wind errors are shown in the Maritime
Continent, the Bay of Bengal and the western equatorial Indian Ocean off the African coast.

On day 1, the contribution of the PHL to the total error is very small, mostly confined to the PHL region as expected, but by
day 5 of the NWP forecasts the PHL errors are responsible for forcing mean errors beyond the nudged region, such as the
420   erroneous cyclonic wind in the Western Pacific subtropics, as well as errors in the Maritime Continent. These errors are
consolidated by day 15 of the forecast (Fig. 16 (d–f)). For completeness, we also show the contribution to the total error

from the areas outside the PHL nudging domain (Fig. 16(g–i)). A smaller area of erroneous cyclonic circulation in the Pacific occurs just south of Japan by day 5, that indicates that the systematic error in the WNPSH also has extra-tropical origins. Other wind errors not forced by the PHL region include the easterlies west of Sumatra and errors in the Bay of

425   Bengal and the Western equatorial Indian Ocean off the African coast.

Similar experiments have been carried out with the other regions identified in Sect. 3.4.1. These also confirm that the local and remote contributions from those regions to the circulation errors in the ASM emerge in the first 5-15 days of the forecasts. These experiments illustrate the important role played by certain regions in the development, from an early stage, of systematic errors in the ASM. Future work will include applying the nudging technique to GloSea5 hindcasts in order to

430   trace the influence of specific regions on the development of errors on seasonal timescales. Identifying such key regions provides a focus for future process analysis, model development and evaluation which may ultimately improve the model forecasts for the ASM as a whole.

4 Summary and Conclusions

In this study we have demonstrated the use of a range of modelling tools and techniques aimed at understanding the sources

435   of error in monsoon regions, using the specific example of the ASM errors in the MetUM model. The tools and techniques allow close examination of the error development after initialization, the separation of the roles of local processes and remote teleconnections, the identification of the contribution from errors developing in particular regions to the ASM error as a whole, and understanding of the role of atmosphere-ocean coupling.

Our analysis suggests that there are both local and remote contributors to the EASM errors seen in the MetUM model

440   simulations. The experiments indicate that Indonesia and the oceans around the Philippines play a separate, but interacting, role in the development of these errors during the EASM season, while SASM region (in which errors are mainly driven locally) helps to reinforce those errors. Although many of the same systematic error patterns have been found in atmosphere-only simulations (e.g. Rodriguez and Milton, 2019), SST errors also contribute, both at initialisation and through their development in a coupled response to the circulation and rainfall errors.

445   This study illustrates how this methodology can be used to identify the regions and model components responsible for the development of these long-standing monsoon errors. While further analysis is needed to investigate the processes involved and how they are mis-represented in the models, we have narrowed down the regions responsible (mainly the Maritime Continent and Philippines regions) which will allow us to target future detailed investigations. We have also demonstrated that it is mainly the atmospheric component that is responsible, with the ocean providing a coupled feedback which mainly

450   exacerbates the errors. We have also shown that the development of the errors in the first few weeks depends on when the hindcasts are initialised in relation to the broadscale monsoon transition that typically occurs in mid-May. Finally, we find that these systematic errors and their development are largely insensitive to changes in horizontal resolution.

This analysis methodology benefits from the use of a seamless modelling system, where different configurations of a modelling system that are used for forecasting on different timescales share very similar physical and dynamical formulations. This allows the development of systematic errors to be studied on a range of timescales, and the roles of resolution and ocean-atmosphere coupling to be studied, without the complication of different physical parameterizations or dynamical cores that other multi-model studies might include. This approach also allows the whole suite of models to benefit from improvements that ultimately result from better understanding of the errors and informed, targeted, model development.

Our study highlights a number of different techniques that can be employed to investigate the sources of model error in a particular region. Once these are known, further work can be done to explore the local processes contributing to this behaviour and their sensitivity to changes in physical parameterizations in the model. While further work is clearly necessary, we hope that this work inspires other modelling groups to carry out similar analysis with their own models in order that some of the major, long-lasting, systematic errors in GCMs can ultimately be reduced.

Code and data availability

Due to intellectual property right restrictions, we cannot provide either the source code or documentation papers for the Met Office Unified Model (MetUM). The MetUM is available for use under licence. For further information on how to apply for a licence see https://www.metoffice.gov.uk/research/approach/collaboration/unified-model/partnership (last access: 16 July 2020). JULES is available under licence free of charge. For further information on how to gain permission to use JULES for research purposes, see https://jules.jchmr.org/software-and-documentation. The model code for NEMO v3.4 is available from the NEMO website (www.nemo-ocean.eu). The model code for CICE is freely available from the United States Los Alamos National Laboratory (http://oceans11.lanl.gov/trac/CICE/wiki/SourceCode). Model data used in this study are archived at the Met Office, and are available to research collaborators upon request.

Author contribution

Gill Martin initiated the study and carried out the analysis of the seasonal hindcasts and free-running climate simulations. Richard Levine designed, ran and analysed the RCM experiments. José Rodriguez and Michael Vellinga analysed the operational coupled and uncoupled NWP simulations. José Rodriguez designed, ran and analysed the nudged simulations. All authors contributed to the writing of the manuscript.

480 **Competing interests**

The authors declare that they have no conflict of interest.

Acknowledgements

This work and its contributors were supported by the UK-China Research & Innovation Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP) China as part of the Newton Fund, and by the Weather and Climate

485 Science for Service Partnership (WCSSP) India, a collaborative initiative between the Met Office, supported by the UK Government's Newton Fund, and the Indian Ministry of Earth Sciences (MoES). NOAA High Resolution SST data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at https://www.esrl.noaa.gov/psd/. ERA5 data were obtained from the Copernicus Climate Change Service Climate Data Store at https://cds.climate.copernicus.eu/cdsapp#!/home.

490 **References**

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P. P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., et al.: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present), J. Hydrometeorol., 4, 1147–1167, 2003.

Amaya, D. J., Kosaka, Y., Zhou, W., Zhang, Y., Xie, S., and Miller, A. J.: The North Pacific Pacemaker Effect on Historical

495 ENSO and Its Mechanisms, J. Climate, 32, 7643–7661, https://doi.org/10.1175/JCLI-D-19-0040.1, 2019.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R .L. H. Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description –Part 1: Energy and water fluxes, Geosci. Model Dev., 4, 677–699, doi:10.5194/gmd-4-677-2011, 2011.

500 Blockley, E. W., Martin, M. J., Mclaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., Mirouze, I., Peterson, K.A., Sellar, A., and Storkey, D.: Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts, Geosci. Model Dev., 7, 2613–2638, 2014.

Bowler, N., Arribas, A., Beare, S., Mylne, K. E., and Shutts, G.: The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system, Q. J. Roy. Meteor. Soc.,135, 767–776, 2009.

505 Bush, S. J., Turner, A. G., Woolnough, S. J., Martin, G. M. and Klingaman, N. P.: The effect of increased convective entrainment on Asian monsoon errors in the MetUM general circulation model, Q. J. Roy. Meteor. Soc., 141, 311-326. doi:10.1002/qj.2371, 2015.

Copernicus Climate Change Service (C3S): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), accessed August 2020. https://cds.climate.copernicus.eu/cdsapp#!/home, 2017.

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

Deser, C., Simpson, I. R., McKinnon, K. A., and Phillips, A. S.: The Northern Hemisphere extratropical atmospheric circulation response to ENSO: How well do we know it and how do we evaluate models accordingly? J. Climate, 30, 5059–5082, https://doi.org/10.1175/JCLI-D-16-0844.1, 2017.

Ding, Y. H., and Chan, J. C. L.: The East Asian summer monsoon: An overview. Meteor. Atmos. Phys., 89, 117–142, https://doi.org/10.1007/s00703-005-0125-z, 2005.

Hermanson, L., Ren, H., Vellinga, M., Dunstone, N. D., Hyder, P., Ineson, S., Scaife, A. A., Smith, D. M., Thompson, V., Tian, B., and Williams, K. D.: Different types of drifts in two seasonal forecast systems and their dependence on ENSO. Clim. Dynam., 51, 1411–1426, https://doi.org/10.1007/s00382-017-3962-9, 2018.

Huang, B., Hu, Z .-Z. and Jha, B.: Evolution of model systematic errors in the tropical Atlantic Basin from coupled climate hindcasts, Clim. Dynam., 28, 668–682, 2007.

Huffman, G. J., and Bolvin, D. T.: TRMM and Other Data Precipitation Data Set Documentation. Laboratory for Atmospheres, NASA Goddard Space Flight Center and Science Systems and Applications. [Available online from http://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf.], 2013.

Huffman, G. J., Adler, R. F., Bolvin, D. T., and Nelkin, E. J.: The TRMM multi-satellite precipitation analysis (TMPA).Satellite Rainfall Applications for Surface Hydrology, M.Gebremichael and F. Hossain, Eds., Springer, 3−22, DOI:10.1007/978-90-481-2915-7_1, 2010.

Hunke, E. C. and Lipscomb, W. H.: CICE: The Los Alamos sea ice model, documentation and software. Version 3.1, Tech. Rep. LA-CC-98-16, Los Alamos National Laboratory, Los Alamos, NM, 2004.

Hurrell, J., Meehl, G. A., Bader, D., Delworth, T. L., Kirtman, B. and Wielicki, B.: A unified modeling approach to climate system prediction, Bull. Amer. Meteorol. Soc., 90, 1819–1832, 2009.

Johnson, S. J., Levine, R. C., Turner, A. G., Martin, G.M., Woolnough, S.J., Schiemann, R., Mizielinski, M.S., Roberts, M.J., Vidale, P.L., Demory, M.-E., and Strachan, J.: The resolution sensitivity of the South Asian monsoon and Indo-Pacific in a global 0.35° AGCM, Clim. Dynam. 46, 807-831. https://doi.org/10.1007/s00382-015-2614-1, 2016.

Johnson, S.J., Turner, A. G., Woolnough, S., Martin, G., and MacLachlan, C.: GloSea5-GC2 seasonal forecasts of the Indian Monsoon: Skill and mechanisms. Clim. Dynam., 48, 1447-1465. doi:10.1007/s00382-016-3151-2, 2017.

Kara, A. B., Rochford, P. A., and Hurlburt, H. E.: An optimal definition for ocean mixed layer depth, J. Geophys. Res., 105 (C7), 16803– 16821, doi:10.1029/2000JC900072, 2000.

Karmacharya, J., Levine, R. C., Jones, R., Moufouma-Okia, W., New, M.: Sensitivity of systematic errors in South Asian summer monsoon simulations to regional climate model domain size and implications for downscaled regional process studies. Clim. Dynam., 45(1–2), 213–231. doi: 10.1007/s00382-015-2565-6, 2015.

Keane, R. J., Williams, K. D., Stirling, A. J., Martin, G. M., Birch, C. E., and Parker, D. J.: Fast Errors in Monsoon Rainfall over Southern and Central India in the Met Office Unified Model, J. Climate, 32, 6385–6402, https://doi.org/10.1175/JCLI-D-18-0650.1, 2019.

Klinker, E.: Investigation of systematic errors by relaxation experiments, Q. J. Roy. Meteorol. Soc., 116, 573–594, 1990.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., and Simpson, J.: The Tropical Rainfall Measuring Mission (TRMM) sensor package. J. Atmos. Oceanic Technol., 15(3), 809−817, https://doi.org/10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2, 1998.

Lazar, A., Vintzileos, A., Doblas-Reyes, F. J., Rogel, P. and Delecluse, P.: 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: 387-397. doi:10.1111/j.1600-0870.2005.00135.x, 2005.

Levine, R. C., and Turner, A. G.: Dependence of Indian monsoon rainfall on moisture fluxes across the Arabian Sea and the impact of coupled model sea surface temperature errors, Clim. Dynam., 38, 2167–2190, https://doi.org/10.1007/s00382-011-1096-z, 2012.

Levine, R. C., and Martin, G. M.: On the climate model simulation of Indian monsoon low pressure systems and the effect of remote disturbances and systematic errors, Clim. Dynam., 50, 4721-4743. https://doi.org/10.1007, 2018.

Liu, H., Zhang, M., and Lin, W.: An investigation of the initial development of the double -ITCZ warm SST errors in the CCSM, J. Climate, 25, 140–155, 2012.

MacLachlan, C. , Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., Gordon, M., Vellinga, M., Williams, A. , Comer, R. E., Camp, J. , Xavier, P. and Madec, G.: Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system, Q. J. Roy. Meteor. Soc., 141, 1072-1084. doi:10.1002/qj.2396, 2015.

Madec, G.: NEMO ocean engine, Tech. Rep. 27, Note du Pole demodélisation, Institut Pierre-Simon Laplace (IPSL), 2008.

565 Marathayil, D., Turner, A. G., Shaffrey, L. C. and Levine, R. C.: Systematic winter sea-surface temperature errors in the northern Arabian Sea in HiGEM and the CMIP3 models, Environ. Res. Lett., 8, 014028, doi:10.1088/1748-9326/8/1/014028, 2013.

Martin, G.M., Milton, S. F., Senior, C.A., Brooks, M. E., Ineson, S., Reichler, T. and Kim, J.:Analysis and Reduction of Systematic Errors through a Seamless Approach to Modeling Weather and Climate, J. Climate, 23, 5933—5957, 2010.

570 Phillips, T.J., and Co-authors: Evaluating parameterizations in General Circulation Models: Climate simulation meets weather prediction, Bull. Amer. Meteorol. Soc., 85, 1903–1915, 2004.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res., 108, D14, 4407, doi:10.1029/2002JD002670, 2003.

575 Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G.: Daily High-Resolution-Blended Analyses for Sea Surface Temperature, J. Climate, 20, 5473-5496, 2007.

Ringer, M. A., Martin, G. M., Greeves, C. Z., Hinton, T. J., James, P. M., Pope, V. D., Scaife, A. A., Stratton, R. A., Inness, P. M., Slingo, J. M. and Yang, G. -Y.:The physical properties of the atmosphere in the new Hadley Centre Global Environment Model, HadGEM1. Part II: Aspects of variability and regional climate, J. Climate, 19, 1302-1326,
580 http://journals.ametsoc.org/doi/full/10.1175/JCLI3713.1, 2006.

Rodwell, M.J., and Palmer, T.N.: Using numerical weather prediction to assess climate models, Q. J. Roy. Meteor. Soc., 133, 129–146, 2007.

Rodríguez, J. M., Milton, S. F., and Marzin, C.: The East Asian atmospheric water cycle and monsoon circulation in the Met Office Unified Model. J. Geophys. Res.-Atmos., 122, 10,246–10,265. https://doi.org/10.1002/2016JD025460, 2017.

585 Rodriguez, J. M. and Milton, S. F.: East Asian Summer Atmospheric Moisture Transport and its Response to Interannual Variability of the West Pacific Subtropical High: an evaluation of the Met Office Unified Model, Atmosphere, 10, 457, doi:10.3390/atmos10080457, 2019.

Saji, N., Goswami, B., Vinayachandran, P. and Yamagata, T.: A dipole mode in the tropical Indian Ocean, Nature, 401, 360–363. https://doi.org/10.1038/43854, 1999.Siongco, A. C., Ma, H., Klein, S. A., Xie, S., Karspeck, A. R., Raeder, K. and
590 Anderson, J. L.: A Hindcast Approach to Diagnosing the Equatorial Pacific Cold Tongue SST Error in CESM1, J. Clim., 33, 1437–1453, https://doi.org/10.1175/JCLI-D-19-0513.1, 2020.

Sperber, K.R, Annamalai, H., Kang, I.S., Kitoh, A., Moise, A., Turner, A.G., Wang, B., and Zhou, T.: The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century, Clim. Dynam., 41, 2711–2744. doi: 10.1007/s00382-012-1607-6, 2013.

595 Vannière, B., Guilyardi, E., Madec, G. et al.: Using seasonal hindcasts to understand the origin of the equatorial cold tongue error in CGCMs and its impact on ENSO, Clim. Dynam., 40, 963–981, https://doi.org/10.1007/s00382-012-1429-6, 2013.

Vannière, B., Guilyardi, E., Toniazzo, T. et al.: A systematic approach to identify the sources of tropical SST errors in coupled models using the adjustment of initialised experiments, Clim. Dynam., 43, 2261–2282. https://doi.org/10.1007/s00382-014-2051-6, 2014.

600 Vellinga, M., Copsey, D., Graham, T., Milton, S. and Johns, T. : Evaluating benefits of two-way ocean-atmosphere coupling for global NWP forecasts, acc. Weather and Forecasting, 2020.

Voldoire, A., Exarchou, E., Sanchez-Gomez, E. et al.: Role of wind stress in driving SST errors in the Tropical Atlantic, Clim. Dynam., 53, 3481–3504. https://doi.org/10.1007/s00382-019-04717-0, 2019.

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T.,
605 Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P. L., Wilkinson, J., Willett, M., Woolnough, S., and Xavier, P.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, Geosci. Model Dev., 10, 1487–1520, https://doi.org/10.5194/gmd-10-1487-2017, 2017.

610 Walters, D., Baran, A.J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., et al.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, Geosci. Model Dev., 12, 1909–1963, 2019.

Wang, B., Zhang, Y., and Lu, M. M.: Definition of South China Sea monsoon onset and commencement of the East Asia summer monsoon. J. Clim., 17, 699–710, 2004.

615 Wang, B., Wu, Z., Li, J., Liu, J., Chang, C., Ding, Y., and Wu, G.: How to Measure the Strength of the East Asian Summer Monsoon. J. Clim., 21, 4449–4463, https://doi.org/10.1175/2008JCLI2183.1, 2008.

Wang, C., Zou, L., and Zhou, T.: SST biases over the Northwest Pacific and possible causes in CMIP5 models. Sci. China Earth Sci., 61, 792–803, https://doi.org/10.1007/s11430-017-9171-8, 2018

Waters, J, Lea DJ., Martin, M.J., Mirouze, I., Weaver, A. and While, J.: Implementing a variational data assimilation system
620 in an operational 1/4 degree global ocean model, Quart. J. Roy. Meteor. Soc., 141, 687, 333-349, https://doi.org/10.1002/qj.2388 , 2014.

Williams, K. D., Harris, C. M., Bodas-Salcedo, A., Camp, J., Comer, R. E., Copsey, D., Fereday, D., Graham, T., Hill, R., Hinton, T., Hyder, P., Ineson, S., Masato, G., Milton, S. F., Roberts, M. J., Rowell, D. P., Sanchez, C., Shelly, A., Sinha, B.,

Walters, D. N., West, A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled model 2.0 (GC2) configuration,
625 Geosci. Model Dev., 8, 1509–1524, https://doi.org/10.5194/gmd-8-1509-2015, 2015.

Wu, B., Zhou, T., and Li, T.: Impacts of the Pacific–Japan and Circumglobal Teleconnection Patterns on the Interdecadal
Variability of the East Asian Summer Monsoon. J. Climate, 29, 3253–3271, https://doi.org/10.1175/JCLI-D-15-0105.1,
2016.

Xie, P., Janowiak, J., Arkin, P., Adler, R.F., Gruber, A., Ferraro, R., Huffman, G.J., Curtis, S.: GPCP pentad precipitation
630 analyses: an experimental data set based on gauge observations and satellite estimates, J. Climate, 16(13), 2197–2214, 2003.

Zhang, Y., Xie, S., Kosaka, Y., and Yang, J.: Pacific Decadal Oscillation: Tropical Pacific Forcing versus Internal
Variability, J. Climate, 31, 8265–8279, https://doi.org/10.1175/JCLI-D-18-0164.1, 2018.

635

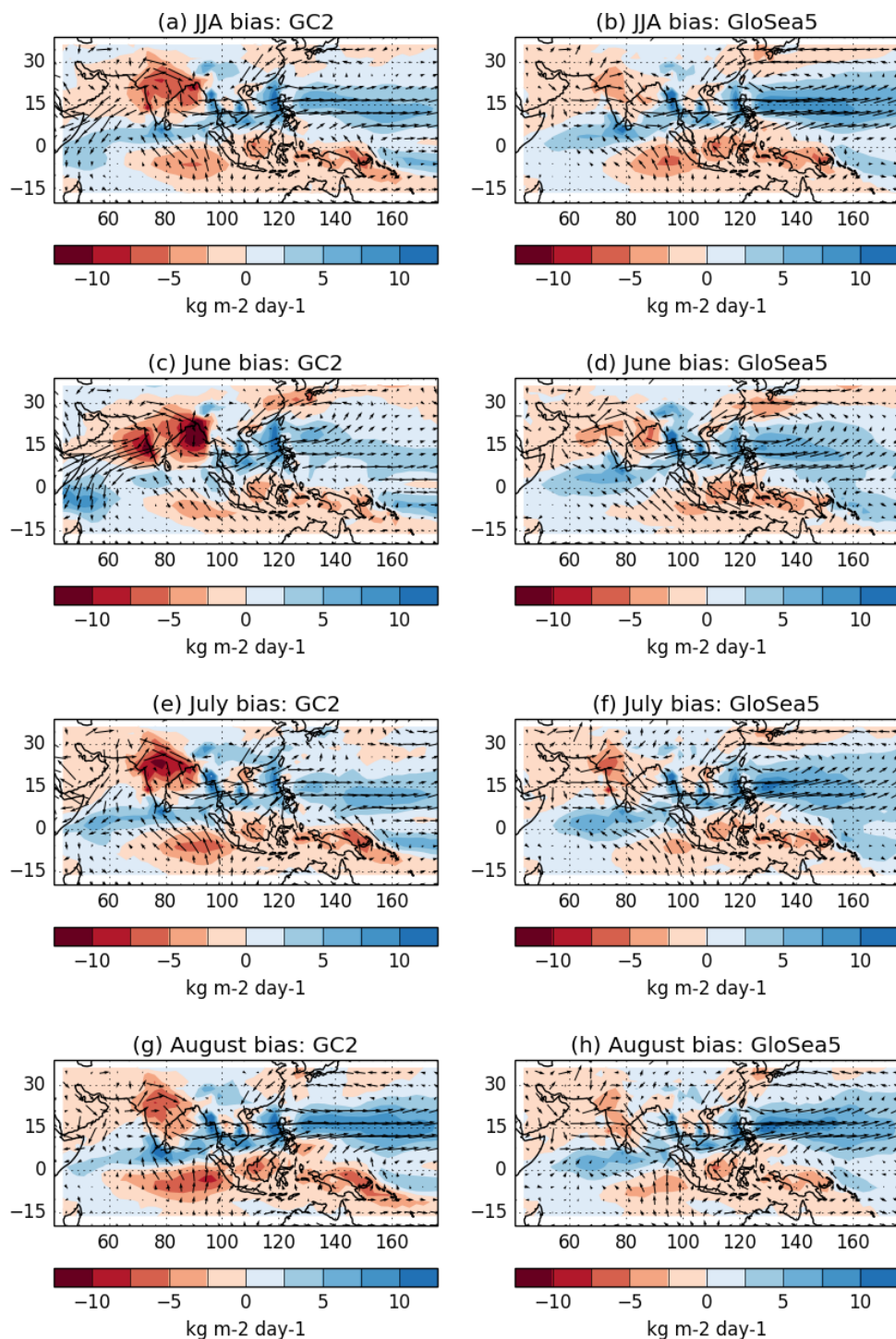



Figure 1: (a,b) JJA, (c,d) June, (e,f) July and (g,h) August climatological errors in precipitation (against GPCP observations) and
850 hPa winds (against ERA-Interim reanalyses) from the current GloSea5 23-year operational hindcast ensemble initialised each
year on four start dates (1, 9, 17, 25) in April, May, June and July respectively, with 7 members per start date.

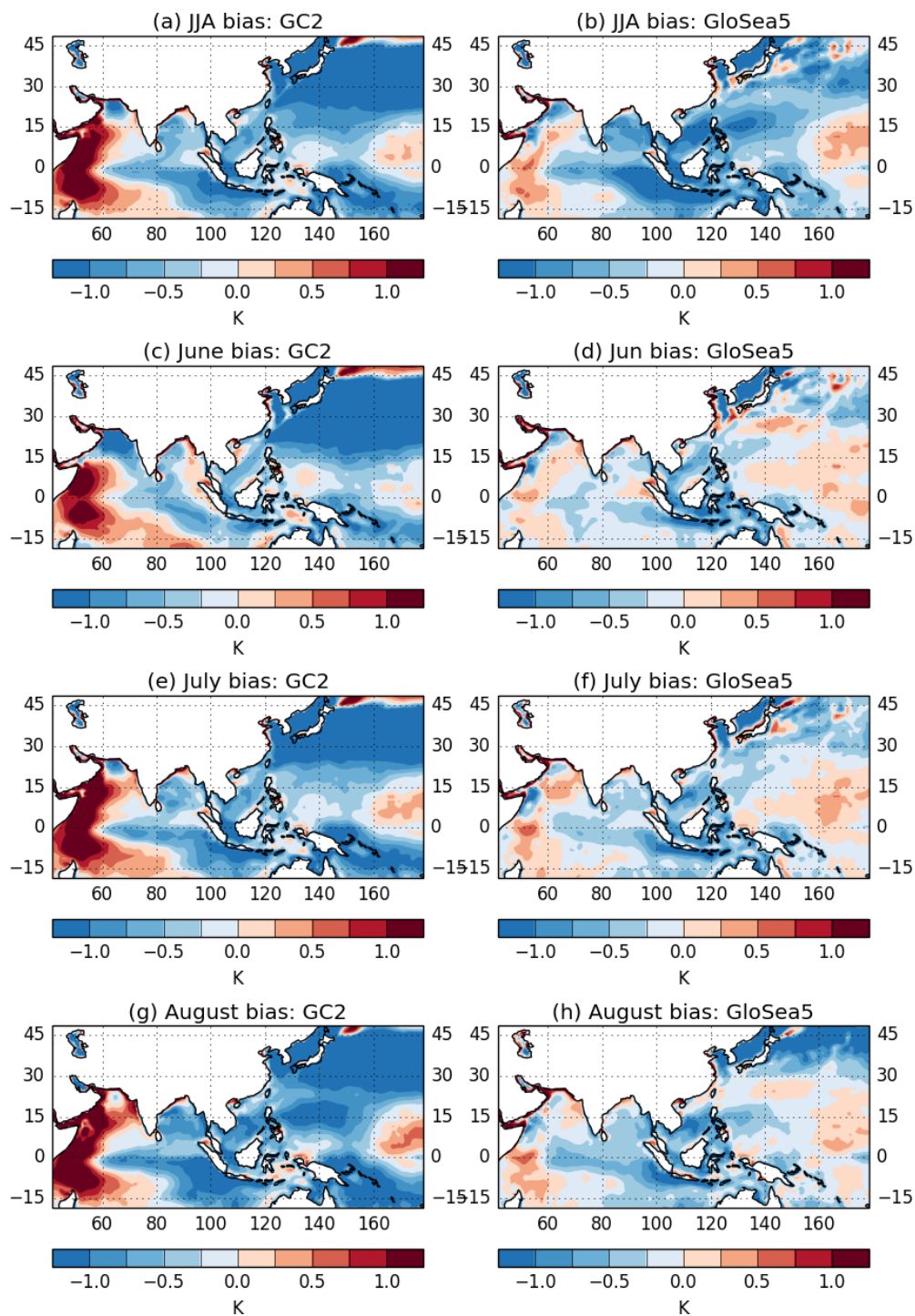



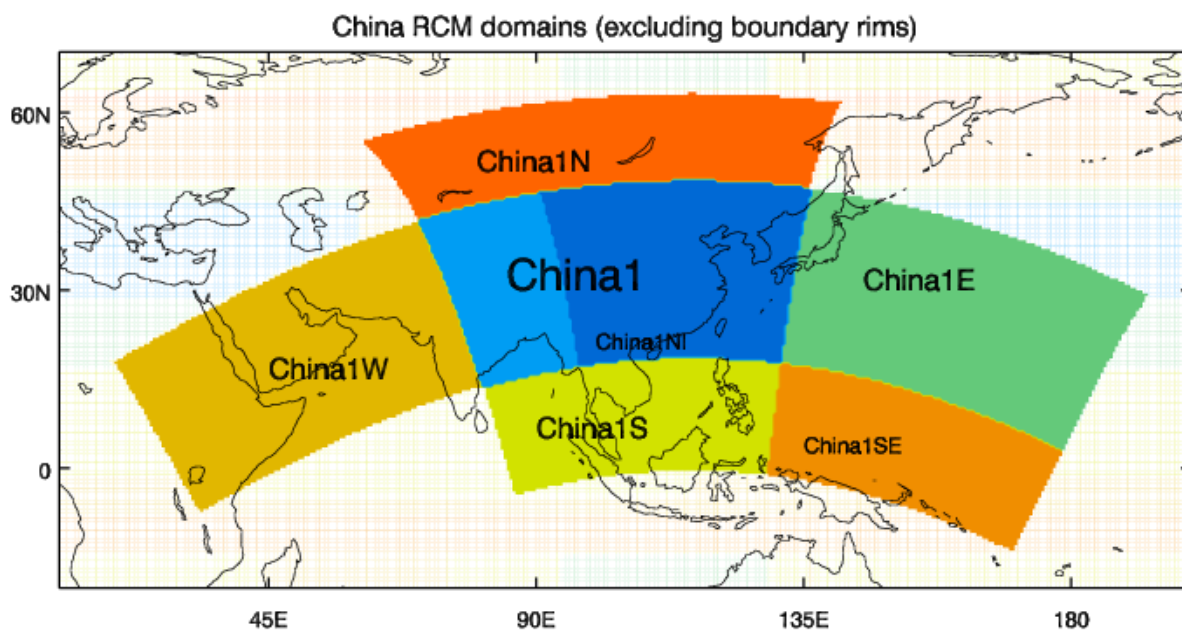
Figure 2: As Fig 1 but for sea surface temperature, compared against HadISST observations.

645

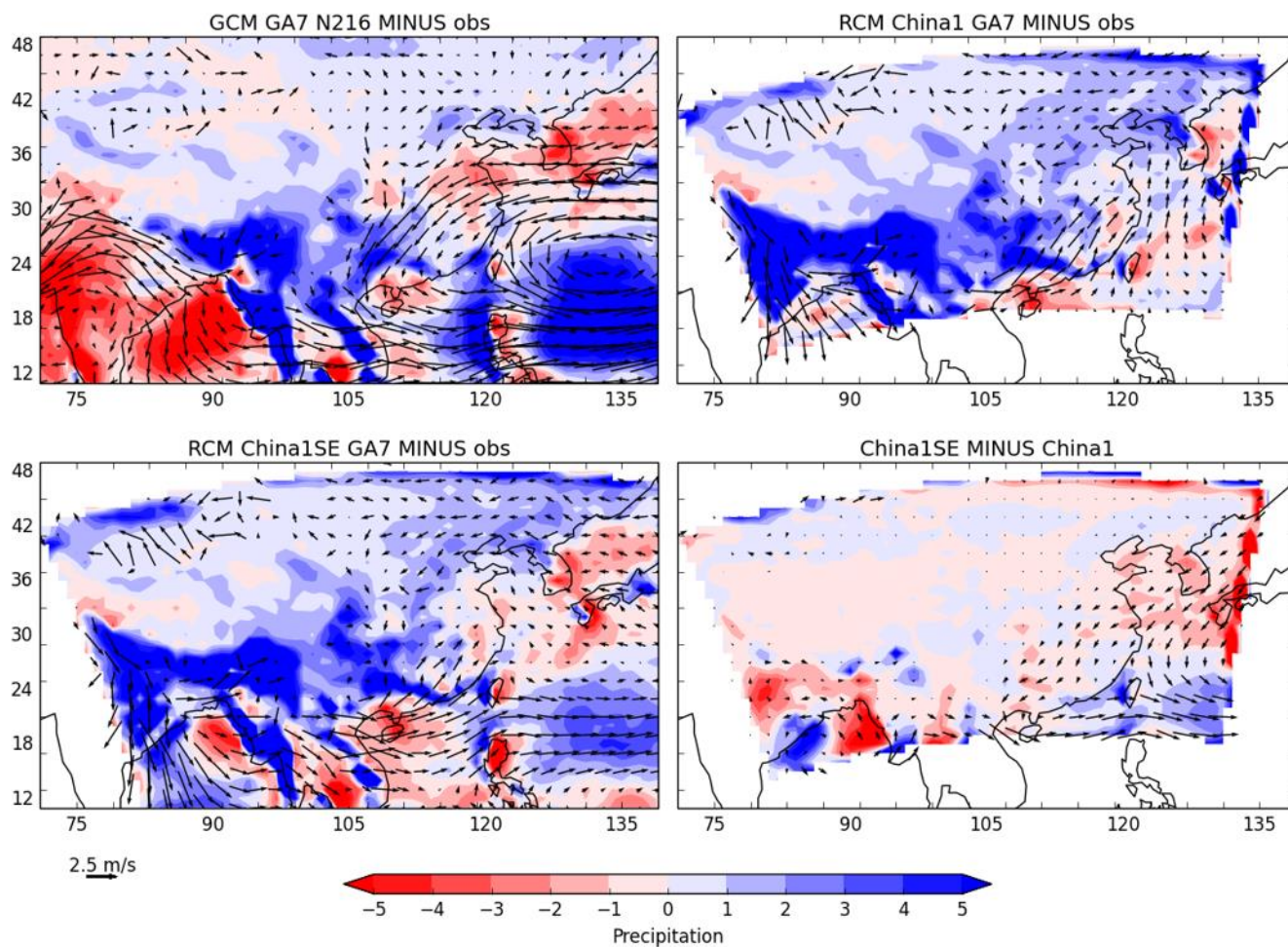
Figure 3: Domains used in Regional Climate Model experiments. China1 is the smallest, central domain, with the other domains obtained by extending this to the north, south, east and west. All RCM simulations use a 0.44° × 0.44° resolution grid and a rotated pole at 61°N, 296.3°E.

650

Figure 4: (top) JJA climatological (1989-2008) precipitation and 850hPa wind errors in N216 AGCM simulation and the China1 RCM simulation vs TRMM and ERA-Interim (obs). (bottom) Effects of extending the domain towards the south and east: (left) China1SE minus obs; (right) China1SE minus China1. Colour bar indicates precipitation differences (mm/day) and vectors indicate differences in 850 hPa winds.

655

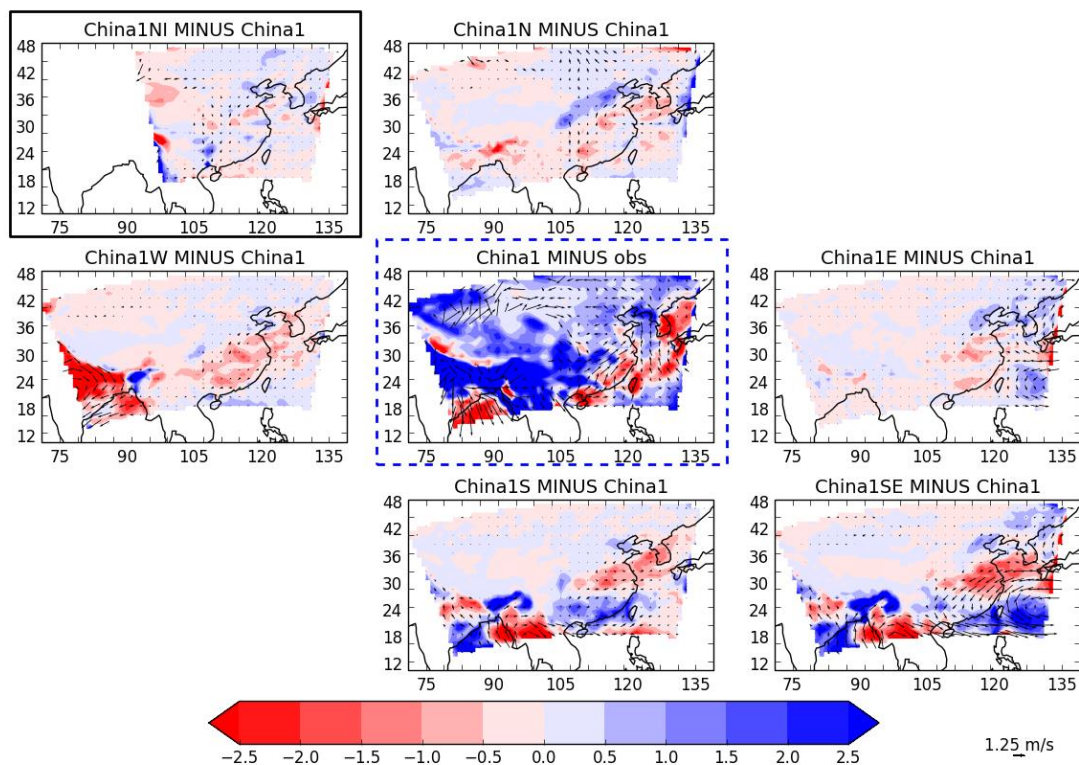



Figure 5: Effects of extending the China1 domain in different directions, as indicated in Fig. 3. Differences are against China1 in each case, while the centre panel shows China1 minus observations (note difference in colour scale from equivalent plot in Fig. 4). Plots are for JJA 1989-2005. Colour bar indicates precipitation differences (mm/day) and vectors indicate differences in 850 hPa winds.

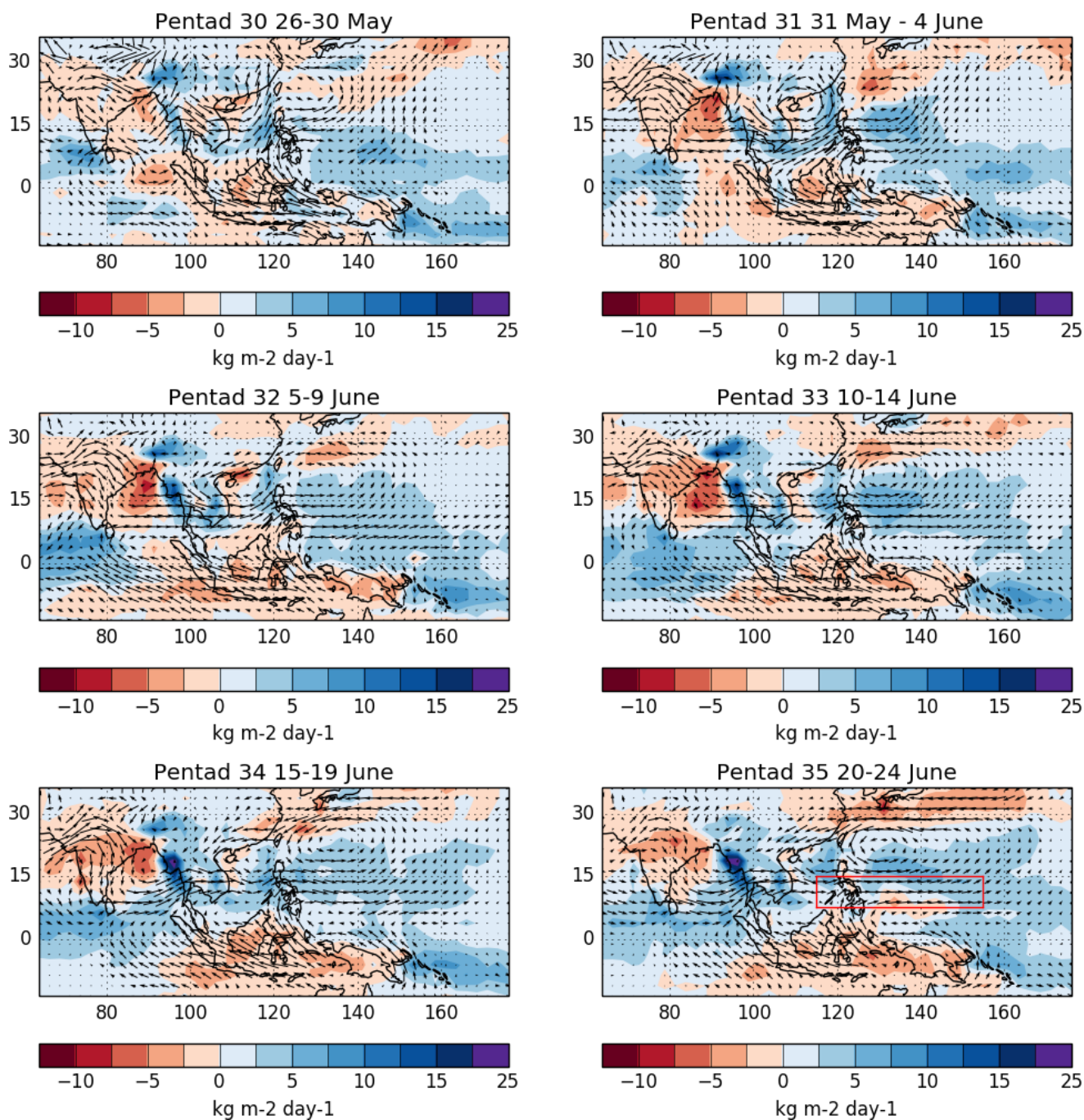



Figure 6: Development of climatological errors in 850 hPa winds and precipitation (against ERA-I and GPCPv2), pentad by pentad after initialisation on 25th May. The red box indicates the "Philippines" region used Fig. 11(c).

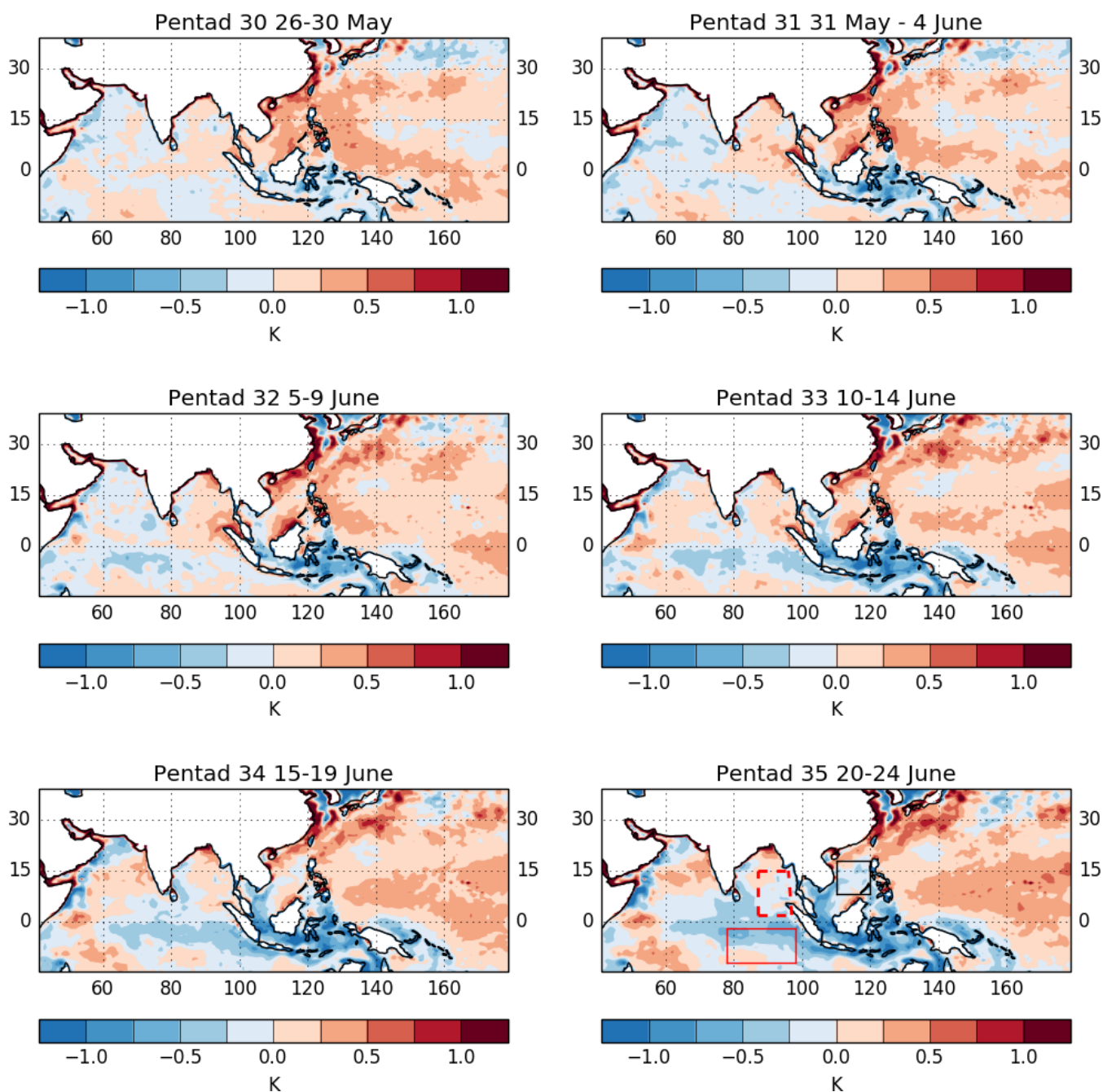



Figure 7: Development of climatological errors in SST (OISSTv2), pentad by pentad after initialisation on 25th May. The red and black boxes indicate the EEIO and SCS boxes used in Fig. 11, respectively.

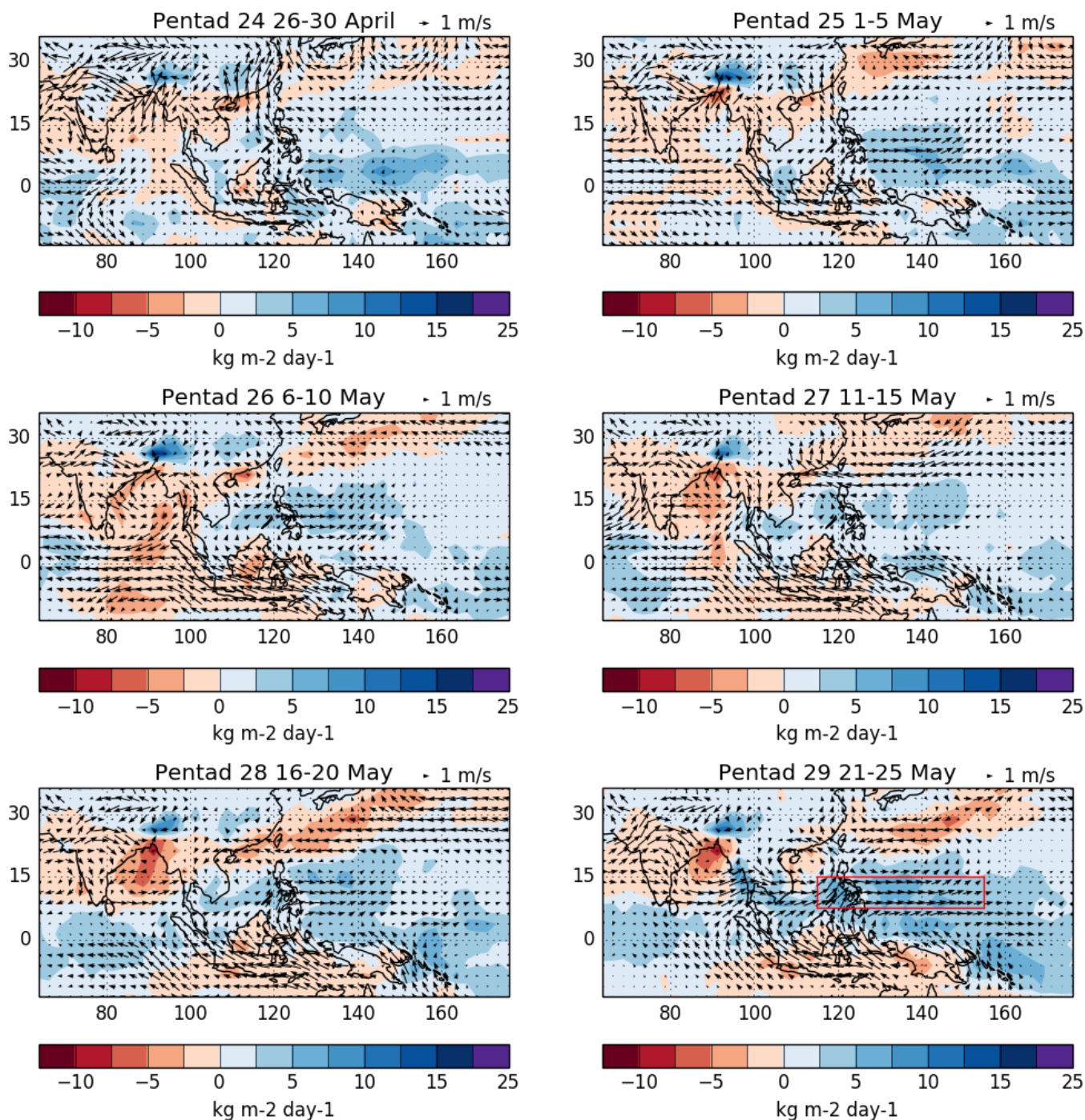



Figure 8: Development of climatological errors in 850 hPa winds and precipitation (against ERA-I and GPCPv2), pentad by pentad after initialisation on 25th April. The red box indicates the "Philippines" region used Fig. 11(c).

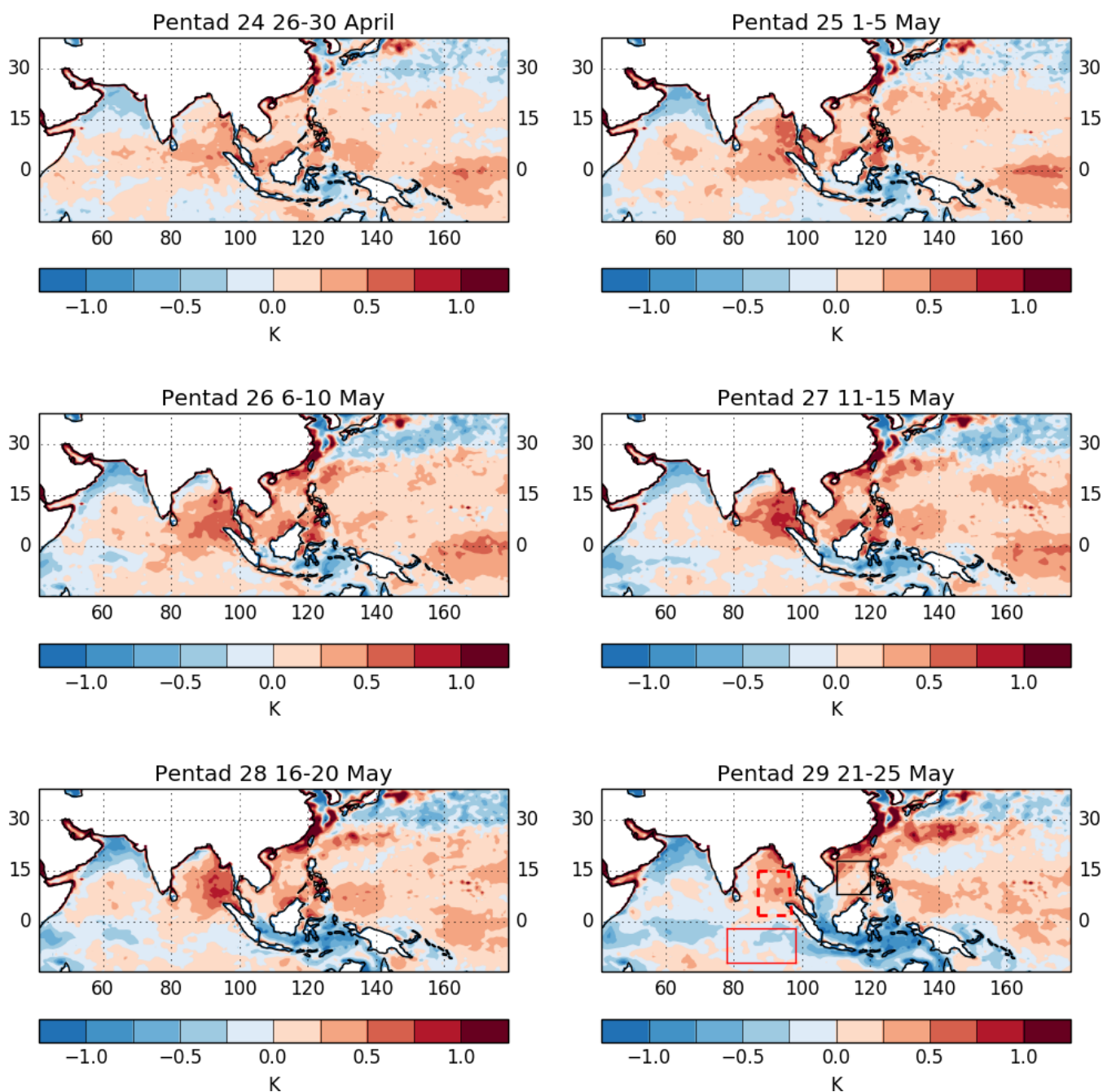



Figure 9: Development of climatological errors in SST (OISSTv2), pentad by pentad after initialisation on 25th April. The red and black boxes indicate the EEIO and SCS boxes used in Fig. 11, respectively.

675

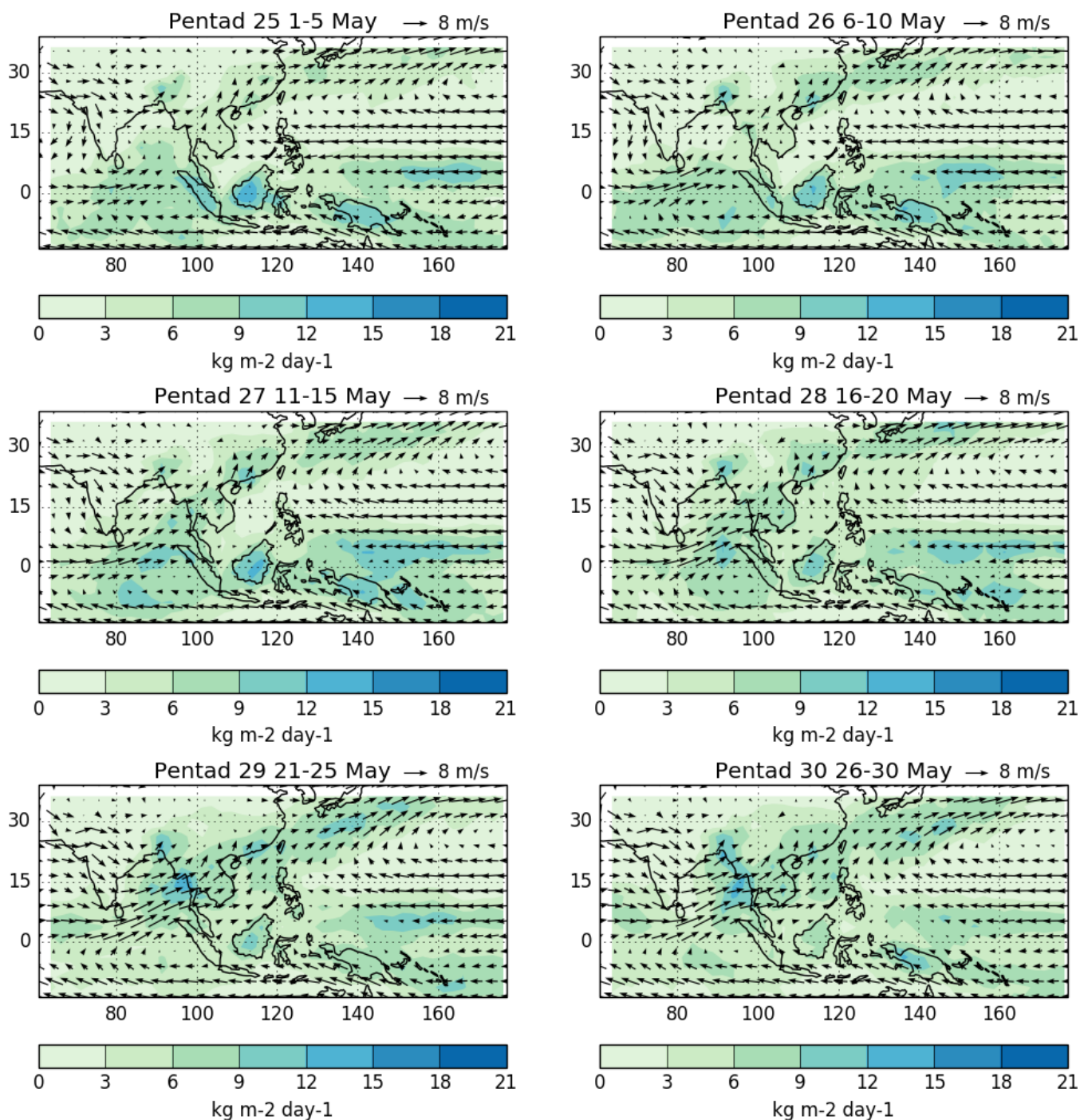



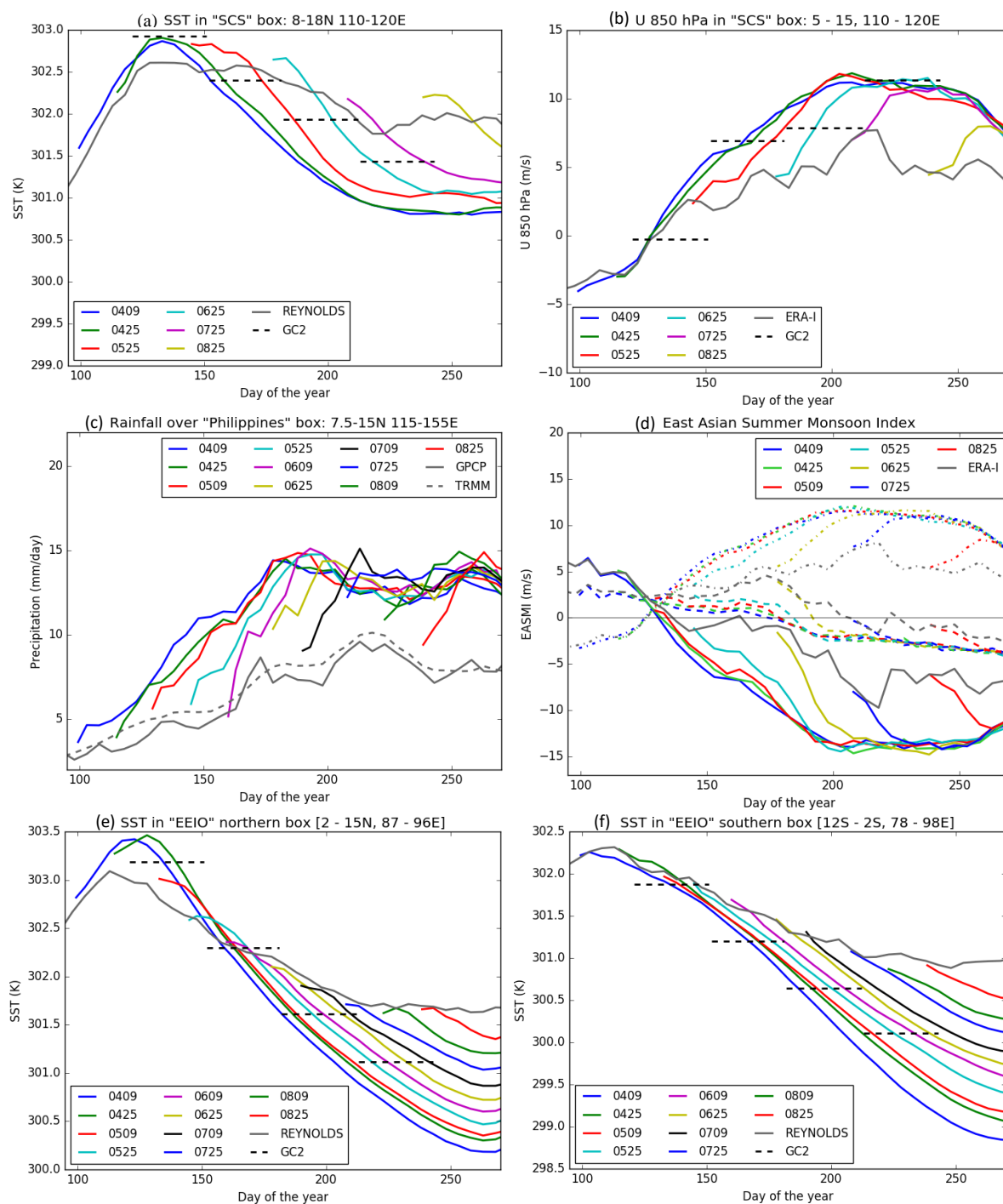
Figure 10: Pentad evolution of 850 hPa winds (ERA-Interim) and rainfall (GPCPv2) through the broadscale seasonal transition towards the Asian summer monsoon season, heralded by the reversal of wind direction over the South China Sea which typically occurs around pentad 28 (Wang et al., 2004).

680

Figure 11: Climatological spin-up of various quantities over key regions in GloSea5 hindcast ensembles initialised on 25 April (0425), 25 May (0525), 25 June (0625), 25 July (0725), 25 August (0825), 7 members for each, averaged over 1993-2015: (a) SST in South China Sea (8°-18°N, 110°-120°E); (b) 850 hPa zonal wind over SCS box used by Wang et al (2004) (5°-15°N, 110°-120°E); (c)

685 **Rainfall over the Philippines box (7.5°-15°N, 115°-155°E); (d) EASMI: 850 hPa zonal wind difference (22.5°-32.5°N, 110°-140°E) – (5°-15°N, 90°-130°E); (e) SST in EEIO northern box (2°-15°N, 87°-96°E); (f) SST in EEIO southern box (12°S-2°S, 78°-98°E). Pentad and monthly SST is shown from OISSTv2 and HadISST respectively, pentad rainfall from GPCP and TRMM observations, and EASMI from ERA-I. For EASMI, the solid lines indicate the index, while the dashed line indicates the northern box, dot-dash indicates southern box. Black dashed horizontal lines indicate monthly averages from free-running GC2 coupled model simulations. Day of the year is calculated using the Gregorian calendar (without leap years); day 121 corresponds to 1 May.**

690

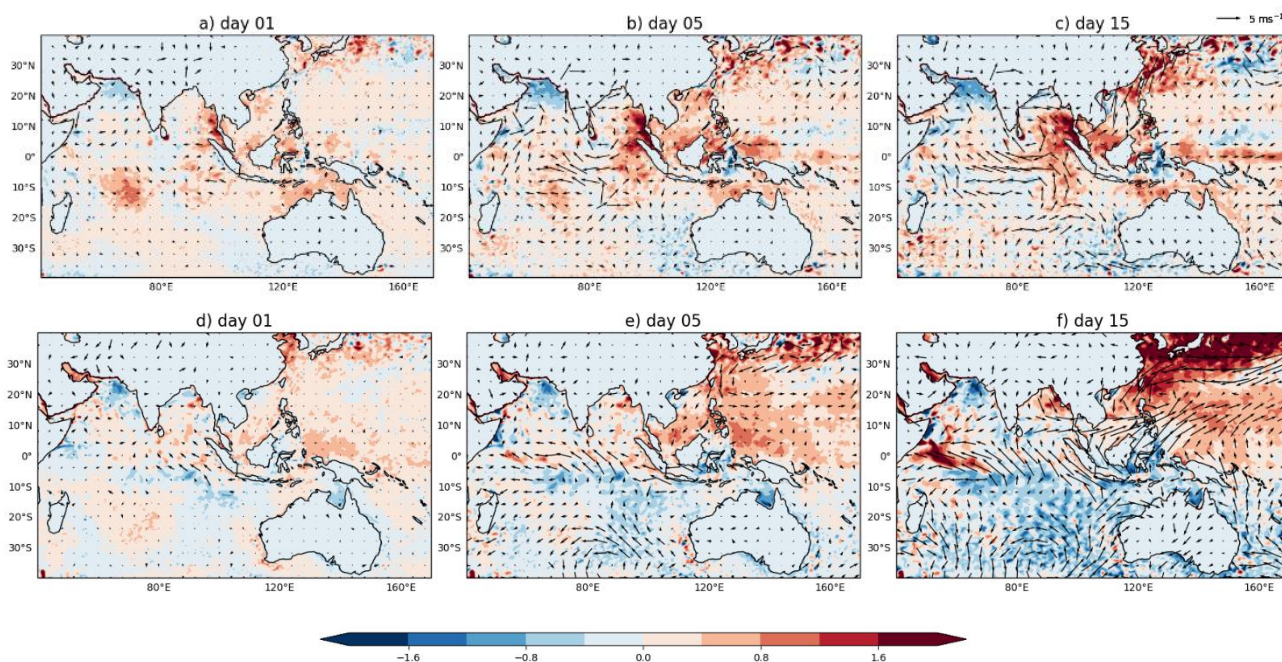



Figure 12: SST error (shading, C) and 10-m wind errors (arrows) from coupled NWP-N768 simulations with respect to OISSTv2 in composites before (10-19 May; top row) and after (10 – 23 July; bottom row) the broadscale seasonal transition in 2016, for forecast lead times of 1, 5 and 15 days.

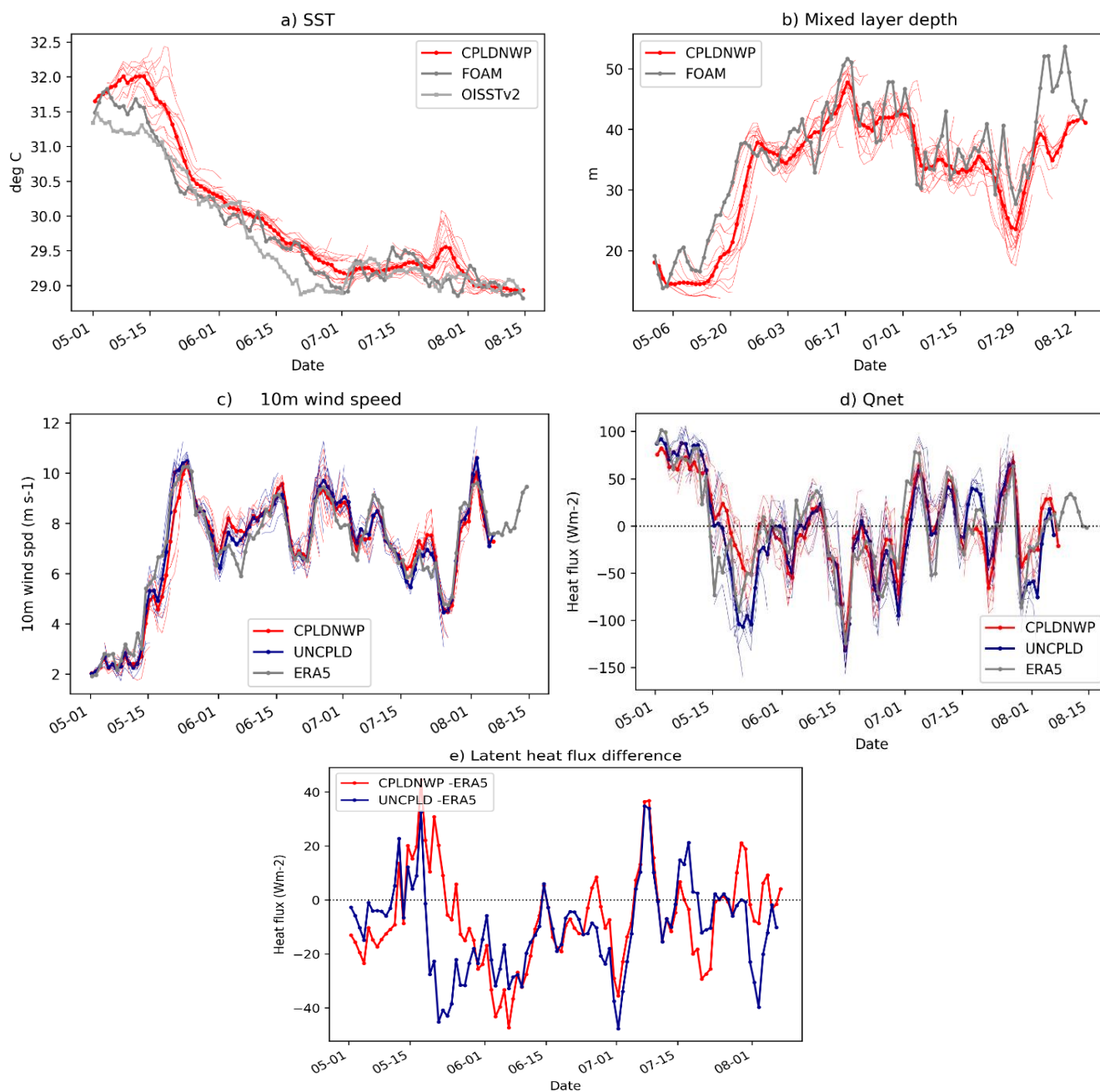
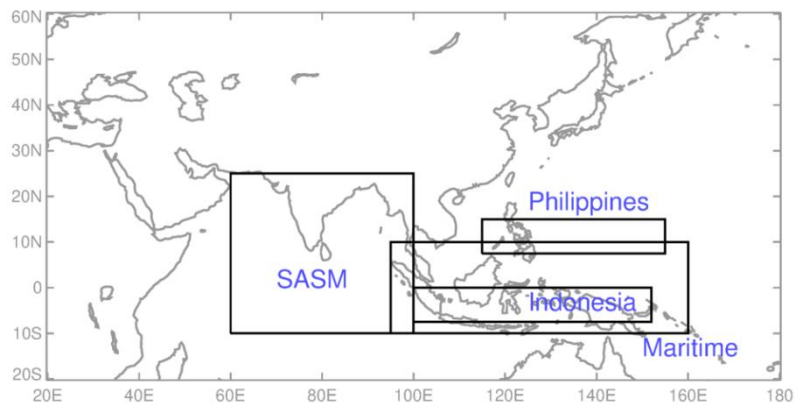



Figure 13: Time series of daily forecasts initialised between 1 May–31 July 2016. Individual forecasts (thin lines) are averaged at each validity time ('Date') into a mean forecast (coloured heavy line). Grey lines are from analyses. Lines show area-averages over EEIO northern box (see Figure 11) (a) SST from CPLDNWP and FOAM analyses, out to day 15; (b) mixed layer depth calculated using daily mean temperature and salinity from CPLDNWP and FOAM analyses, out to day 15; (c) 10m wind speed from CPLDNWP (red) and UNCPLD (blue) out to day 7; (d) net surface heat flux from CPLDNWP (red) and UNCPLD (blue) out to day 7. Positive (negative) value denotes net flux into (out of) the ocean; (e) Mean forecast error of latent heat flux in CPLDNWP and UNCPLD forecasts, relative to ERA5.

710 **Figure 14: Regions used in nudged climate simulations: "Philippines" [115°-155°E, 7.5°-15°N], "Indonesia" [100°-152.5°E, 7.5°S-0°N], "S Asian Monsoon (SASM)" [60° - 100°E, 10°S - 25°N], and "Maritime Continent" [95° - 160°E, 10°S - 10°N].**

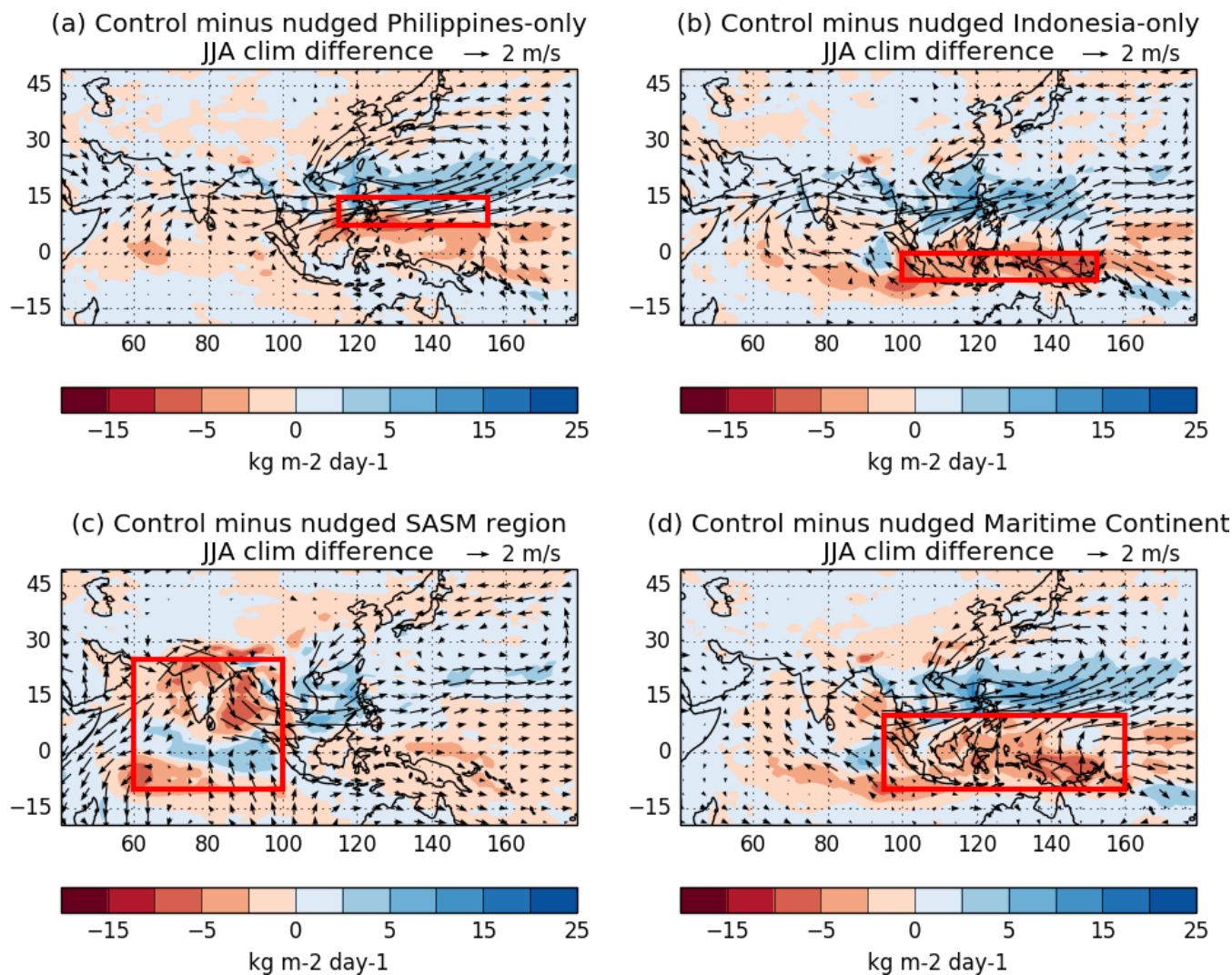



Figure 15: Climatological differences in 850 hPa winds and rainfall in JJA between "Nudged" experiments and their Control, indicating the influence on the Control of errors developing in four different regions used for the Nudged experiments: (a) Philippines; (b) Indonesia; (c) South Asian summer monsoon; (d) Maritime Continent. The nudged regions are shown in red.

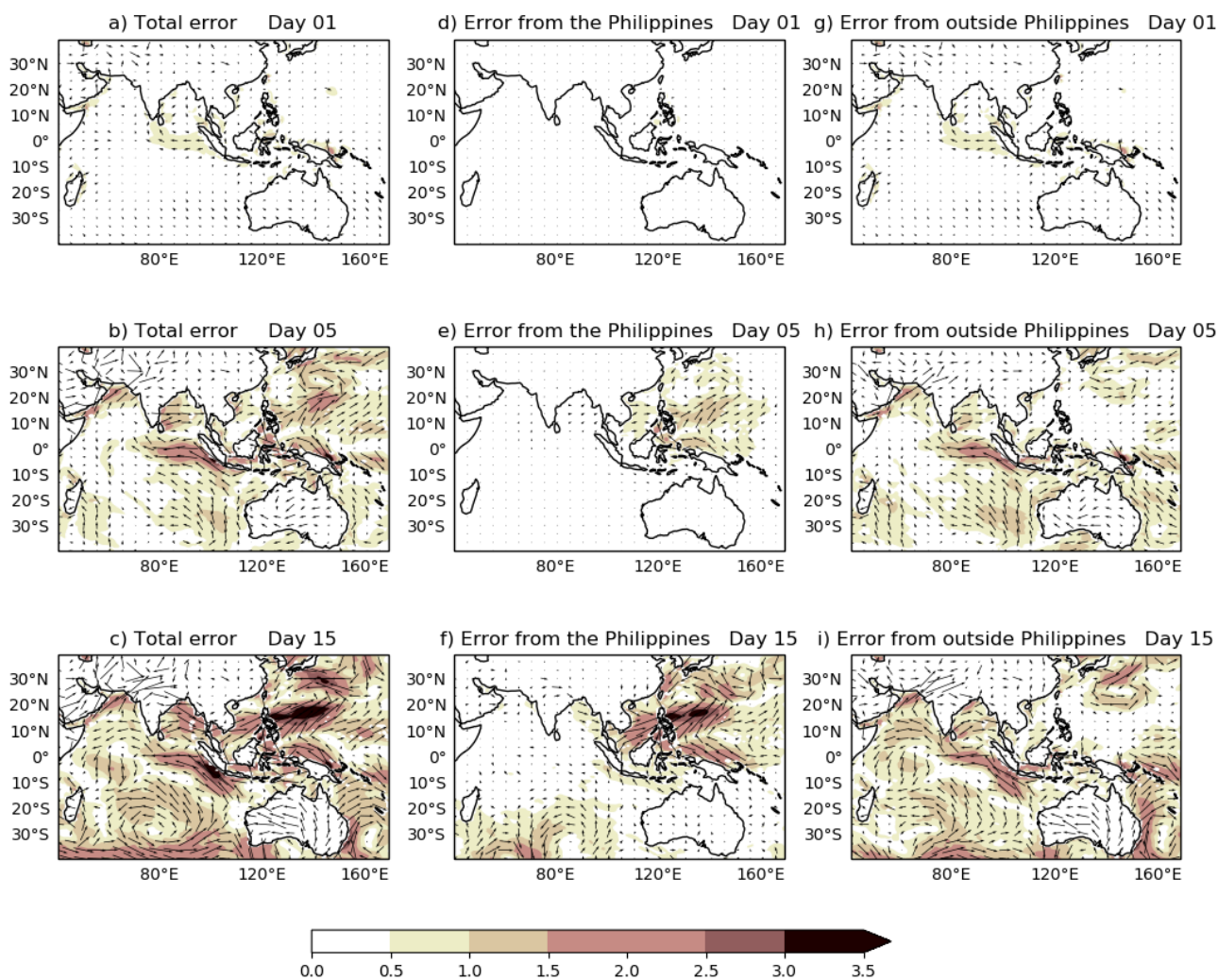



Figure 16: June–August 2016 mean 10m-wind errors (arrows) and their magnitudes (shading, ms -1). (a–c) atmosphere-only NWP-N216 total errors, with respect to MetUM analysis, for forecast lead times of 1, 5 and 15 days. (d–f) NWP-N216 free run minus Philippines nudging (nudged region shown in Fig. 15a), showing errors forced from the Philippines for forecast lead times of 1, 5 and 15 days. (g–i) Philippines nudging minus global nudging, showing errors forced outside of the Philippines for forecast lead times of 1, 5 and 15 days.