The influence of air-sea coupling on forecasts of the 2016 Indian summer monsoon and its intraseasonal variability

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1002/qj.3914

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

RESEARCH ARTICLE

The influence of air–sea coupling on forecasts of the 2016 Indian summer monsoon and its intraseasonal variability

Maria Valdivieso | Simon C. Peatman | Nicholas P. Klingaman

National Centre for Atmospheric
Science-Climate, Department of
Meteorology, University of Reading, UK

Correspondence
M. Valdivieso, National Centre for
Atmospheric Science-Climate,
Department of Meteorology, University of
Reading, Reading RG6 6BB, UK.
Email:
m.valdiviesodacosta@reading.ac.uk

Present address
Simon C. Peatman, School of Earth and
Environment, University of Leeds, Leeds
LS2 9JT, UK

Abstract

Daily initialized coupled and uncoupled numerical weather prediction (NWP) forecasts from the global Met Office Unified Model (MetUM) are compared for the 2016 Indian summer monsoon. Three MetUM configurations are used: atmosphere only (ATM), coupled to a mixed-layer ocean model (KPP), and coupled to a dynamical ocean model (NEMO). The analysis focuses on the impact of air–sea coupling, particularly in the Bay of Bengal (BoB), on NWP for monsoon rainfall. Seasonal-mean biases in all three configurations are highly consistent and driven by errors in atmospheric processes. Rainfall is initially overestimated over India, but underestimated over the BoB, the latter associated with too much short-wave radiation and too little cloud cover in MetUM. The excess short-wave radiation ($>40\,\mathrm{W\cdot m^{-2}}$ over the northwest BoB) is partially compensated by additional latent cooling, primarily due to overestimated surface wind speeds. In NEMO and KPP, coupling improves the timing of intraseasonal active and break phases over India, primarily the end of these phases, which are systematically too late in ATM. NEMO and KPP show a more realistic intraseasonal local phase relationship between sea surface temperature (SST) and rainfall throughout the BoB, but no configuration reproduces the observed significant lagged relationship between BoB SST and Indian rainfall. The lack of this relationship may be partly attributed to weak heat flux feedbacks to northern BoB SST, with the forecast short-wave feedback having systematically the wrong sign (positive) compared to satellite radiation, and thus contributing to SST warming at all lead times. Based on these MetUM forecasts, there is a limited impact of coupling on NWP for monsoon rainfall, both for the mean rainfall and intraseasonal variability. Further research to improve NWP for monsoon rainfall should focus on reducing MetUM atmospheric systematical biases.

KEYWORDS
air–sea coupling, atmospheric convection, Bay of Bengal, Indian monsoon, intraseasonal variability, weather forecasting

1 | INTRODUCTION

Indian summer monsoon rainfall provides a critical freshwater resource for more than 1 billion people on the Indian subcontinent. Although interannual variability in seasonal-total, area-averaged rainfall is relatively small – the standard deviation is approximately 10% of the mean (e.g., Yang and Lau, 2006) – monsoon rainfall variability on synoptic and intraseasonal scales is considerably larger. Synoptic variability is controlled by monsoon lows and depressions, which typically form in the Bay of Bengal (BoB) and track along the monsoon trough that extends northwest–southeast across India (e.g., Sikka, 1977; Boos *et al.*, 2015). These systems cause considerable rainfall, with recent estimates suggesting that depressions are responsible for 10–20% of rainfall across central India (Hunt and Fletcher, 2019). On intraseasonal scales, rainfall is controlled by monsoon active–break cycles, which are linked to the northward-propagating Boreal Summer Intraseasonal Oscillation (BSISO; e.g., Webster *et al.*, 1998; Rajeevan *et al.*, 2010; Lee *et al.*, 2013). The BSISO has a nominal period of approximately 30–60 days, although the frequency of events is highly irregular (Annamalai and Sperber, 2005). Northward-propagating BSISO events are frequently associated with eastward-propagating equatorial convective anomalies, hence the BSISO is described as the boreal summer manifestation of the Madden–Julian Oscillation (MJO; e.g., Lawrence and Webster, 2002). Active BSISO phases over India are associated with stronger westerly winds, higher rainfall over central India, reduced rainfall over southern and northeastern India, and an increase in the frequency of monsoon lows and depressions (e.g., Goswami *et al.*, 2003; Krishnamurthy and Shukla, 2007).

Predicting Indian monsoon precipitation remains a considerable challenge, both for the seasonal mean (e.g., Rajeevan *et al.*, 2012; Saha *et al.*, 2016) and for synoptic and intraseasonal variability (e.g., Routray *et al.*, 2010; Lee *et al.*, 2015; Dutta *et al.*, 2019). We focus our discussion on numerical weather prediction (NWP) temporal scales, for which most previous analysis has been based on case-studies. Dutta *et al.* (2019) demonstrated that refining the horizontal resolution of an atmospheric NWP model from 12 to 1.5 km improved predictions of the track and intensity of a monsoon depression and its associated rainfall, though substantial errors remained. Hunt and Turner (2017b) found similar improvements with resolution for a wider set of depression cases in an NWP model, at coarser scales ranging from 208 to 16 km, as did Johnson *et al.* (2016) in a climate model at similar resolutions. Several recent studies have found that accurate land-surface initialisation and improved treatments of atmospheric boundary-layer processes also improve NWP

forecasts of rainfall, for case-studies of monsoon depressions (Rajesh *et al.*, 2017; Hunt and Turner, 2017a; Rai and Pattnaik, 2019).

Many studies have used climate models to demonstrate that including air–sea coupled feedbacks improves simulated BSISO intensity and propagation (e.g., Fu and Wang, 2004; Klingaman *et al.*, 2011; DeMott *et al.*, 2014); however, few studies have considered the effects of ocean coupling on NWP of Indian summer monsoon rainfall. Coupled feedbacks enable the ocean to respond to variability in atmospheric forcing (via surface fluxes) through changes in SST, which in turn influences surface fluxes. Thus, it is essential to understand air–sea interactions and related upper-ocean processes in the BoB and their representation in monsoon prediction models on scales from NWP to seasonal. While most sub-seasonal and seasonal prediction systems use coupled models (e.g., Vitart *et al.*, 2017) – in recognition of the importance of ocean feedbacks for longer-scale predictions – only a few centres use coupled models for NWP (e.g., the European Centre for Medium-range Weather Forecasting (ECMWF); Buizza *et al.*, 2018). Most contemporary NWP systems use atmosphere-only models with fixed SST, including the Met Office system of interest here, due perhaps to complexities in initialising high-resolution oceans or a perception that coupled feedbacks on NWP scales offer limited benefit. Yet, the 1–2 week NWP range is similar to the time-scale for northward propagation of a BSISO convective event (e.g., Klingaman *et al.*, 2008b). While initialized forecasts from coupled models suffer from similar drawbacks to atmosphere-only forecasts, in particular due to poorly constrained surface fluxes, they can represent high-frequency SST variability which influences surface upward long-wave emission, turbulent (sensible and latent) heat fluxes and evaporation. Moreover, it is unknown how synoptic to intraseasonal prediction of monsoon precipitation depends on the representation of physical processes and model configurations such as spatial resolution, initial ocean stratification (temperature and salinity change with depth) and coupling approaches (e.g., one-dimensional versus three-dimension ocean models). Understanding the benefits and drawbacks of including air–sea coupling in NWP for the Indian monsoon motivates the present study.

The role of air–sea feedbacks in the Indian summer monsoon and its intraseasonal variability have gained attention in recent years, given the availability of high-resolution models and new BoB observations as part of the joint Air–Sea Interactions Regional Initiative (ASIRI) and Ocean Mixing and Monsoon (OMM) programs (e.g., Lucas *et al.*, 2014; Mahadevan *et al.*, 2016; Wijesekera *et al.*, 2016). Several studies have investigated the processes responsible for the intraseasonal SST

response in the BoB in both observations and models (e.g., Vecchi and Harrison, 2002; Shankar *et al.*, 2007; Sharma *et al.*, 2016). One-dimensional processes are generally regarded to dominate the BoB upper-ocean heat budget, particularly in the northern bay, where a shallow surface mixed layer responds quickly to changes in surface heat fluxes (e.g., Agarwal *et al.*, 2007; Weller *et al.*, 2016). A number of studies have emphasized the importance of summer salinity stratification and the formation of barrier layers associated with the export of freshwater from large rivers, such as the Ganga-Brahmaputra-Meghna and Irrawaddy (e.g., Shenoi, 2002; Sengupta *et al.*, 2016), and the sensitivity of this influence to BSISO convective events. For example, Rao *et al.* (2011) found a correlation between heavy summer monsoon rainfall in the river catchment area and subsequent positive sea surface salinity (SSS) anomalies in the northern bay, with a lag of 50–60 days. The role of ocean dynamics in BoB intraseasonal variability is much less clear. Ocean transports of heat and freshwater may have a more substantial role in regions of strong currents and substantial eddy activity; recent studies have used eddy-resolving ocean models to assess small-scale heat and freshwater transports within the BoB that are difficult to observe. For example, Roman-Stork *et al.* (2019) used satellite observations and 1/12° ocean model simulations to study BISO variability (10–20 days) in the BoB during summer monsoons from 2016 to 2018. They found that intraseasonal SST and SSS anomalies in the northern BoB reflected atmospheric forcing, whereas the cooling and salinification of the central BoB are controlled by wind-driven upwelling in the central BoB. These estimates strengthen the importance of a dynamically active ocean and indicate the need to re-examine bay-scale upper-ocean budgets, although large uncertainties in reanalysis-based surface fluxes over the BoB, which are used to force the ocean-only models, are a recognized drawback (e.g., Sanchez-Franks *et al.*, 2018).

Other studies have evaluated the role of ocean feedbacks on MJO and BSISO atmospheric convection through a moist static energy (MSE) budget-based approach, highlighting the substantial role of intraseasonal SST variations in surface turbulent (sensible and latent) flux anomalies, through changes in near-surface humidity and temperature gradients (e.g., DeMott *et al.*, 2016; Gao *et al.*, 2019). Gao *et al.* (2019) performed such a MSE budget analysis in the context of the BSISO using recent atmospheric reanalysis products for the Indo-Pacific Warm Pool (20°S–40°N, 30°E–120°W), and found that enhanced surface turbulent fluxes induced by intraseasonal SST anomalies in this region were located ahead (i.e., to the north) of convection, favouring northward BSISO propagation. A key result in both observation- or model-based studies is the lead–lag phase relationships between SST, air–sea fluxes

and atmospheric convection on intraseasonal scales. In a forecast, air–sea fluxes are needed to predict the evolution of SST and atmospheric boundary-layer responses, and to infer the SST effect on the bulk formulae used for modelling the air–sea fluxes, which subsequently help drive atmospheric convection. Several papers have suggested that these time-phase relationships could be quantified in both observations and models, and used to better understand the systematic process errors and drifts which are robust across forecast temporal scales (e.g., Shelly *et al.*, 2014). We examine such diagnostics in this study, with special focus on the BoB SST–rainfall relationships on intraseasonal time-scales.

We analyse daily initialized forecasts from the global Met Office Unified Model (MetUM) during June to September (JJAS) 2016, as part of the Bay of Bengal Boundary-Layer Experiment (BoBBLE; Vinayachandran *et al.*, 2018). Three MetUM configurations are used: no air–sea feedbacks (atmosphere-only), coupled to a mixed-layer ocean and coupled to a dynamical ocean. Forecast lead times up to 7 (atmosphere-only) or 15 (air-sea coupled) days are analysed. Forecasts are compared against each other and against observational estimates based on satellite measurements, atmospheric reanalyses or hybrid products (i.e., a combination of the two). This comparison examines monsoon precipitation over India, and SST, surface heat fluxes and related meteorological variables used in the bulk formulae (such as surface winds, air temperature and humidity) over the BoB and wider Indian Ocean.

The paper outline is as follows. A summary of the MetUM configurations, the design of the forecast experiments, and descriptions of verification products and diagnostics used in the analysis, are given in Section 2. Section 3.1 examines spatial patterns of seasonal-mean (JJAS 2016) biases in precipitation, surface wind and SST as a function of forecast lead time. Section 3.2 does the same for forecast surface heat flux components and variables used in the bulk formulae, along with more detailed comparisons in the northern BoB with other observation-based products to document observational uncertainty in the surface heat budget. Section 3.3 investigates intraseasonal variability of precipitation in the forecasts and satellite-based estimates, regionally and averaged over the land of India. It also considers the role of air–sea coupling in the propagation of intraseasonal active and break events from the BoB to land. Section 3.4 investigates lead–lag correlations between northern BoB SST and regional precipitation from forecasts and observational products on intraseasonal time-scales. It also evaluates regional co-variability of surface heat fluxes and SST over the northern BoB from forecasts and other flux products, including the coupling coefficients, on intraseasonal

TABLE 1 Summary of NWP configurations used. All forecasts are initialized daily at 0000 UTC with an atmospheric grid spacing of 0.23° × 0.16° with 85 vertical levels

Name	Ocean model	Ocean grid	SST	Forecast length
ATM	None	—	OSTIA (persistent)	7 days
KPP	GOML2.0	0.23° × 0.16°; 100 vertical levels	Initialized from FOAM analysis	15 days
NEMO	NEMO-CICE	≈0.25°; 75 vertical levels	Initialized from FOAM analysis	15 days

time-scales. Section 4 provides further discussion and Section 5 gives the conclusions.

2 | DATA AND METHODS

2.1 | MetUM forecasts

We analyze NWP performance for the 2016 Indian summer monsoon in forecasts from three configurations of the global Met Office Unified Model (MetUM): two configurations with atmosphere–ocean coupling and one configuration without. For the latter, we analyze forecasts with the operational NWP model (hereafter ATM for atmosphere-only), which in 2016 used the Global Atmosphere 6.1 (GA6.1; Walters *et al.*, 2017) scientific configuration at a horizontal grid spacing of 0.23 ° × 0.16°, with 85 points in the vertical and a model lid at 85 km. ATM forecasts use initialised SST and sea ice from the Operational Sea-surface Temperature and sea Ice Analysis (OSTIA; Donlon *et al.*, 2012) which are fixed during the forecast. ATM forecasts are initialized from Met Office operational atmospheric analyses and run for 7 days.

The two coupled configurations comprise the same GA6.1 atmospheric model from ATM, at the same resolution, coupled to different ocean models. One coupled configuration uses the Nucleus for European Modelling of the Oceans (NEMO) three-dimensional dynamical ocean model, with a tri-polar horizontal (ORCA) grid at ≈0.25° grid spacing with 75 points in the vertical, as well as the Los Alamos sea-ice model (CICE). This configuration is similar to the Global Coupled 2.0 configuration employed in operational seasonal forecasts and in climate simulations (Williams *et al.*, 2015). We refer to this configuration as NEMO. NEMO forecasts are initialized from the same atmospheric analyses as ATM, and oceanic analyses from the Met Office operational Forecasting Ocean Assimilation Model (FOAM; Storkey *et al.*, 2010). NEMO forecasts are run for 15 days.

The other coupled configuration uses the Multi-Column configuration of the *K* Profile Parametrization (MC-KPP) one-dimensional ocean mixed-layer model, which uses the vertical mixing scheme of Large *et al.* (1994). We refer to this configuration as KPP. MC-KPP

does not simulate vertical or horizontal advection, or sea ice; the latter is prescribed from the same OSTIA analysis used in ATM. As there is one MC-KPP column under each atmospheric gridpoint, the oceanic and atmospheric horizontal resolutions are identical. MC-KPP has 100 vertical levels in a 1,000 m water column, on a stretched grid to increase resolution near the surface. MC-KPP and NEMO have comparable vertical spacing near the ocean surface (≈1 m), but MC-KPP has much finer resolution within the typical tropical mixed layer (e.g., 40 points in the top 100 m in MC-KPP, versus 15 points in NEMO). The NEMO and KPP configurations are coupled to the atmosphere via the Ocean Atmosphere Sea Ice Soil (OASIS; Craig *et al.*, 2017) coupler every hour, including fluxes of heat, freshwater and momentum. The coupling is identical between the configurations, except that NEMO simulates the drag of the ocean surface currents on the atmospheric near-surface wind, whereas KPP does not. Further details of the KPP configuration can be found in Hirons *et al.* (2015) and Peatman and Klingaman (2018), although it is important to note that the forecasts here do not use any form of flux correction. KPP forecasts are initialized from the same atmospheric and oceanic analyses as the NEMO forecasts and are run for 15 days.

Table 1 summarizes the MetUM configurations used to generate the forecasts, including the oceanic initial conditions. All forecasts are initialized daily at 0000 UTC. We analyze forecasts valid during 1 June to 30 September 2016, which includes forecasts initialized between 15 May and 30 September 2016. Comparing KPP and ATM shows the effects of thermodynamic air–sea coupling; comparing NEMO and ATM shows the effects of thermodynamic and dynamic coupling (i.e., including the ocean dynamical response); comparing NEMO and KPP shows the effects of ocean dynamics within the coupled system. Further details on these forecasts can be found in Feng *et al.* (2019), which analysed the same three sets of forecasts for West Pacific tropical cyclones.

Surface fluxes are of key importance to this study, especially at the ocean surface since these mainly determine the northern BoB SST in the coupled configurations. Surface heat fluxes consist of short-wave and long-wave radiation terms along with turbulent fluxes for heat (sensible and latent) computed from bulk formulae, with the

surface upward long-wave emission (computed using the Stefan–Boltzmann Law) and the turbulent fluxes sensitive to SST. All model configurations use the same bulk formulae, as the fluxes are computed in the GA6.1 atmospheric model, which is the same in the three sets of forecasts. In ATM, turbulent fluxes are derived from bulk formulae with inputs from the evolving atmospheric state and fixed OSTIA SST. In KPP and NEMO, turbulent fluxes are computed through the same bulk formulae, but with time-varying oceanic states. Air–sea flux differences between ATM and either NEMO or KPP arise from differences in the atmospheric and surface ocean states, including initial SSTs (which are the same in KPP and NEMO, but differ between ATM and the two coupled forecasts; see the fourth column of Table 1), and currents (for NEMO only). In earlier work using the same three sets of forecasts (Feng *et al.*, 2019), we found that differences in forecast SSTs between ATM and either NEMO and KPP are small, with magnitudes generally less than 0.5 K in the seasonal mean. Further details of the surface flux computation can be found in Walters *et al.* (2017) and references therein.

2.2 | Verification products

Gridded outgoing long-wave radiation (OLR) from the National Oceanic and Atmospheric Administration (NOAA) Interpolated dataset (Liebmann and Smith, 1996) is used as a proxy for convection when considering the propagation of intraseasonally varying convection. These data are provided as daily means on a 2.5° grid. Gridded precipitation data from the Tropical Rainfall Measuring Mission (TRMM) 3B42 product version 7 (Huffman *et al.*, 2007), provided at 0.25° grid spacing every 3 hr, are used to verify forecast precipitation. To analyse the lagged covariance between rainfall and SST, we use data from the ECMWF Fifth Reanalysis (ERA5; Hersbach *et al.*, 2019), and from the Global Precipitation Climatology Project (GPCP; Adler *et al.*, 2003) and TRMM 3B42. Near-surface (10 m) zonal and meridional wind components are verified against the ERA5 equivalent neutral wind vector at 10 m above the sea surface.

For comparison with the model surface fluxes and related meteorological variables, we use four widely used observation-based gridded flux products: OAFlux (Yu *et al.*, 2008) for latent and sensible heat fluxes; the Clouds and the Earth's Radiant Energy System (CERES) surface radiation synoptic product (SYN1deg; Rutan *et al.*, 2015); the ERA5 atmospheric reanalysis (Hersbach *et al.*, 2019); and the regional TropFlux product (a combination of ERA-Interim reanalysis (Dee *et al.*, 2011) and remote-sensing data; Kumar *et al.*, 2011). The comparisons are limited to gridded flux products

because *in situ* flux measurements from the Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction (RAMA) array (McPhaden *et al.*, 2009) in the tropical Indian Ocean are not available in the BoB (e.g., the RAMA buoy at 15°N, 90°E) for JJAS 2016. It is hard to determine which gridded surface flux products are the most accurate in the BoB, because no one flux product performs best against *in situ* measurements for all flux component/ bulk input variables (Sanchez-Franks *et al.*, 2018). However, the flux products used here are independent of the NWP models and contribute to assessment of the forecasts, both for the JJAS 2016 mean and on intraseasonal time-scales. The CERES, OAFlux and TropFlux products are available as daily means on a 1° grid, while ERA5 data are provided as hourly means on a 0.25° grid. Quantitative model–observation flux comparisons in the following sections are performed with daily mean data on a common 1° × 1° grid.

2.3 | Analysis methods

Active and break phases of the Indian summer monsoon are defined as in Rajeevan *et al.* (2010), who identify the core monsoon zone (CMZ) in northern India, a region of highly spatially correlated precipitation during July and August. This is preferred to all-India rainfall, since the sign of rainfall anomalies varies across India during active and break phases (Lee *et al.*, 2013). Active or break events are defined as periods when the area-averaged precipitation in the CMZ is greater than 1 or less than −1 standard deviation from the mean for at least three consecutive days. Composites of precipitation anomalies for active and break events are discussed in Section 3.3. The active and break days during JJAS 2016 are found using TRMM 3B42 precipitation; we define anomalies as the difference from the JJAS 2016 mean. Composites from each set of forecasts are averaged over the active and break days found from observations; the forecast anomalies are computed from a lead-time-dependent JJAS 2016 mean.

To investigate the BSISO, northward Hovmöller diagrams (latitude versus time, averaged over 70°–100°E) are produced for a constant forecast lead time by concatenating forecasts such that the validity time is always 1 June to 30 September. For example, for lead time 3 days, the third day of each forecast from forecasts initialized on 30 May to 28 September is extracted; the data from these 122 forecasts are concatenated in time. The time-mean is subtracted to produce an anomaly from the seasonal mean. To highlight the intraseasonal variability, the resulting field is bandpass filtered with cut-off frequencies of 70 and 24 day^{-1} using a Lanczos filter (Duchon, 1979) of order 30 (i.e., 61 weights).

Since the filtering loses 30 days at either end of the time series, the Hovmöller is first padded by tapering to zero over 30 days at each end.

To examine the consistency of the SST–precipitation phase relationships on intraseasonal time-scales, we calculate lead–lag correlations of the SST anomaly averaged over the northern BoB with gridpoint and area-averaged precipitation anomalies over the region (65°–105°E, 0°–30°N). Following Vecchi and Harrison (2002), the northern BoB is defined as ocean grid points north of 15°N. To focus on the intraseasonal variability, a 7-day running mean is applied to daily mean data over JJAS 2016; daily SST and precipitation anomalies are then computed as the departure from their 50-day centred means. Correlations are calculated from 16 days before (SST leading rainfall) and 16 days after (rainfall leading SST) the JJAS 2016 period. This comparison examines both maps (spatial patterns) from forecasts at 7 days' lead time and regionally averaged results over India (15°–25°N, 75°–90°E) and the northern BoB domain (15°–22°N, 80°–100°E), as a function of forecast lead time. Correlations are plotted only where they are significantly non-zero at the 95% confidence level according to a Student's t-test.

3 | RESULTS

3.1 | Large-scale mean biases

In this section we consider the JJAS 2016 mean forecast biases in the three MetUM configurations. The observed mean precipitation and forecast biases are shown in Figure 1. Biases are shown for averages of several forecast lead times. For example, days 1&2 means all lead times from $T + 0$ to $T + 48$ hr, where T is the initialization time at midnight UTC. The pattern correlation r between observations and each forecast plot is given above each panel, as is the root-mean-square error (RMSE) averaged over the domain shown, measured in mm·day^{-1}. r and RMSE are also shown for all lead times in Figure 1b.

The observed spatial distribution for JJAS from TRMM 3B42 (Figure 1a) shows that 2016 was broadly a typical year, with India's heaviest rain being over the Western Ghats, followed by the northeast region. Over the BoB there was a clear wet–dry gradient from the northeast to the southwest. Over the rest of the Indian Ocean, the strongest rainfall was in the east of the basin, just south of the Equator. In the long-term mean of climate-length simulations, the MetUM typically has JJAS dry biases over India and the Maritime Continent (MC), and wet biases over the Indian Ocean and west tropical Pacific (e.g., Williams *et al.*, 2015; Johnson *et al.*, 2016; Walters *et al.*, 2017; Peatman and Klingaman, 2018). In these 2016 ATM

forecasts (Figures 1c,d), such biases are seen over the MC, and the Indian and Pacific Oceans. However, over India there is a wet bias, especially in the east of the country. A dry bias just off the west coast, over the Arabian Sea, grows stronger with lead time; a further dry bias over northeast India emerges by days 3&4 (not shown), strengthening by days 5&6. Over the BoB there is a wet bias at early lead times, especially in the west (days 1&2), with a dry bias emerging in the east and spreading westwards until it covers most of the Bay (days 5&6).

Figure 1e,f show the effect on mean precipitation of introducing air–sea coupling by comparing the ATM and KPP forecasts, and Figure 1g–i show the effect of introducing ocean dynamics by comparing KPP and NEMO. During the first 2 days of the forecasts, air–sea interactions have only a small effect on seasonal-mean precipitation, with negligible differences in r and RMSE. Differences in mean precipitation over ocean increase with lead time, with air–sea coupling generally improving forecasts. For example, over the Indian Ocean just south of India and just south of the Equator, the wet bias is reduced; and in the east of the ocean basin, off the west coast of Sumatra, the dry bias is reduced. In the west Pacific Ocean, air–sea coupling also improves forecasts, with reduced wet biases north of about 15°N and immediately east of New Guinea, and reduced dry biases between about 0° and 10°N.

Over the BoB, ATM forecasts have a more uniform distribution of rainfall than observed, with a dry bias in the northeast and a wet bias in the southwest. The KPP forecasts again show an improvement in this region, partially restoring the northeast to southwest gradient. However, there is no substantial impact of air–sea interactions on the precipitation bias over land, despite the changes over the BoB and the fact that in boreal summer convection propagates from the BoB to the land as part of the BSISO.

Ocean dynamics alone have a minor effect on precipitation over the first week of the forecasts (Figure 1g–i), with substantial differences not emerging until the second week. These are mainly over the Pacific and the Southern Hemisphere Indian Ocean. As with the impact of introducing air–sea coupling, the effect is almost entirely limited to over ocean, with minimal impact over land, even over areas where convection propagates from ocean to land.

In summary, air–sea coupling slightly reduces errors in forecast seasonal mean precipitation. Forecasts worsen monotonically with lead time according to both metrics r and RMSE.

Figure 2 shows the JJAS OSTIA SST and ERA5 10 m wind, and the forecast biases. ATM forecasts have a very small SST bias ($<\pm0.1$ °C everywhere with RMSE $= 0.01$ °C) relative to OSTIA in days 1&2 because the OSTIA SST is prescribed, but biases grow with time because the SST is persistent, with a warm bias of up
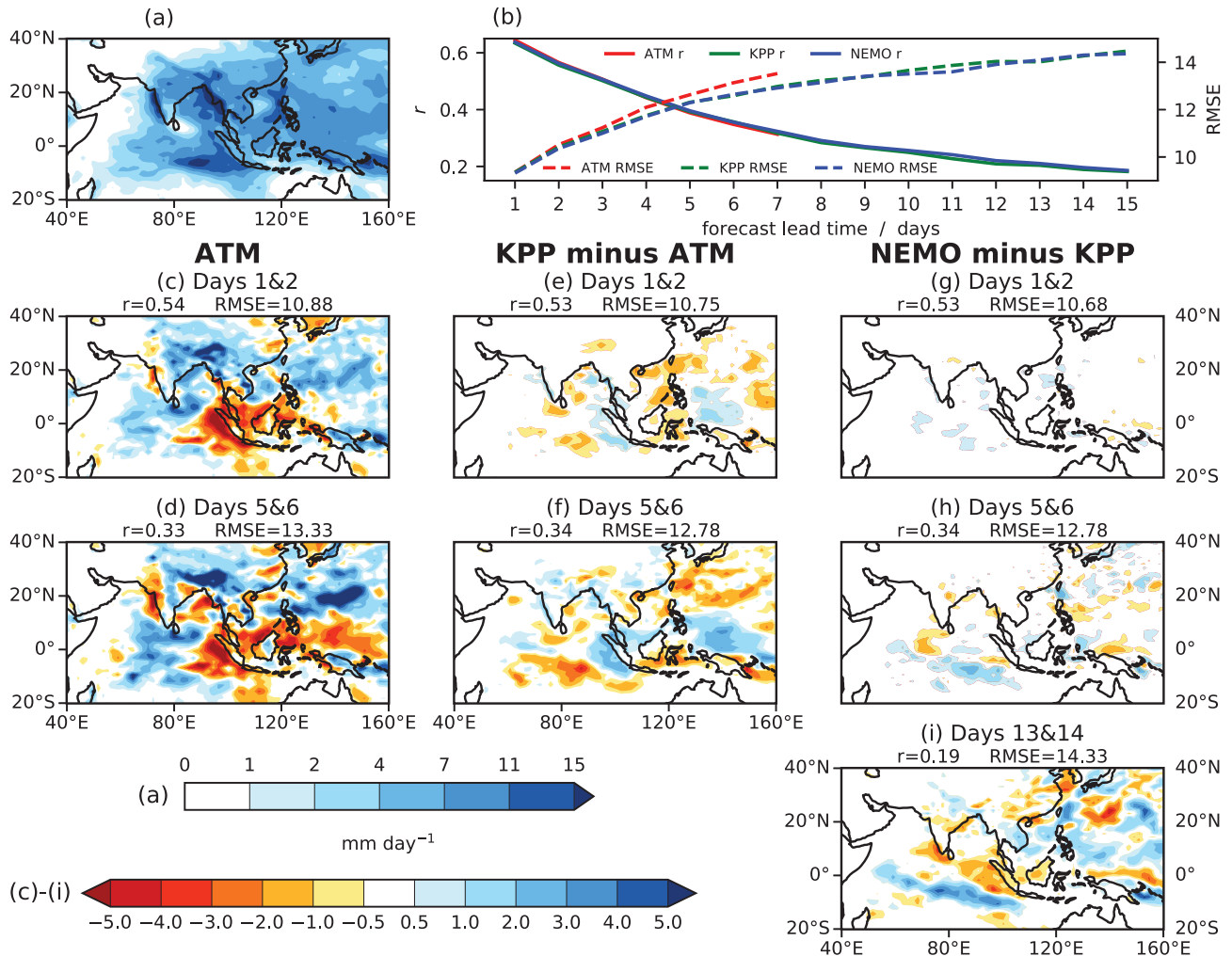
FIGURE 1 (a) JJAS 2016 mean precipitation rate from TRMM 3B42. (b) Pattern correlation r and domain-wide root-mean-square error (RMSE) of each set of forecasts against the observations in (a), as a function of forecast lead time. (c, d) Mean precipitation bias in ATM forecasts, averaged over days (c) 1&2 and (d) 5&6 of the forecast. (e, f) Difference between KPP and ATM forecasts, averaged over days (e) 1&2 and (f) 5&6. (g, h, i) Difference between NEMO and KPP forecasts, averaged over days (g) 1&2, (h) 5&6 and (i) 13&14. r and RMSE are also given above each panel, again computed for the model's full (i.e., not difference) field against observations, as in (b)

to 0.2 °C by days 5&6 over much of the Indian Ocean. By contrast, NEMO and KPP show a cold bias of around 0.3–0.5 °C over most of the Indian Ocean in days 5&6, although there are warm biases in some of the shallow seas of the MC. The RMSE across the domain is 0.25 °C, compared with only 0.09 °C for the ATM forecasts.

The Indian Ocean bias for KPP is similar to that seen in climate-length integrations with the MetUM–KPP coupled model (Hirons *et al.*, 2015; Peatman and Klingaman, 2018). The strongest biases in the KPP forecasts tend to be corrected in the NEMO forecasts even at very short lead times, although in some regions they are over-corrected and in others a bias is introduced where one did not exist for KPP. The result is that domain-wide RMSE, which is worse for coupled models than ATM, does not depend substantially on ocean dynamics. The pattern correlation is almost always above 0.98 for all three sets of forecasts,

indicating that SST biases vary only in their amplitude, with little change in spatial structure.

Ocean coupling has very little effect on wind biases and, where an effect is seen, it is almost exclusively over ocean.

3.2 | Biases due to errors in the surface heat fluxes

In this section we compare the forecast surface heat fluxes and related meteorological variables with independent observational estimates over JJAS 2016. We focus on comparisons with the CERES SYN1deg surface radiation product (Rutan *et al.*, 2015) combined with the OAFlux turbulent flux product (Yu *et al.*, 2008). The CERES SYN1deg surface product (hereafter referred to as CERES)
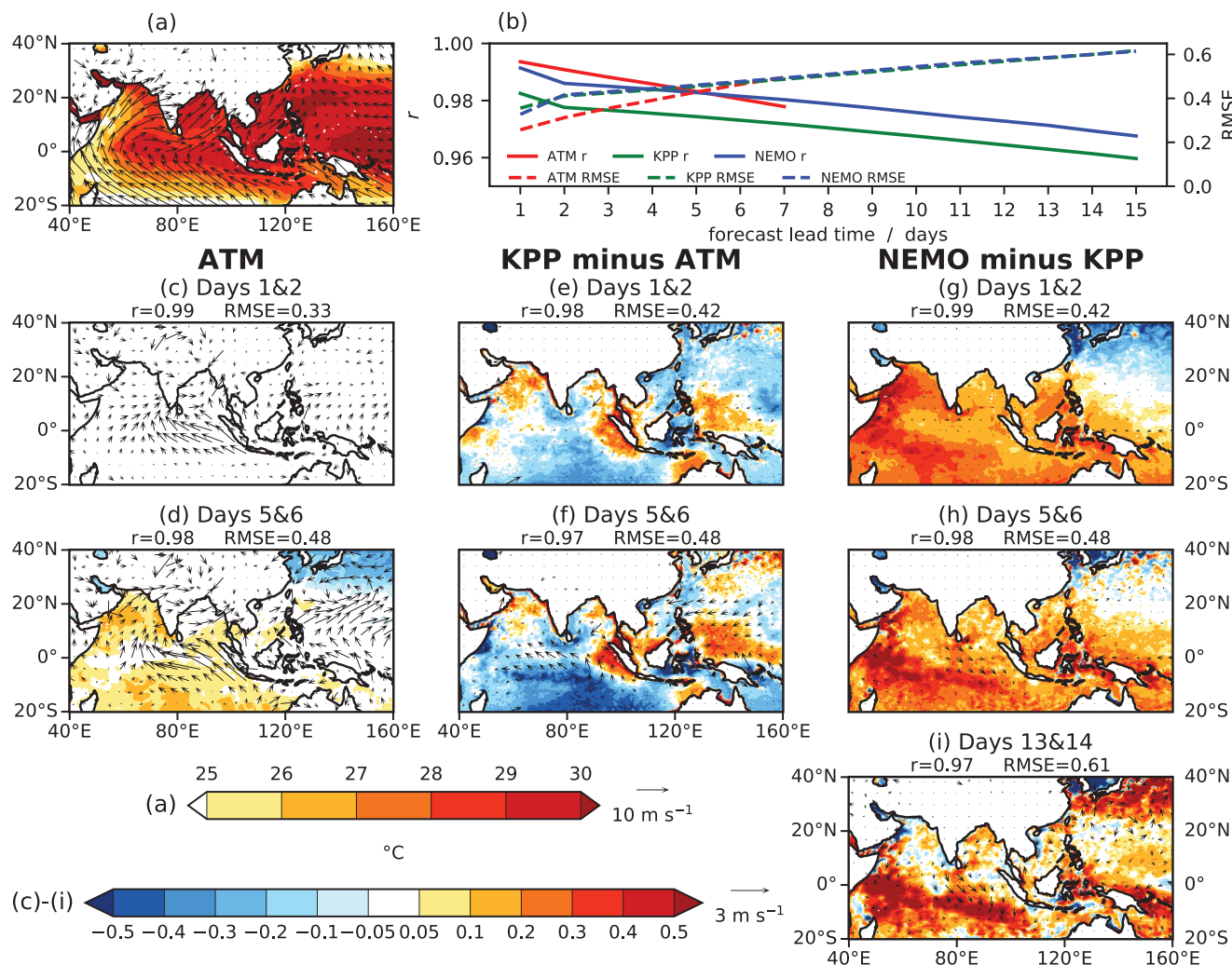
FIGURE 2   As Figure 1, but for SST from OSTIA and 10 m wind from ERA5. *r* and RMSE are for SST. RMSE excludes the lowest percentile and highest percentile of the bias, to remove the effect of very large biases in coastally tiled grid boxes. *r* and RMSE above each panel are computed for the model's full (i.e., not difference) field against observations, as in (b)

is derived from multiple satellite measurements using a radiative transfer model. OAFlux is derived from objective synthesis of satellite observations and atmospheric reanalyses using the COARE bulk flux algorithm (Fairall *et al.*, 2003). Among the gridded flux products in Section 2.3, CERES and OAFLux are the products that have been most completely evaluated against *in situ* flux measurements (Yu *et al.*, 2008; Rutan *et al.*, 2015).

Figure 3 shows mean (JJAS 2016) differences in surface heat flux components between ATM and CERES/OAFlux fluxes in the northeast Indian Ocean sector, along with differences in surface wind speeds and sea–air specific humidity gradients upon which the turbulent heat fluxes depend. These difference maps indicate the short-term biases that evolve in the MetUM at lead times up to 7 days. The surface heat balance (Q_{net}) in most areas is largely determined by short-wave radiation (Q_{sw}) and latent heat flux (Q_{lat}), as the net long-wave radiation

(Q_{lw}) is relatively constant across the BoB (except in the southern bay), and the sensible heat flux (Q_{sen}) is very small (except around the monsoon upwelling areas off Sri Lanka). ATM SST is prescribed and identical to OAFlux (not shown).

Figure 3a–c show the development of positive net radiation ($Q_{sw} + Q_{lw}$) biases, starting off the east coast of India and gradually spreading across the BoB, with the largest errors (>40 W·m^{-2}) over the northwest BoB (10°–20°N) and in the eastern Indian Ocean south of 10°N. These differences in net radiative fluxes are related to the developing dry bias across the region (Figure 1) and the associated underestimation of cloud cover. Q_{sw} is much too strong compared to satellite radiation, whereas errors in Q_{lw} are minimal (not shown); the radiative flux errors are almost entirely due to Q_{sw}.

There are also significant errors in turbulent ($Q_{lat} + Q_{sen}$) heat fluxes (Figure 3d–f), with biases typically

ATM minus CERES/OAFlux
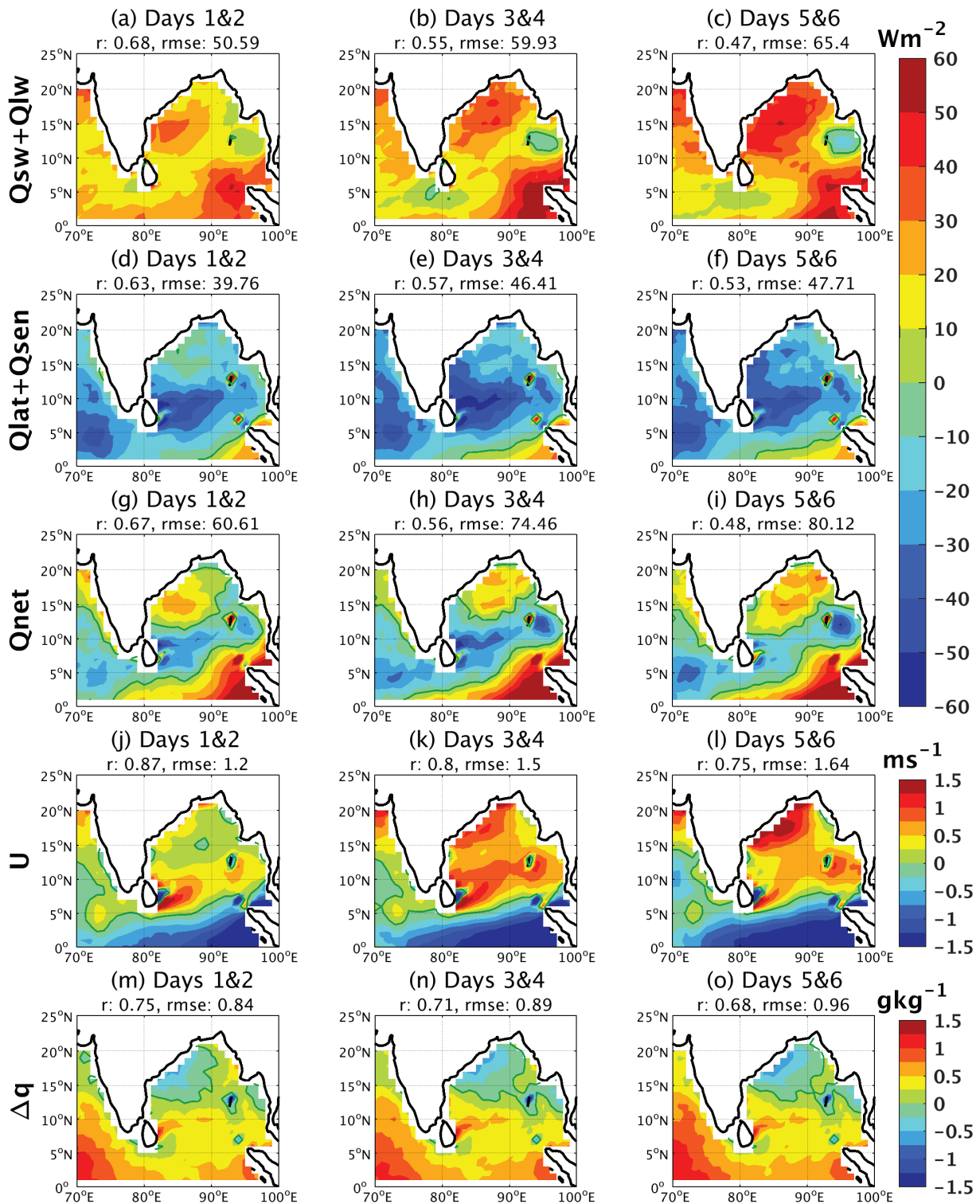
FIGURE 3 JJAS 2016 mean bias in ATM forecasts of: (a)–(c) net radiative (short-wave plus long-wave) heat flux; (d)–(f) net turbulent (latent plus sensible) heat flux; (g)–(i) net surface heat flux; (j)–(l) 10 m wind speed; and (m)–(o) sea–air specific humidity gradient. Biases are relative to CERES radiation and OAFlux turbulent fluxes in the northeast Indian ocean sector, and are shown for days 1&2, 3&4 and 5&6 of the forecasts. Net surface heat flux is positive into the ocean. *r* and RMSE given above each panel are calculated against CERES or OAFlux data

$\approx$25–40 W·m^{-2} of greater surface cooling in most regions that are matched by errors in bulk variables (Figure 3j–o). Surface wind speeds are too strong (Figure 3j–l), beginning southeast of Sri Lanka and gradually expanding north with lead time into the BoB, as compared to OAFlux, which uses a combination of SSM/I and QuikSCAT retrievals. Over the equatorial Indian Ocean (south of 5°N), the surface wind speeds are underestimated by more than 1.5 m·s^{-1}, always dominated by the too-weak easterly trade winds (not shown).

Negative differences in sea–air humidity gradients over most of the north BoB relative to OAFlux (a combination of satellite water vapour retrievals from SSM/I and atmospheric reanalysis) indicate a wet bias in the atmospheric boundary layer, while positive differences elsewhere indicate dry biases (Figure 3m–o). This suggests an anti-correlation between specific humidity errors and wind errors in most regions. The patterns of biases in sea–air temperature gradients (affecting both Q_{sen} and Q_{lat}) resemble the specific humidity difference fields, with most areas showing opposite-signed biases in temperature gradients and surface winds (not shown). This demonstrates that the overestimated Q_{lat} in ATM is primarily due to too-strong near-surface wind speeds, not boundary-layer humidity or temperature biases. Turbulent heat fluxes within the BoB are also sensitive to the choice of bulk parametrization schemes (e.g., Mallick *et al.*, 2020). OAFlux and ATM use different bulk formulae, so differences in the turbulent flux algorithm may also contribute to the differences in turbulent fluxes shown here.

Turning to differences in net surface heat flux ($Q_{net} = Q_{sw} + Q_{lw} + Q_{lat} + Q_{sen}$) (Figure 3g–i), there is a positive bias (excess heat into the ocean) at early lead times throughout the northwestern BoB that strengthens with lead time, always dominated by Q_{sw} errors. There is also a significant negative bias contributed by excess turbulent heat fluxes ($\approx$25 W·m^{-2}), particularly in the western BoB and over the Arabian Sea. The large positive Q_{net} errors at the southeastern latitudes (extending up to 0°–5°N) seem to be connected with underestimated easterly trade winds (and the resultant turbulent fluxes).

The domain-average RMSE for Q_{net} increases from 61 W·m^{-2} on days 1&2 to 80 W·m^{-2} on days 5&6, with pattern correlations with CERES/OAFlux of 0.67 and 0.48, respectively. This reflects the degradation of both negative biases in $Q_{lat} + Q_{sen}$ and positive ones in $Q_{sw} + Q_{lw}$ over this region with lead time. The RMSE results contrast with the bias maps in Figure 3g–i which show considerable compensation of net flux error that can be hidden when regionally averaging. Finally, pattern correlations for variables used in the bulk formulae, that is, wind speeds and humidity gradients, are larger than for $Q_{lat} + Q_{sen}$; the latter correlation may be degraded due to a mismatch in

the transfer coefficients used in OAFlux (COARE3.0) and ATM (MetUM) turbulent flux algorithms.

We now examine differences among the forecast models for the same quantities (except the specific humidity gradients), along with SST differences that affect Q_{net} in the coupled configurations. Figure 4a–e show the difference between KPP and CERES/OAFlux averaged over lead times 5–6 days. Figure 4f–j shows the KPP difference from ATM averaged over the same lead times; Figure 4k–o shows the equivalent difference with NEMO.

The KPP forecasts are more consistent with the CERES/OAFlux combination than the ATM forecasts, with smaller positive biases for Q_{sw} and Q_{net} throughout the BoB, and smaller Q_{lat} biases in the western BoB and the Arabian Sea. NEMO differences from CERES/OAFlux are very similar and are not shown here. The air–sea coupling at these forecast scales (up to one week lead time) appears to compensate for the excess of heat due to underestimated cloud effects (and related Q_{sw} errors), because the coupled models develop SST biases that, through their influence on surface fluxes, effectively transport heat from the BoB to south of India, where Q_{net} is highest in CERES/OAFlux (not shown).

KPP and NEMO show local-scale SST differences relative to OAFlux and the fixed daily SST used in ATM (Figure 4e,j,o), particularly in the southeast corner in the Andaman Sea and downstream of Sri Lanka (up to $\pm$0.6°C). The spatial distribution of SST differences between KPP and ATM resembles that of the net flux difference fields (Figure 4h,j), with a significant pattern correlation ($r \approx -0.8$), consistent with surface fluxes dominating local heat budgets nearly everywhere. Thus we can interpret SST differences in terms of surface flux changes, rather than due to changes in near-surface vertical mixing in KPP. However the correspondence is not perfect; for example, there is an indication of excessive downward mixing from the surface layers in the coupled systems in the areas off the east Indian coast and north Sri Lanka, cooling the ocean surface and modifying the mean turbulent fluxes there (Figure 4g,l). This is consistent with higher surface winds in the coupled forecasts (Figure 4i,n), especially in KPP. There is also a clear dipole pattern in the surface wind and turbulent and net flux differences between KPP and NEMO, particularly near strong currents off Sri Lanka (Figure 4l–n), reflecting in this case contributions from ocean dynamics processes which are represented in NEMO (horizontal and vertical heat transports), but not in KPP. Apart from these minor regional differences, the NEMO and KPP results are highly consistent, with most areas showing flux differences of up to $\pm$10 W·m^{-2}.

Figure 5 shows a more detailed comparison between the full range of forecasts and CERES/OAFlux averaged
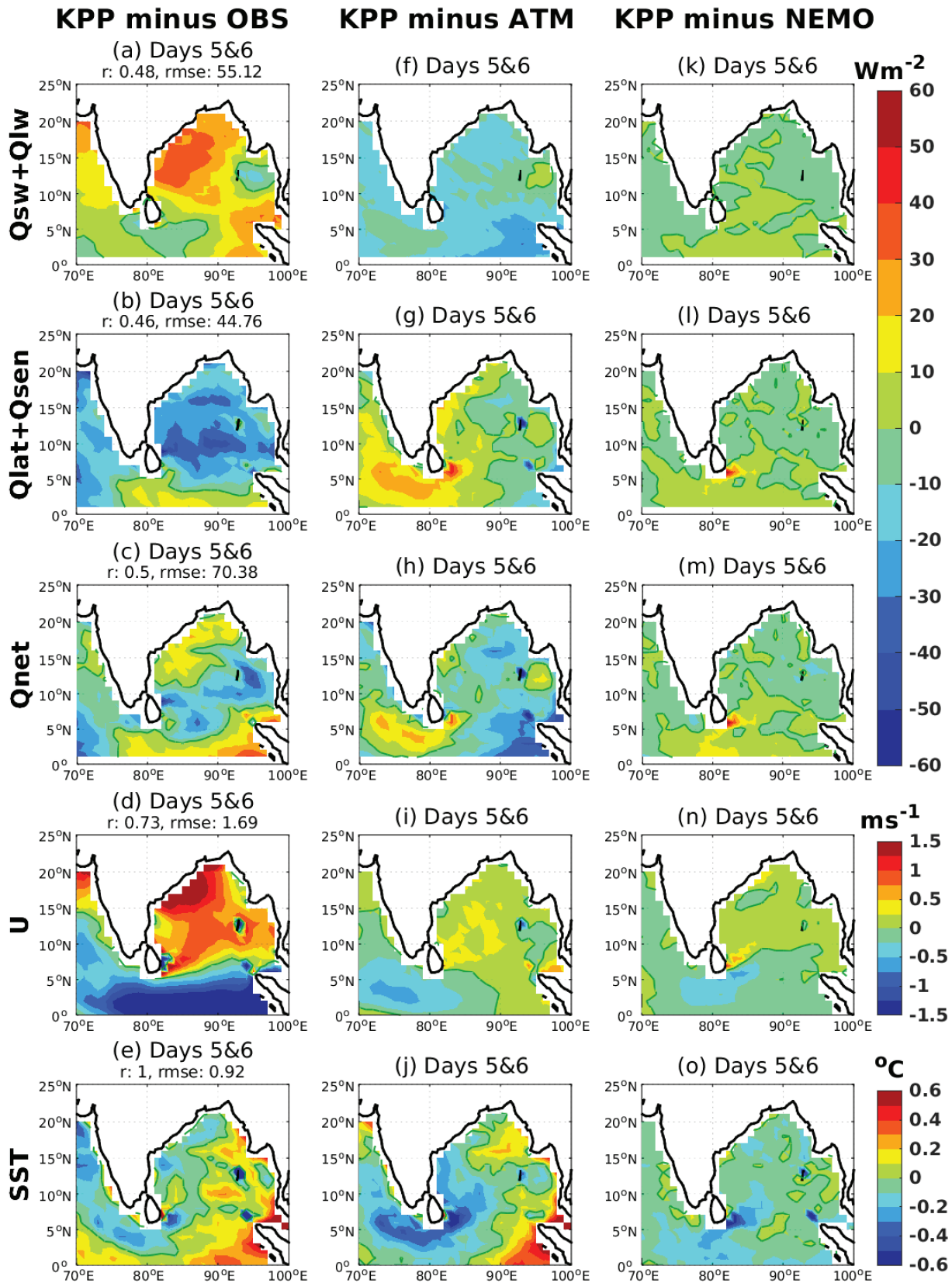
FIGURE 4 As Figure 3, but for: (a)–(e) KPP minus CERES/OAFlux, (f)–(j) KPP minus ATM and (k)–(o) KPP minus NEMO. All differences shown are averaged over days 5&6 of the forecasts. *r* and RMSE given above (a)–(e) are calculated against CERES or OAFlux data

over the northern BoB (north of 15°N) during JJAS 2016. Also shown are the standard errors for daily variations in CERES/OAFlux data over the season (red horizontal lines), along with the seasonal-mean spreads derived from the ensemble of observational flux products in Section 2.3 (dashed grey horizontal lines) to illustrate how the forecasts compare to inter- and intra-product variations.

All forecasts show season-mean flux biases of similar sign, increasing with lead time over the first week of the forecasts. Mean offsets (forecast–CERES/OAFlux) are estimated to be +10–40 W·m^{-2} for Q_{sw} and −15 to −25 W·m^{-2} for Q_{lat}, always well outside the error bounds of CERES/OAFlux over the season ($\approx$4–7 W·m^{-2}). Q_{lw} and Q_{sen} biases are much smaller (<+1–3 W·m^{-2}) within the observational uncertainty. Among the forecasts, ATM shows larger Q_{sw} errors relative to the coupled KPP and NEMO and a positive bias in Q_{net} into the ocean, while KPP and NEMO show a negative bias in Q_{net} during the first week and a positive bias later. There are compensating errors between flux components so that the range in Q_{net} errors in the north BoB is typically within 10 W·m^{-2} of the CERES/OAFlux values, as shown in Figure 5e. The dominant effects are anticorrelated Q_{sw} and Q_{lat} errors and, to a lesser degree, anticorrelated Q_{lat} and Q_{sen} errors.

Turning to the bulk variables, SST biases in the coupled configurations grow with lead time (up to +0.27 °C in NEMO), particularly after day 6 once the Q_{net} bias becomes positive, but these warm biases in the north BoB are considerably smaller than the range of observational estimates ($\approx$0.7 °C). The air temperature and specific humidity biases are also positive, indicating a increasingly wetter (+0.1–0.6 g·kg^{-1}) and warmer (+0.2–0.7°C) atmospheric boundary layer in MetUM relative to OAFlux. This implies weaker sea–air temperature and humidity contrasts in the forecasts relative to OAFlux, but these do not contribute to additional bulk sensible or latent cooling in the models. In contrast, the forecast surface wind speeds are systematically overestimated (+0.25–1.2 m·s^{-1}) and dominate the Q_{lat} biases for all lead times. It is also seen that the coupled configurations show nonlinear evolution of the biases (e.g., decreasing wind speed differences) at longer lead times, likely reflecting the model drift toward a weaker monsoon circulation and lower precipitation during the second week of lead time.

Finally, it is worth noting that the observational range in most flux components and bulk variables (except the surface winds) is quite large, even after temporal and spatial averaging, thus making it difficult to unambiguously evaluate the surface heat balance in the north BoB, where the Q_{net} spread (inter-product standard deviation) reaches $\approx$27 W·m^{-2} on average, consistent with the findings of Sanchez-Franks et al. (2018) and other regional studies.

3.3 | Intraseasonal variability

We now consider the representation of intraseasonal active and break monsoon events in the forecasts. In Figures 6a–c, daily precipitation is averaged over the CMZ (green region in Figure 7), where the seasonal-mean bias was positive on days 1&2 (Figure 1c), but negative by days 5&6 (Figure 1d). The time series is plotted for TRMM 3B42 (thick black line) and the forecasts (coloured lines, for selected lead times). The observations show many local maxima, with major active rainfall events around the start of July (the monsoon onset over the CMZ), mid-July, the start of August and, to a lesser extent, the end of August and end of September.

The wet bias in Figure 1c first emerges clearly in the second half of June, when the forecasts transition to active conditions a couple of weeks before the observed sudden peak at the start of July. This early onset occurs in all three configurations, at all lead times. Thereafter in the season, the timing of the variability tends to be accurate but the rainfall amount shows biases. An exception to the accuracy in the timing of rainfall events is in the first half of August, when the forecasts tend to persist wet conditions for too long at lead times of less than 1 week. In general, rainfall rates are lower at longer lead times. Beyond about 1 week lead time, the coupled forecasts show dry biases for most of the season.

Again, coupling has little impact. Forecasts of the timing and amplitude of rainfall events worsen with lead time, as seen from the correlation and RMSE compared with observations, shown in Figure 6d.

The spatial pattern of rainfall within active events is shown in Figure 7. The definition of an active event, following Rajeevan et al. (2010), is described in Section 2.3. These active events approximately correspond to phases 5–6 of BSISO1 in Lee et al. (2013). The green region in each panel is the CMZ. TRMM 3B42 observations (Figure 7a) show that active events are associated with an equally strong precipitation anomaly over the northern BoB as over the central CMZ (over 10 mm·day^{-1}), and a strong negative anomaly of similar magnitude in the southern BoB and western MC. The forecast plots (Figure 7c–i) for each experiment and lead time show the bias in the precipitation anomaly for active days. That is, they show the forecast anomaly minus the observed anomaly, where the anomalies are calculated relative to the forecast and observed climatologies, respectively. Stippling indicates that the forecast anomaly has the wrong sign.

In ATM forecasts, the anomalies are of the correct sign (no stippling) over most areas, including almost all of India and the BoB. However, the CMZ has distinct regimes, with active days enhancing precipitation too much in the
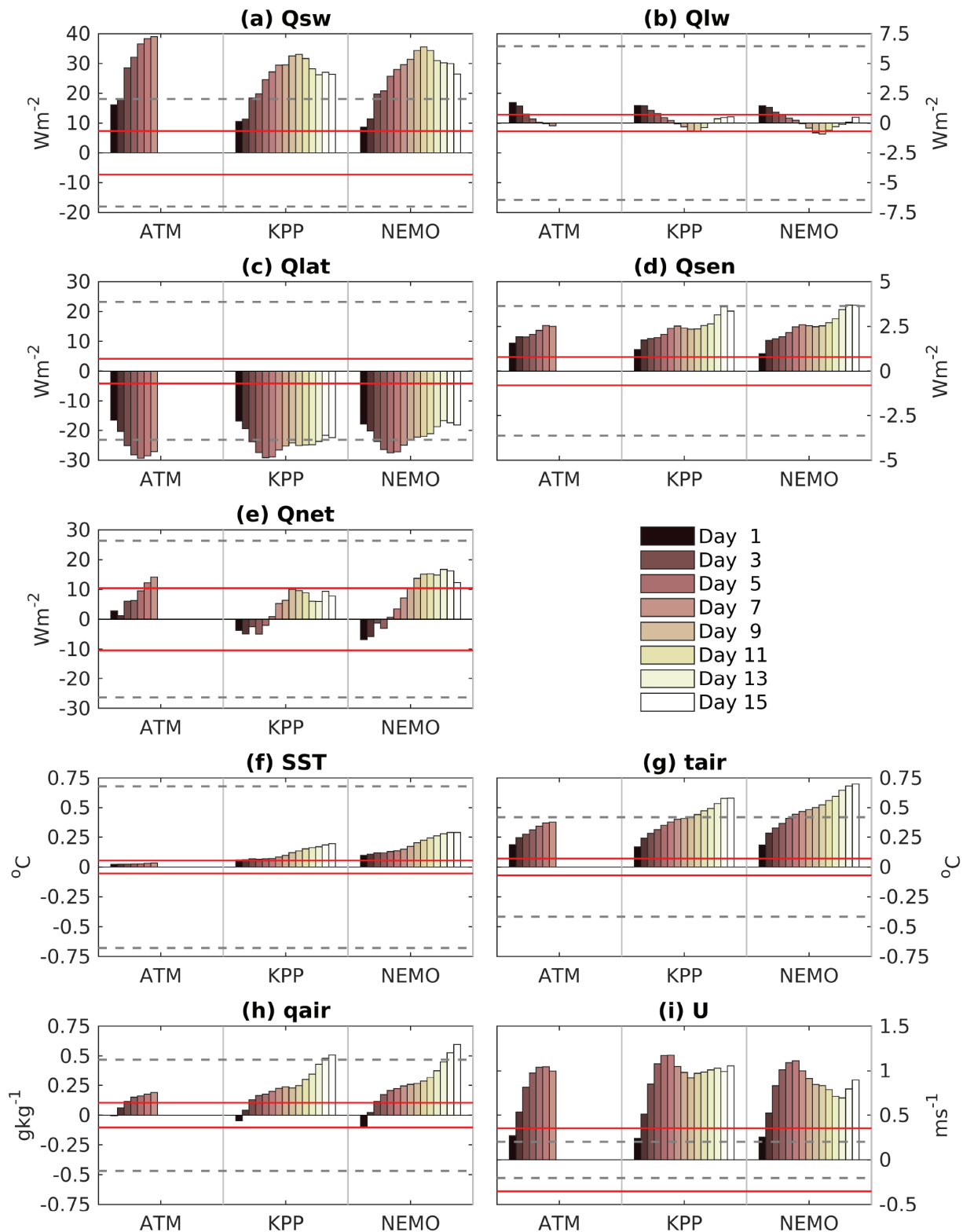
FIGURE 5 JJAS 2016 mean differences from CERES radiation combined with OAFlux flux product, (i.e., forecast − CERES/OAFlux) for (a)–(e) surface heat flux components and (f)–(i) surface variables used in the bulk formulae to calculate the turbulent heat fluxes. All differences are averaged over the northern BoB (i.e. north of 15°N). Heat fluxes are positive into the ocean. Red horizontal lines indicate standard error of daily variations in CERES or OAFlux data, representing mainly the intraseasonal variability throughout JJAS 2016. Dashed grey lines indicate spread (inter-product standard deviation) for observation-based products (CERES, OAFlux, ERA5 and TropFlux) over the 2016 JJAS season: $18.0\,\mathrm{W\cdot m^{-2}}$ for Q_{sw}; $6.4\,\mathrm{W\cdot m^{-2}}$ for Q_{lw}; $23.1\,\mathrm{W\cdot m^{-2}}$ for Q_{lat}; $3.6\,\mathrm{W\cdot m^{-2}}$ for Q_{sen}; $26.4\,\mathrm{W\cdot m^{-2}}$ for Q_{net}; $0.67\,^\circ\mathrm{C}$ for SST; $0.42\,^\circ\mathrm{C}$ for t_{air}; $0.47\,\mathrm{g\cdot kg^{-1}}$ for q_{air}; and $0.20\,\mathrm{m\cdot s^{-1}}$ for U
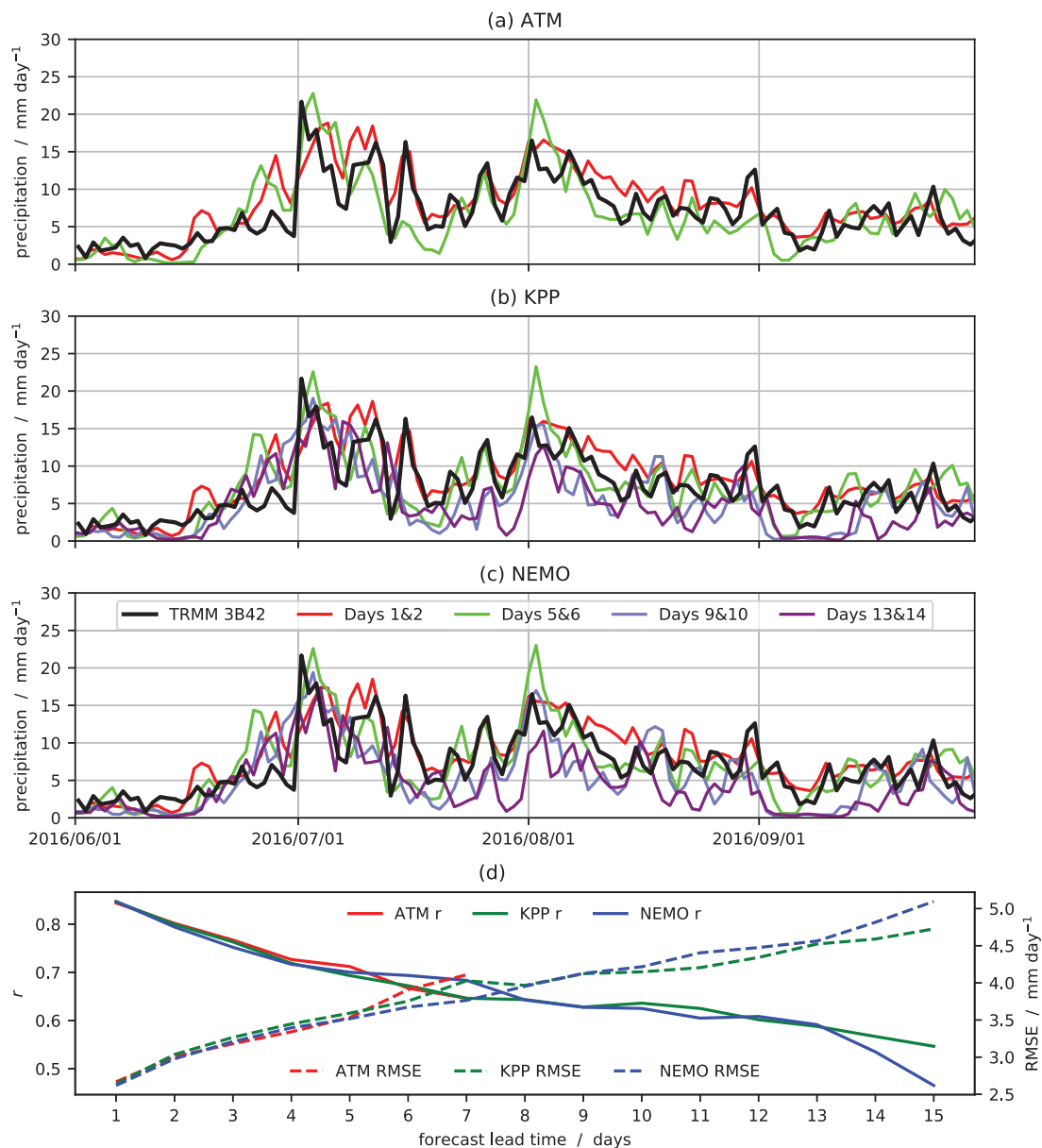
FIGURE 6 (a) Daily precipitation averaged over the core monsoon zone (CMZ) defined by Rajeevan *et al.* (2010 – the green region in Figure 7), for TRMM 3B42 (black line) and ATM forecasts (coloured lines, by forecast lead time). (b, c) are as (a) but for (b) KPP-coupled, and (c) NEMO-coupled forecasts. (d) Correlation *r* and root-mean-square error (RMSE) against observations for each set of forecasts as a function of forecast lead time

east and too little in the west. This is consistent with the season-mean bias in Figure 1c. Furthermore, precipitation is not suppressed enough in the southeast BoB so the meridional gradient in BoB precipitation is smaller than observed. The impacts of ocean coupling and dynamics are again very minimal over land, although precipitation is reduced around the coasts in the BoB. However, even the latter change is small in the domain-wide average. Figure 7b shows there is almost no difference between ATM, KPP and NEMO for these composites. Forecast performance deteriorates rapidly over the first 5 days of the forecasts, with minimal change over the following week.

Results for break events (not shown) are very similar to those for active events, but with the signs of biases generally reversed.

We now consider the northward propagation of convection, using northward Hovmöller diagrams of bandpass-filtered OLR, as described in Section 2.3. NOAA Interpolated OLR, on a 2.5° grid, was used to diagnose the propagation of observed active and suppressed events in Figure 8a. Active (negative OLR anomaly) events are highlighted with solid green lines and suppressed (positive OLR anomaly) with dashed green lines, each at 6 W·m⁻². The same green contours indicating observed events
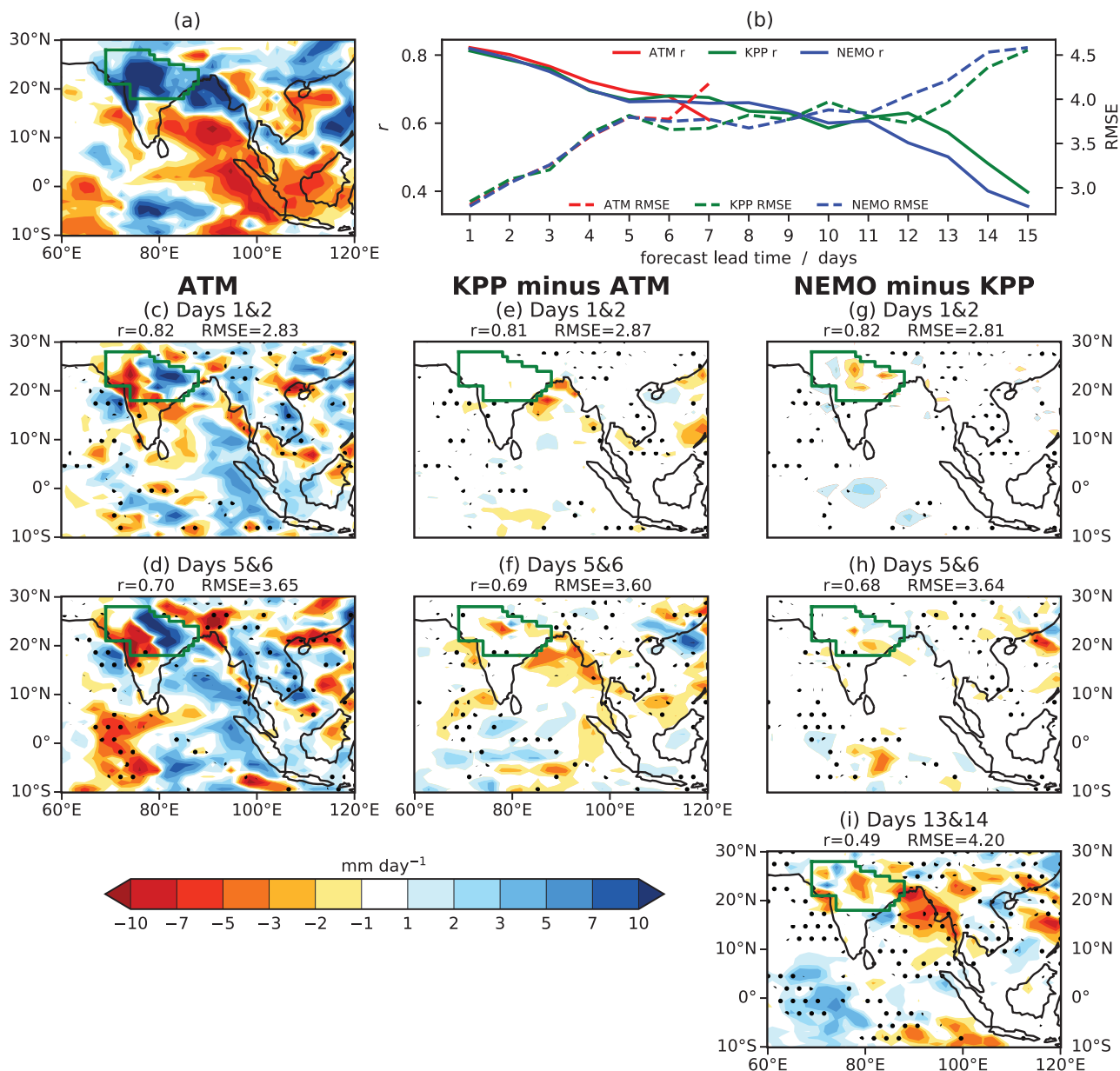
FIGURE 7 (a) Precipitation anomaly composited over 2016 active monsoon days (see main text for details), relative to June to September 2016 mean, for TRMM 3B42. (b) Pattern correlation r and domain-wide root-mean-square error (RMSE) of each set of forecasts against the observations in (a), as a function of forecast lead time. (c)–(i) Biases in active phase precipitation anomalies, with forecast anomalies calculated relative to the forecasts' own climatologies. Forecasts and lead times are as in Figure 1. Stippling indicates that the forecast anomaly has the wrong sign. The green line indicates the core monsoon zone (CMZ) defined by Rajeevan *et al.* (2010)

are overlaid on the forecast bias plots for comparison. There were three major active events during the analysis period.

In ATM, on day 2 (Figure 8c), positive biases tend to precede suppressed events in the Southern Hemisphere (e.g., around 15–25 June and 20–30 August) but follow them in the Northern Hemisphere around India and the BoB (e.g., the first half of July and second half of August). Similar patterns exist for the active events. The convective anomalies propagate too slowly northward, such that

active convection starts too early in the south and ends too late in the north. This too-slow propagation becomes slower still at later lead times (Figure 8d). This is partially consistent with the result in Figure 6a for early August, when precipitation remained enhanced over the CMZ (north India) longer than in observations. However, over the season in general this was not seen in precipitation in Figure 6. The slow propagation is also consistent with the climate-length simulations in Peatman and Klingaman (2018), in which the propagation was generally
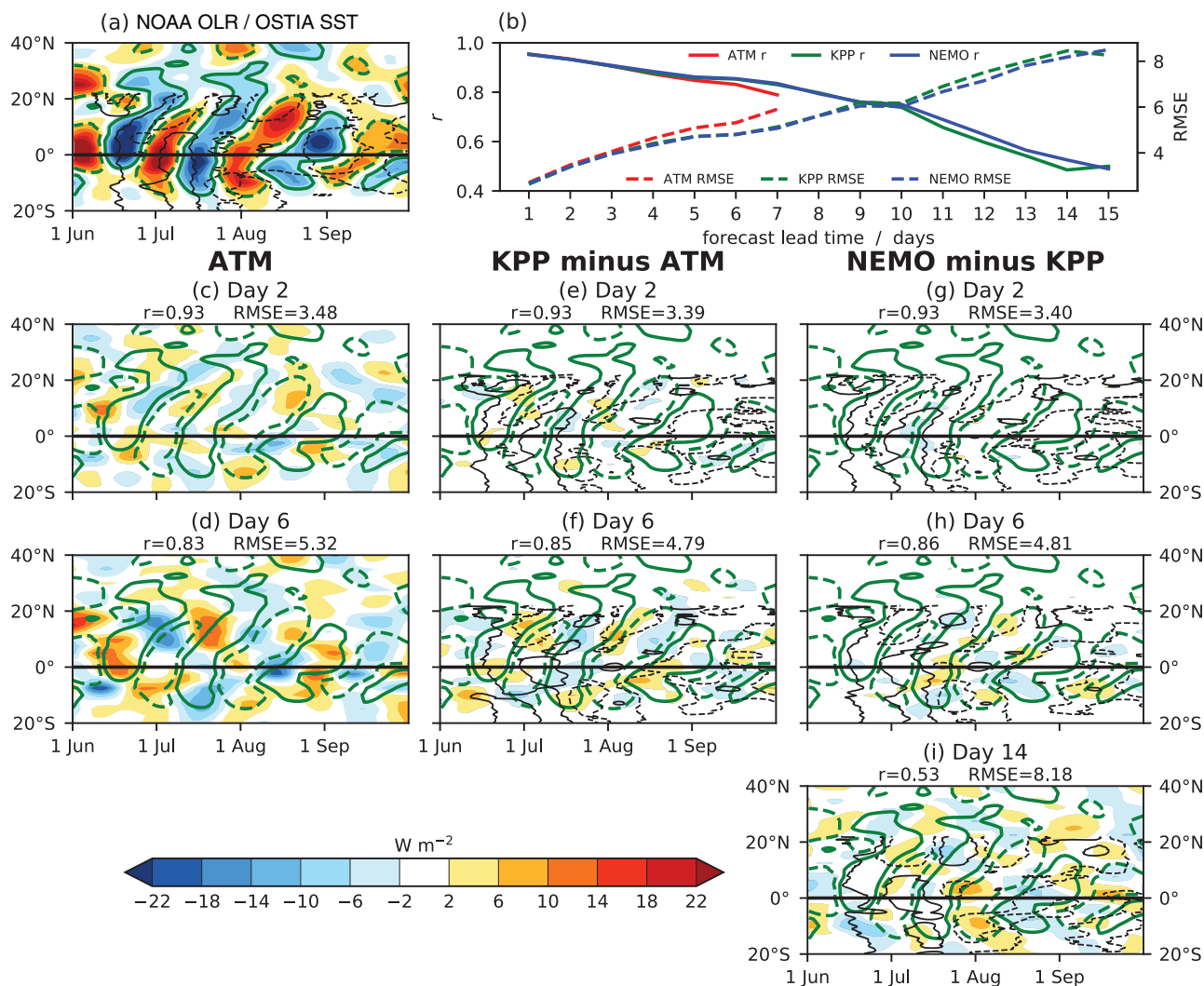
FIGURE 8 (a) Northward Hovmöller diagrams, averaged over 70°–100°E, of 24–70-day bandpass filtered OLR, showing northward propagation of active (blue) and suppressed (red) 2016 BSISO events. (b) Pattern correlation r and domain-wide root-mean-square error (RMSE) of each set of forecasts against the observations in (a), as a function of forecast lead time. (c, d) Bias in ATM forecasts, for days (c) 2 and (d) 6 of the forecast. (e, f) Difference between KPP and ATM forecasts, for days (e) 2 and (f) 6. (g–i) Difference between NEMO and KPP forecasts, for days (h) 2, (h) 6 and (i) 14. Observed active (suppressed) events are highlighted by the solid $-6\,\mathrm{W\cdot m^{-2}}$ (dashed $+6\,\mathrm{W\cdot m^{-2}}$) contours, overlaid on each panel. Black contours show filtered SST anomalies of $+0.1$ (solid) and $-0.1\,°\mathrm{C}$ (dashed); for coupled models these are shown *per se*, not as differences from observations or from other experiments

too slow in both KPP-coupled and atmosphere-only configurations.

As with the other diagnostics which have already been presented, the differences between each set of forecasts are smaller than the magnitude of the biases in ATM (Figures 8e–i). However, the coupling generally improves the timing of the intraseasonal variability. For example, consider the suppressed event in the first half of August. Especially at longer lead times, positive OLR differences (orange) precede it and negative differences (blue) follow it, which reduces the biases from ATM. For the NEMO forecasts, the impact relative to KPP is very minimal until about 1.5 weeks into the forecast. Thenceforth, broadly

speaking the impact of ocean dynamics further reduces the biases from ATM. The improvements from coupling are summarized by Figure 8b, with KPP and NEMO having improved r and RMSE relative to ATM, but with minimal differences between KPP and NEMO until around 10 days' lead time.

In observations, SST (black contours in Figure 8a) are approximately in quadrature with convection, with warm SST anomalies leading active convection. This relationship holds very accurately in both the KPP and NEMO coupled forecasts (Figures 8e–i). The next section considers this phase relationship in more detail.

3.4 | Evaluation of coupled feedbacks in the northern BoB

Air–sea interactions over the northern BoB (north of 15°N) are of particular interest because the SST anomalies in the north can modulate the meridional gradient of SST across the Bay, which in turn can affect the northward propagation of the monsoon (Shankar *et al.*, 2007). Vecchi and Harrison (2002) examined intraseasonal variability of satellite-based OLR with respect to satellite SST anomalies in the northern BoB. Positive and negative SST anomalies on the intraseasonal scale in the northern BoB exhibit a strong statistical relation with the BSISO over India, with SST leading rainfall (convection) by about ten days. We illustrate such lead–lag correlations from the MetUM forecasts and three rainfall datasets – TRMM 3B42, GPCP and ERA5 – correlated with OSTIA SSTs. We also assess regional co-variability of forecast surface heat fluxes and SST over the northern BoB, and compare these with CERES radiation combined with OAFlux turbulent fluxes and other flux products on intraseasonal scales, including the coupling coefficients, to illustrate coupled feedbacks with the ocean in KPP and NEMO.

3.4.1 | Lagged correlations of precipitation and SST

Figure 9a–c show maps of lead–lag correlations for ERA5 SST averaged over the north BoB with gridpoint TRMM precipitation in 65°–105°E, 0°–30°N during JJAS 2016. Figure 9d–l shows the same but for the forecasts at 7 days' lead time.

For observations, the maximum positive lagged correlations across the BoB and over eastern and central India are seen when the BoB SST leads regional precipitation by about 9 days (Figure 9a) and suggest an atmospheric response to oceanic forcing (i.e., high SST leads enhanced rainfall; low SST leads reduced rainfall). The maximum negative correlations are found when BoB SST follows rainfall with a lag of about 5 days (Figure 9c) and suggest an oceanic response to atmospheric forcing (i.e., low SST lags enhanced rainfall; high SST lags reduced rainfall). Taken together, these correlations suggest an asymmetry in air–sea interactions, by which the oceanic response to atmospheric forcing is faster than the atmospheric response to oceanic forcing. The instantaneous correlations (Figure 9b) are negative, such that the atmosphere drives SST variability.

We now examine whether the three sets of forecasts reproduce the observed SST–rainfall phase relationship at lead times of 7 days (Figure 9d–l). ATM develops a highly biased correlation pattern, even at lag 0 (Figure 9d–f). Lagged correlations are very weak over the northern BoB. Correlations over the central Bay are similar to observations, but are too zonal and tend to spread eastwards, suggesting errors in the location of rainfall anomalies. This is consistent with the mean precipitation results, where ATM shows a more uniform distribution of rainfall than in TRMM (Figures 1c,d).

Differences in the strength and location of the lagged SST–rainfall correlations are likely related to the fact that variability in atmospheric fluxes cannot influence the fixed daily SST in ATM (e.g., Kumar *et al.*, 2013; Shelly *et al.*, 2014). It may also be partly related to problems simulating clouds in the MetUM, which affect both radiative and precipitation fluxes, such as erroneously strong sensitivity of parametrized atmospheric convection to SST (Klingaman *et al.*, 2008a; Hirons *et al.*, 2018). In contrast, both KPP (Figure 9g–i) and NEMO (Figure 9j–l) coupled forecasts are generally able to reproduce a very similar pattern of lagged correlations to that obtained from combined ERA5 SST and TRMM fields, but with some differences in magnitude. For example, when BoB SST leads rainfall by 9 days (Figure 9g,j), both KPP and NEMO develop visibly weaker correlations (<0.4) than observed over most of the BoB, whereas over India, the maximum correlation region shifts northwestward compared to observations.

To develop a more detailed comparison of these SST–rainfall phase relationships, Figure 10 shows regionally averaged lagged correlations for the BoB and India (outlined in Figure 9a) for the forecasts at all lead times, along with other observational and reanalysis products to highlight observational uncertainty. For the north BoB (Figure 10a), the satellite precipitation products (TRMM, GPCP) combined with ERA5 SST show a similar lead–lag behaviour. The ERA5 (atmosphere-only) reanalysis shifts the maximum positive SST–rainfall correlation to shorter lags, relative to TRMM and GPCP, consistent with the hypothesis that atmosphere-only models favour in-phase SST–rainfall relationships. Among the forecasts, ATM maintains the correlation amplitude (>0.6) but quickly shifts to an in-phase SST–rainfall relationship by forecast day 3, reflecting fixed SST and the absence of coupled feedbacks. The two coupled forecasts are largely consistent and show significant improvements over ATM, with a clear lead–lag relationship similar to observations, but the phase relationship breaks down after day 6 of the forecasts.

The overall patterns for the India region (Figure 10b) resemble those for the BoB area, but with considerably lower correlations, especially at lead times of one week or less.
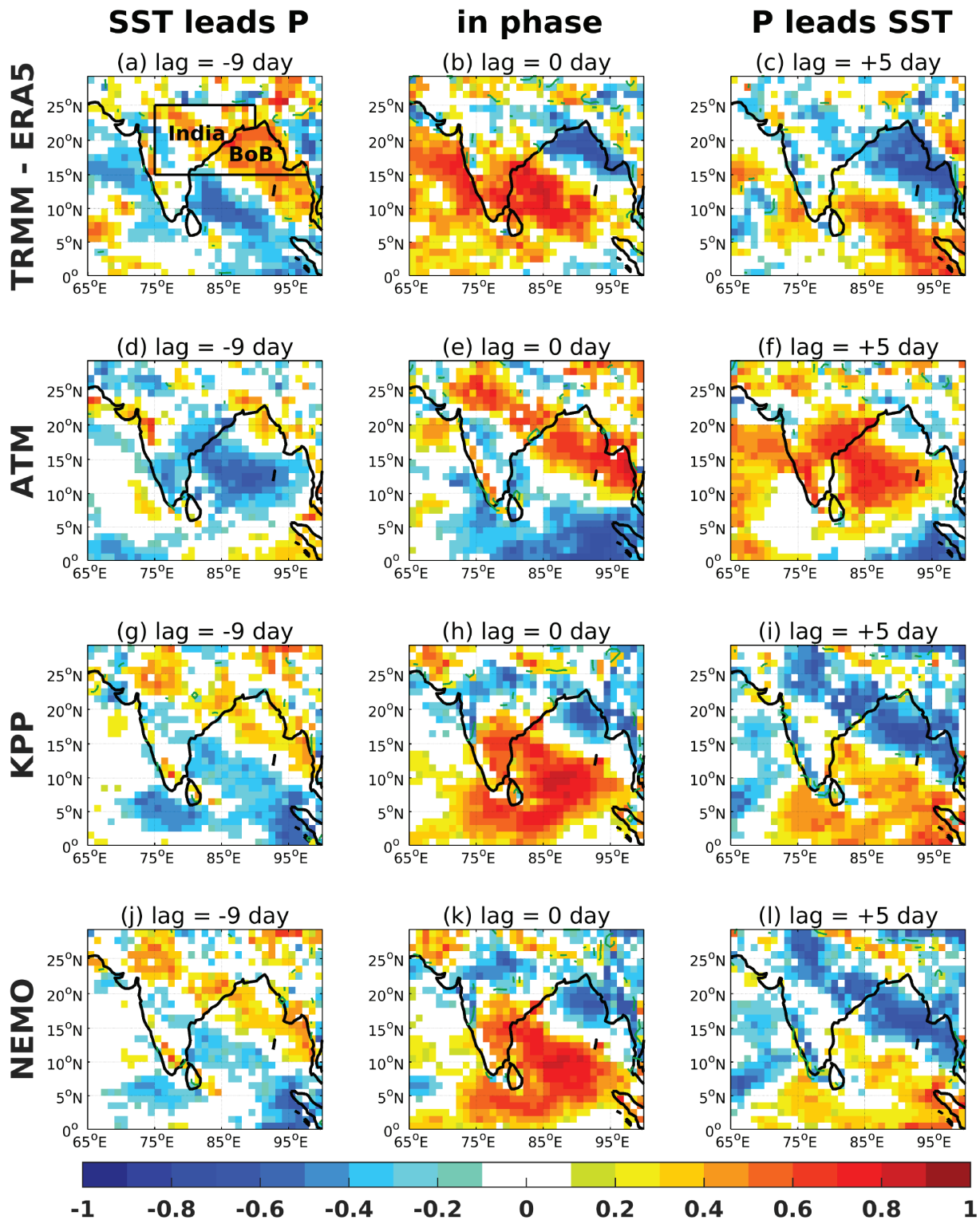
FIGURE 9 Maps of lead–lag correlation between SST anomaly averaged over the northern BoB and local precipitation (P) anomaly at each grid point (see main text for details). (a)–(c) TRMM precipitation and ERA5 SST; (d)–(f) ATM forecasts; (g)–(i) KPP forecasts and (j)–(l) NEMO forecasts. Lag correlations for all forecasts are shown at a lead time of 7 days. Negative lag indicates SST leading precipitation and positive lag indicates precipitation leading SST. Lag times of −9, 0 (in phase) and +5 days are shown, with correlations plotted only where they are significantly non-zero at the 95% confidence level (according to Student's t-test). Two bounded regions, "India" and "BoB", are indicated on (a)
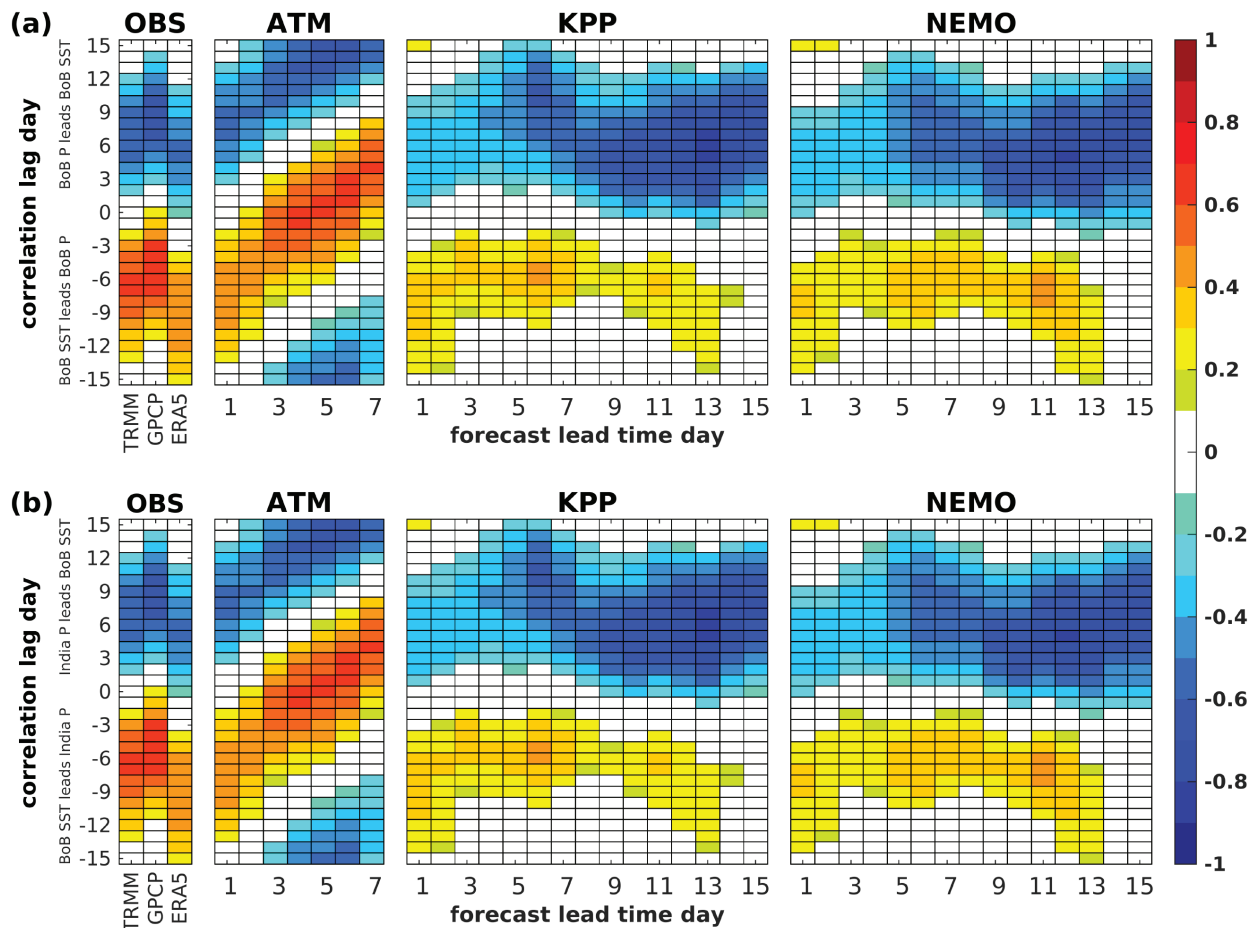
FIGURE 10 Regionally averaged lead–lag correlations for observational products and all forecasts: (a) between SST and precipitation averaged over the north BoB (15°–22°N, 80°–100°E); (b) between SST averaged over the north BoB and precipitation averaged over eastern and central India (15°–25°N, 75°–90°E). The y-axis is the day of lagged correlation. For observational data, SST is from ERA5 and precipitation is taken from three products (TRMM, GPCP and ERA5). For forecasts, the x-axis is forecast lead time. Correlations are estimated to be significant at 95% confidence levels when their absolute values exceed the range: 0.18–0.21 for observational products, 0.19–0.24 for ATM, 0.19–0.22 for KPP and 0.18–0.25 for NEMO

We note that the biases in forecast lagged correlations can result from the interplay of model biases that develop over the northern BoB at longer leads (details provided in Figure 5 in Section 3.2). Of particular interest is the degradation in the lagged correlation in KPP and NEMO after 6 days' lead time. This may be related to the developing biases of reduced precipitation and excessive short-wave radiation at longer lead times, which may reduce the sensitivity of atmospheric convection to SST variability, shifting the model into a regime in which the ocean responds passively to atmospheric forcing. This interpretation may be an oversimplification, as errors in other components of surface heat fluxes, especially in the latent heat fluxes, as well as errors in ocean mixing in response to overestimated forecast winds, may also play a role. Uncovering cause and effect would require a large number of sensitivity experiments across these model configurations and is

beyond the scope of this paper. We return to this discussion in Section 3.4.3.

3.4.2 | Surface heat flux variability

Figure 11 shows daily timeseries of surface heat fluxes and wind speeds from the forecasts averaged over the northern BoB during summer 2016, along with the corresponding observational estimates, including satellite CERES radiation combined with OAFlux turbulent fluxes, the ERA5 atmospheric reanalysis and the hybrid product TropFlux. All data have been lowpassed with a 7-day running mean to highlight intraseasonal variability. The 2016 JJAS mean Q_{net} from the CERES/OAFlux combination averaged over the north BoB is $18\,\mathrm{W\cdot m^{-2}}$, with Q_{sw} ($173\,\mathrm{W\cdot m^{-2}}$) balanced by Q_{lat} ($-113\,\mathrm{W\cdot m^{-2}}$) and, to a lesser extent, Q_{lw} ($-36\,\mathrm{W\cdot m^{-2}}$). Note that the heat flux

associated with rain falling on the ocean and the surface heat storage over summer 2016 (due to the annual cycle in mixed-layer depth) are not included in this net heat flux estimate because of the lack of comparable estimates in the models. Specifically, the rain heat flux is not included in the models, and the upper ocean heat storage cannot be estimated from the available data over JJAS 2016 (although this could be derived with the appropriate output over the full calendar year).

All forecasts reproduce the subseasonal variability in Q_{net} (Figure 11a–c) dominated by variability in Q_{sw} (± 41 W·m^{-2} in CERES observations; Figure 11d–f), with periods of net surface cooling (i.e., negative Q_{net}) that persist for several days during July–September 2016 associated with the BSISO, although with occasional large variability not present in the observational products.

The stronger subseasonal variability in the forecasts compared to the observational products is clearly seen in both radiative (Figure 11d–f) and turbulent (Figure 11g–i) flux components, with the resulting Q_{net} differences having significant correlated variability with either net radiation or turbulent flux errors depending on the BSISO phase. For example, overestimation of Q_{net} shown in all models during light wind events (break phases), as shown in Figure 11j–l, is attributed to significant overestimation of Q_{sw} (Figure 11d–f) compared to CERES, while the overestimation of the net heat loss during strong wind events (active phases) is primarily explained by overestimation of the turbulent fluxes (Figure 11g–i) due to overestimated winds in the models compared to satellite data (assimilated in both ERA5 and OAFlux products). The strongest wind event over the BoB is seen in the first week of July, leading to roughly 50 W·m^{-2} of overestimated cooling compared to OAFlux. For the two subsequent active events at the beginning of August and September, cancellation of Q_{sw} and Q_{lat} errors creates a Q_{net} in the models closer to CERES/OAFlux. This is consistent with the summer-mean comparisons in Figure 5, which show nearly all forecasts within 10 W·m^{-2} of the CERES/OAFlux values in the northern BoB.

Figure 11 also shows that the range of observation-based estimates can be very large on these intraseasonal scales (even larger than the corresponding summer-mean ranges shown in Figure 5). The smaller Q_{net} in ERA5 compared to CERES/OAFlux is dominated by Q_{lat}, suggesting different biases in the bulk variables in the flux computation. TropFlux shows higher Q_{net} over the BoB (mean of 32 W·m^{-2} compared with 18 W·m^{-2} from CERES/OAFlux) due to both smaller Q_{sw} and Q_{lat}. This suggests that more robust gridded surface flux products, underpinned by an enhanced observational network, are required to assess forecast intraseasonal flux variability.

3.4.3 | Feedback on heat fluxes through SST

Figure 12 shows coupling coefficients between SST and surface heat fluxes from the coupled forecasts and the CERES/OAFlux combined product during the summer monsoon season in 2016. The overall coupling coefficients (given in W·m^{-2} °C^{-1}) give the change in net surface heat flux (Q_{net}) per unit change in SST over JJAS 2016. These are calculated from linear regressions between daily SST and Q_{net} averaged over the BoB. Regressions of individual heat flux components onto SST are also shown.

As expected over the BoB, there is a negative feedback in CERES/OAFlux of ≈ -67 W·m^{-2} °C^{-1} dominated by Q_{lat} (≈ -50 W·m^{-2} °C^{-1}) and with a significant contribution from Q_{sw} (≈ -40 W·m^{-2} °C^{-1}), consistent with these terms dominating local heat budgets nearly everywhere in the BoB. In the coupled models, the negative feedback of Q_{net} is initially about the same as CERES/OAFlux, but it varies with forecast lead time, being either non-existent at short leads ($< \approx$ one week) or clearly less negative at longer leads compared to observations. The contribution from Q_{sw} has opposite signs between the models (positive) and CERES (negative), so that the Q_{sw} feedback in the forecasts contributes to SST warming at all leads. Although Q_{lat} cools the ocean in both models and observations, its contribution to the coupling strength is highly sensitive to forecast lead time: weaker than observations at leads of 2–7 days and stronger than observations at longer leads. This effect is partially compensated by the positive feedback from Q_{sw}, especially in NEMO (Figure 12b).

Overall, the results shown in Figure 12 indicate too weak coupling (i.e., too weak negative feedback of Q_{net} on SST) in the coupled forecasts in the north BoB area, primarily due to Q_{sw} errors and, to a lesser degree, Q_{lat} errors. Importantly, the simulated coupling coefficients depend strongly on lead time, being up to ≈ -50 W·m^{-2} °C^{-1} smaller than CERES/OAFlux at lead times <1 week. This is strong evidence that some aspects of intraseasonal prediction in KPP and NEMO configurations, such as the breakdown in the SST–rainfall phase relationship after 6 days' lead time (Figure 10a), may not only respond to seasonal-mean biases, but may be sensitive to the intraseasonal air–sea coupling strength.

4 | DISCUSSION

In these global-scale NWP forecasts, air–sea coupling has a limited effect on prediction performance for Indian monsoon rainfall. Most of the benefits are limited to oceanic regions, including improvements in mean rainfall and
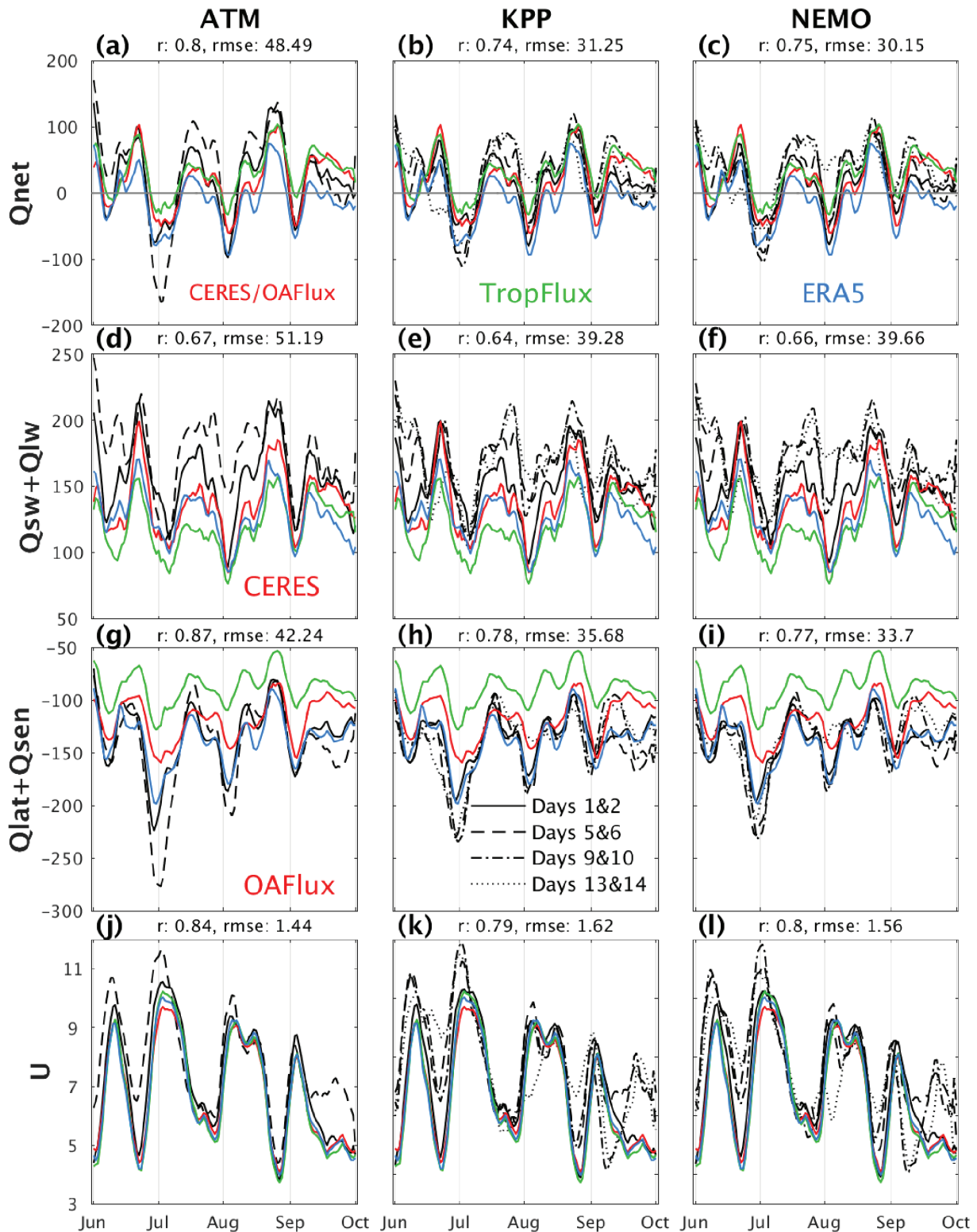
FIGURE 11 Time series of (a–i) surface heat fluxes and (j–l) 10 m wind speed U from forecasts averaged over the northern BoB during JJAS 2016, with corresponding observational estimates including satellite CERES radiation combined with OAFlux turbulent fluxes, ERA5 and the hybrid product TropFlux (Section 2.2). All data are low-pass filtered with a 7-day running mean to highlight intraseasonal variability. Black lines indicate forecasts averaged over days 1&2 (solid), 5&6 (dashed), 9&10 (dashed-dotted) and 13&14 (dotted). Coloured lines indicate CERES/OAFlux (red), TropFlux (green) and ERA5 (blue) estimates. Units are W·m^{-2} (heat flux components; positive into the ocean) and m·s^{-1} (wind speed). Correlation r and root-mean-square error (RMSE), averaged over days 5&6, against the ensemble of the three observational products, are given above each panel
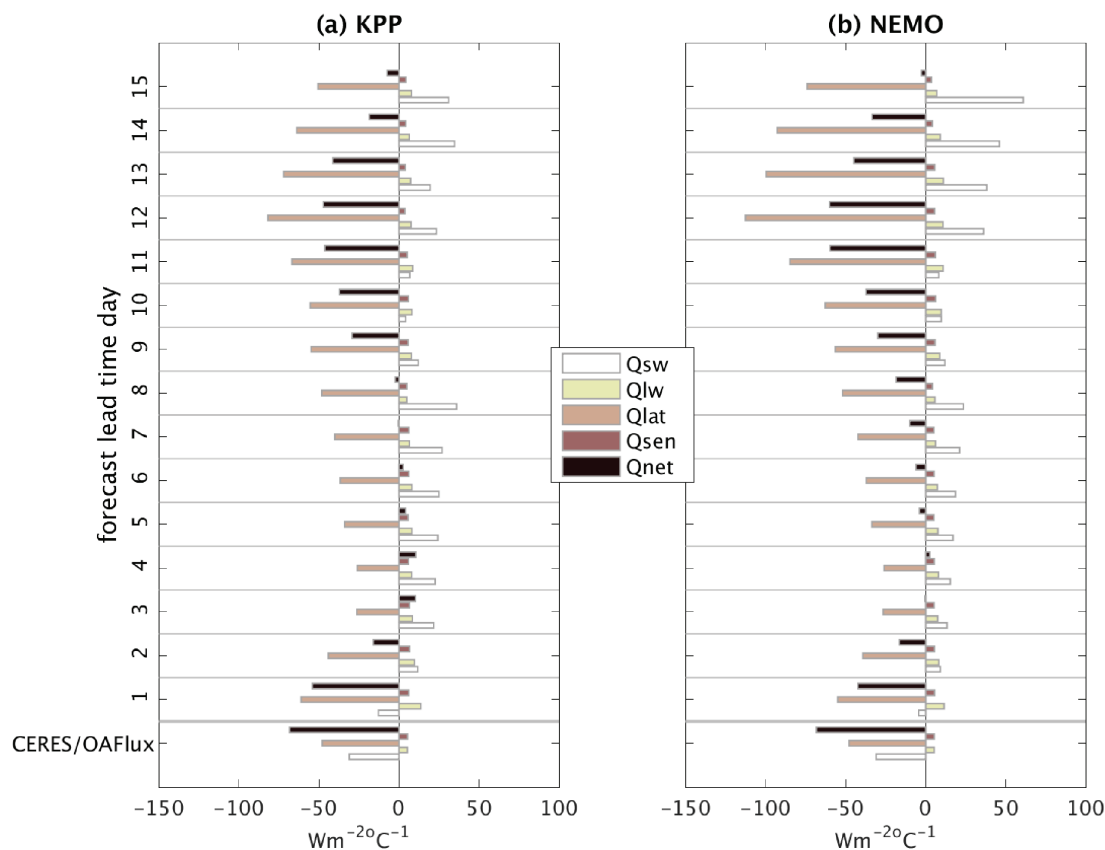
FIGURE 12 Coupling coefficients between SST and surface heat fluxes from the coupled (a) KPP and (b) NEMO forecasts and the CERES/OAFlux combined product during JJAS 2016. Coupling coefficients in $W \cdot m^{-2} \, °C^{-1}$ give the change in net surface heat flux (Q_{net}) per unit change in SST during the period. Coefficients are calculated by linear regression between SST and Q_{net}, spatially averaged over the northern BoB using daily data throughout JJAS 2016. Breakdowns from individual heat flux components (Q_{sw}, Q_{lw}, Q_{lat} and Q_{sen}) are also shown. Negative coefficients indicate a negative feedback (i.e., SST forcing the atmosphere) and positive coefficients a positive feedback (i.e., the atmosphere forcing SST)

in the local SST–rainfall phase relationship. The lack of improvement over Indian land may stem from the degraded phase relationships between BoB SST and Indian rainfall with lead time (Figure 10b), which suggests that, even if BoB SST were predicted perfectly, Indian rainfall predictions would remain biased. The degraded phase relationship, in turn, may be associated with the developing large-scale dry bias with lead time, a salient feature of MetUM sub-seasonal, seasonal and climate simulations. Strongly suppressed convection may force a transition from a regime in which the atmosphere responds to SST variability (i.e., in which SST anomalies lead rainfall anomalies of the same sign) into one in which the ocean passively responds to atmospheric variability (i.e., in which SST anomalies lag rainfall anomalies of the same sign). If the large-scale dry bias makes BoB convection insufficiently sensitive to SST (e.g., Figure 12), intraseasonal variations in SST and convection decouple and the observed link between BoB SST and Indian rainfall disappears. Thus, the season-mean bias in MetUM

may artificially suppress the effect of air–sea coupling for monsoon rainfall prediction.

The role of coupling may also be limited by our choice of a global NWP forecast configuration, the resolution of which necessitates using parametrized atmospheric convection. Relatively coarser resolution simulations with parametrized convection often display an incorrect diurnal cycle of convection, as well as weak or disorganised convection, particularly on synoptic to intraseasonal scales (e.g., Holloway *et al.*, 2013; Pearson *et al.*, 2014). Errors in the diurnal cycle of convection may lead to errors in the diurnal cycle of SST, and hence to the timing and magnitude of peak intraseasonal SST warming (Bernie *et al.*, 2005). Erroneously weak or disorganised convection would produce weak and disorganised SST anomalies that have a limited effect on surface fluxes, and hence on subsequent convection (e.g., Klingaman and Woolnough, 2014; DeMott *et al.*, 2016). The role of coupling in monsoon prediction may differ substantially in a finer-resolution regional-model configuration without

convective parametrization, particularly for synoptic systems such as monsoon depressions, the intensity of which often increase with resolution (Hunt and Turner, 2017b).

We analysed forecasts for a single season, for only a single deterministic forecast per day. While this gives us a larger sample of events than previous research based on case-studies, the sample is still far too small to draw robust conclusions about the role of coupling in monsoon prediction. Although 2016 was a relatively 'normal' monsoon season, the effects of coupling may be sensitive to interannual variability, for example if an El Niño Southern Oscillation or the Indian Ocean Dipole event changes the background state or the distribution of active and break events. Coupling may also have a greater effect on predictions during certain synoptic situations (e.g., monsoon depressions, break phases), but our limited sample does not allow us to quantify this. Using an ensemble-based forecasting system would have allowed us to examine the effect of coupling on probabilistic forecast skill (e.g., of heavy rainfall). More interestingly, using ensemble forecasts would also allow pursuing 'ensemble stratification' approaches to cluster together those ensemble members and synoptic situations that benefit most (or least) from coupling. Investigating the effects of coupling on monsoon NWP in higher-resolution and ensemble-based prediction systems is an active area of future research, as computational resources permit.

We note that in some cases the uncertainty in gridded, observation-based surface flux products is too large to allow us to distinguish between biases in the atmosphere-only and coupled forecasts. This emphasises the need to reduce errors and uncertainties in these products, through expanding and maintaining long-term observational platforms such as RAMA buoys, and through further field campaigns such as BoBBLE with collocated atmospheric and oceanic measurements.

5 | CONCLUSIONS

We analyse MetUM coupled (15-day) and uncoupled (7-day) forecasts (Table 1) of the Indian summer monsoon (JJAS) of 2016 in the context of the BoBBLE project (Vinayachandran *et al.*, 2018), in terms of seasonal-mean biases, active–break phases, intraseasonal SST–rainfall phase relationships and covariability of net surface heat fluxes and SSTs.

Seasonal-mean biases of precipitation (Figure 1) and SST (Figure 2) increase with lead time; coupling improves mean precipitation but degrades mean SST. ATM forecasts, with persisted initial SST from OSTIA, have initially more precipitation over India, but eventually too little precipitation and too much Q_{sw} warming over the

BoB, consistent with too little cloud cover. The excess Q_{sw} (>40 W·m^{-2} over the northwest bay) is partially compensated by about 25–35 W·m^{-2} stronger Q_{lat} over large areas of the BoB (Figures 3 and 5), which itself is primarily explained by overestimated surface winds relative to ERA5 (Figure 2) and satellite-based estimates (Figures 3 and 5). Positive air temperature and humidity biases contribute only slightly to the differences in turbulent fluxes (Figure 5).

While the regional patterns in KPP and NEMO remain similar to ATM, there are many notable local differences in magnitude (Figure 4). Over the BoB, the coupled forecasts show less radiative warming ($\approx$20 W·m^{-2}), with the largest reductions (>25 W·m^{-2}) over the monsoon upwelling areas off south India and Sri Lanka, associated with increased turbulent fluxes there. The mean SST biases (relative to OSTIA) in KPP and NEMO are much larger than those in ATM (Figure 2), even at the longest ATM lead times when the persisted SST has drifted furthest from OSTIA. The strong spatial correlation between differences in SST and Q_{net} (Figure 4) between KPP, NEMO and ATM is consistent with Q_{net} dominating local heat budgets (i.e., contributions from ocean storage, mixing and advection are smaller on average). The closer agreement with observations of simulated Q_{net} and the individual flux components in KPP and NEMO than in ATM indicates that the larger SST biases in KPP and NEMO offset atmospherically driven surface flux biases. Air–sea coupling also slightly improves the spatial distribution of precipitation across the BoB (Figure 1), partially restoring the observed southwest–northeast gradient. However, at all lead times these effects are considerably smaller than the model bias, so the effect on mean precipitation over India is negligible.

Further comparisons over the north BoB with three gridded surface flux datasets show mean biases of similar sign in all forecasts (Figure 5). Mean offsets are estimated to be +10–40 W·m^{-2} for Q_{sw}, −15 to 25 W·m^{-2} for Q_{lat}, and +1–3 W·m^{-2} for Q_{lw} and Q_{sen}. There are compensating errors in flux components; the magnitude of the error in Q_{net} is estimated to be <10 W·m^{-2} at all lead times. The dominant effects are anti-correlated Q_{sw} and Q_{lat}. Q_{sw} is overestimated mostly in monsoon break phases, while Q_{lat} is overestimated during active phases, associated with positive wind-speed biases (Figure 11). Q_{sw} and surface wind biases exceed the range of estimates from observation-based products at all lead times (Figure 5). However, the observational range of Q_{net} (dominated by the spread in turbulent fluxes) is quite large, suggesting the need for additional observations to improve gridded surface flux products.

The intraseasonal northward propagation of convection (BSISO) links air–sea feedbacks to rainfall over India,

since the air–sea feedbacks affect the generation and propagation of oceanic convection and its propagation to land. In these forecasts, air–sea coupling improves the convective propagation diagnosed in OLR, but there is very little impact from ocean dynamics (Figure 8b). However, this improvement does not translate into an improvement in the representation of precipitation, with coupling having very little effect on the intraseasonal cycle's modulation of the spatial pattern of precipitation (Figure 7).

Surface heat fluxes are also a key part of the oceanic heat budget. In the coupled forecasts, we examined feedback mechanisms over the northern BoB by deriving coupling coefficients between SST and Q_{net} and its components, and comparing against CERES/OAFlux (Figure 12). The coupling coefficient between SST and Q_{net} in observations is large and negative ($-67 Wm^{-2} {}^{\circ}C^{-1}$), primarily due to negative feedbacks of SST onto Q_{lat} and Q_{sw} throughout the monsoon season. These are consistent with warm SST generating increased convection, with associated cloudiness reducing Q_{sw}, and stronger winds increasing ocean-to-atmosphere Q_{lat}. In both KPP and NEMO, the Q_{net} coefficient is too weak, especially during the first few days of the forecast. This is caused by a combination of the feedback on Q_{lat} being too weak in approximately the first week of the forecasts, and the feedback on Q_{sw} almost always having the wrong sign. This is consistent with cloudy (clear) skies reducing (increasing) Q_{sw} in the model and therefore lowering (raising) SST, but observations imply the inverse. These results are consistent with the positive SST bias over the north BoB and insufficient or radiatively ineffective simulated clouds in MetUM.

We also investigated the consistency of the forecast intraseasonal SST–rainfall phase relationship within the forecasts on intraseasonal time-scales (Figures 9 and 10). For observations, maximum positive lagged correlations are seen when the north BoB SST leads local precipitation across the BoB and surrounding India land areas, by about 6–9 days. In ATM, the strength of the SST–rainfall lagged correlation does not change with lead time, but the time–phase relationship quickly reduces to zero (i.e., in-phase pattern of convection with underlying SST), reflecting the fixed daily SST and the absence of coupled feedbacks with the ocean. In contrast, the KPP and NEMO coupled forecasts have a roughly constant SST–rainfall phase relationship throughout the BoB for lead times up to around one week, which is consistent with observations. This consistent relationship was also seen between precipitation and OLR in Figures 8e–i, which remained in quadrature in coupled forecasts even out to long lead times. However, over the northern BoB the correlations break down after forecast day 6, whereas over India, they are generally small (<0.4) at all lead times, so northern

BoB SST makes only a small contribution to variations in precipitation over India in these forecasts. These effects may be associated with substantial, large-scale positive Q_{sw} errors which develop quickly over the north BoB, which reduces the sensitivity of local convection to SST.

Based on forecasts of summer 2016, our results show a limited role for air–sea coupling in global NWP for Indian monsoon rainfall. While coupling slightly corrects the timing of active and break phases over India and improves the local precipitation–SST relationship in the BoB, coupling is unable to arrest the development of stronger systematic errors at longer lead times. Future work should focus on reducing these errors, as well as examining the role of coupling in higher-resolution regional models, in ensemble forecasts and across longer reforecast periods.

ORCID
Maria Valdivieso https://orcid.org/0000-0002-1738-7016
Simon C. Peatman https://orcid.org/0000-0002-2511-7649

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

Agarwal, N., Sharma, R., Basu, S., Parekh, A., Sarkar, A. and Agarwal, V.K. (2007) Bay of Bengal summer monsoon 10–20 day variability in sea surface temperature using model and observations. *Geophysical Research Letters*, 34(6). https://doi.org/10.1029/2007GL029296.

Annamalai, H. and Sperber, K.R. (2005) Regional heat sources and the active and break phases of Boreal Summer Intraseasonal Variability. *Journal of the Atmospheric Sciences*, 62, 2726–2748.

Bernie, D.J., Woolnough, S.J. and Slingo, J.M. (2005) Modeling diurnal and intraseasonal variability of the ocean mixed layer. *Journal of Climate*, 18, 1190–1202.

Boos, W.R., Hurley, J.V. and Murthy, V.S. (2015) Adiabatic westward drift of Indian monsoon depressions. *Quarterly Journal of the Royal Meteorological Society*, 141, 1035–1048.

Buizza, R., Balsamo, G. and Haiden, T. (2018) IFS Upgrade brings more seamless coupled forecasts. *ECMWF Newsletter*, 156, 18–22.

Craig, A., Valcke, S. and Coquart, L. (2017) Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3.0. *Geoscientific Model Development*, 10, 3297–3308.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.

DeMott, C.A., Stan, C., Randall, D.A. and Branson, M.D. (2014) Intraseasonal variability in coupled GCMs: the roles of ocean feedbacks and model physics. *Journal of Climate*, 27, 4970–4995.

DeMott, C.A., Benedict, J.J., Klingaman, N.P., Woolnough, S.J. and Randall, D.A. (2016) Diagnosing ocean feedbacks to the MJO: SST-modulated surface fluxes and the moist static energy budget. *Journal of Geophysical Research: Atmospheres*, 121, 8350–8373.

Donlon, C.J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E. and Wimmer, W. (2012) The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of Environment*, 116, 140–158.

Duchon, C.E. (1979) Lanczos Filtering in one and two dimensions. *Journal of Applied Meteorology*, 18, 1016–1022.

Dutta, D., Routray, A., Kumar, D.P., George, J.P. and Singh, V. (2019) Simulation of a heavy rainfall event during southwest monsoon using high-resolution NCUM modeling system: a case study. *Meteorology and Atmospheric Physics*, 131, 1035–1054.

Fairall, C.W., Bradley, E.F., Hare, J.E., Grachev, A.A. and Edson, J.B. (2003) Bulk parameterization of air–sea fluxes: updates and verification for the COARE algorithm. *Journal of Climate*, 16, 571–591.

Feng, X., Klingaman, N.P. and Hodges, K.I. (2019) The effect of atmosphere–ocean coupling on the prediction of 2016 western North Pacific tropical cyclones. *Quarterly Journal of the Royal Meteorological Society*, 145, 2425–2444.

Fu, X. and Wang, B. (2004) The boreal-summer intraseasonal oscillations simulated in a hybrid coupled atmosphere–ocean model. *Monthly Weather Review*, 132, 2628–2649.

Gao, Y., Klingaman, N.P., DeMott, C.A. and Hsu, P.-C. (2019) Diagnosing ocean feedbacks to the BSISO: SST-modulated surface fluxes and the moist static energy budget. *Journal of Geophysical Research: Atmospheres*, 124, 146–170.

Goswami, B.N., Ajayamohan, R.S., Xavier, P.K. and Sengupta, D. (2003) Clustering of synoptic activity by Indian summer monsoon intraseasonal oscillation. *Geophysical Research Letters*, 30, 1431.

Hersbach, H., Bell, W., Berrisford, P., Horányi, A., Muñoz Sabatier, J., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C. and

Dee, D.P. (2019) Global reanalysis: goodbye ERA-Interim, hello ERA5. *ECMWF Newsletter*, 159, 17–24.

Hirons, L.C., Klingaman, N.P. and Woolnough, S.J. (2015) MetUM-GOML1: a near-globally coupled atmosphere–ocean mixed-layer model. *Geoscientific Model Development*, 8, 363–379.

Hirons, L.C., Klingaman, N.P. and Woolnough, S.J. (2018) The impact of air–sea interactions on the representation of tropical precipitation extremes. *Journal of Advances in Modeling Earth Systems*, 10, 550–559.

Holloway, C.E., Woolnough, S.J. and Lister, G.M.S. (2013) The effects of explicit versus parameterized convection on the MJO in a large-domain high-resolution tropical case study. Part I: characterization of large-scale organization and propagation. *Journal of the Atmospheric Sciences*, 70(5). https://doi.org/10.1175/JAS-D-12-0227.1.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F. and Wolff, D.B. (2007) The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8, 38–55.

Hunt, K.M.R. and Fletcher, J.K. (2019) The relationship between Indian monsoon rainfall and low-pressure systems. *Climate Dynamics*, 53, 1859–1871.

Hunt, K.M.R. and Turner, A.G. (2017a) The effect of soil moisture perturbations on Indian monsoon depressions in a numerical weather prediction model. *Journal of Climate*, 30, 8811–8823.

Hunt, K.M.R. and Turner, A.G. (2017b) The effect of horizontal resolution on Indian monsoon depressions in the Met Office NWP model. *Quarterly Journal of the Royal Meteorological Society*, 143, 1756–1771.

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. (2016) The resolution sensitivity of the South Asian Monsoon and Indo-Pacific in a global 0.35° AGCM. *Climate Dynamics*, 46, 807–831.

Klingaman, N.P., Inness, P.M., Weller, H. and Slingo, J.M. (2008a) The importance of high-frequency sea-surface temperature variability to the intraseasonal oscillation of Indian monsoon rainfall. *Journal of Climate*, 21, 6119–6140.

Klingaman, N.P., Weller, H., Slingo, J.M. and Inness, P.M. (2008b) The intraseasonal variability of the Indian summer monsoon using TMI sea surface temperatures and ECMWF reanalysis. *Journal of Climate*, 21, 2466–2483.

Klingaman, N.P. and Woolnough, S.J. (2014) The role of air–sea coupling in the simulation of the Madden–Julian oscillation in the Hadley Centre model. *Quarterly Journal of the Royal Meteorological Society*, 140, 2272–2286.

Klingaman, N.P., Woolnough, S.J., Weller, H. and Slingo, J.M. (2011) The impact of finer-resolution air–sea coupling on the intraseasonal oscillation of the Indian summer monsoon. *Journal of Climate*, 24, 2451–2468.

Krishnamurthy, V. and Shukla, J. (2007) Intraseasonal and seasonally persisting patterns of Indian monsoon rainfall. *Journal of Climate*, 20, 3–20.

Kumar, A., Zhang, L. and Wang, W. (2013) Sea surface temperature–precipitation relationship in different reanalyses. *Monthly Weather Review*, 141, 1118–1123.

Kumar, B.P., Vialard, J., Lengaigne, M., Murty, V.S.N. and McPhaden, M.J. (2011) TropFlux: air–sea fluxes for the global tropical

oceans – description and evaluation. *Climate Dynamics*, 38, 1521–1543.

Large, W.G., McWilliams, J.C. and Doney, S.C. (1994) Oceanic vertical mixing: a review and a model with a nonlocal boundary-layer parameterization. *Reviews of Geophysics*, 32, 363–403.

Lawrence, D.M. and Webster, P.J. (2002) The boreal intraseasonal oscillation: relationship between northward and eastward movement of convection. *Journal of the Atmospheric Sciences*, 59, 1593–1606.

Lee, J.-Y., Wang, B., Wheeler, M.C., Fu, X., Waliser, D.E. and Kang, I.-S. (2013) Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Climate Dynamics*, 40, 493–509.

Lee, S.-S., Wang, B., Waliser, D.E., Neena, J.M. and Lee, J.-Y. (2015) Predictability and prediction skill of the boreal summer intraseasonal oscillation in the intraseasonal variability hindcast experiment. *Climate Dynamics*, 45, 2123–2135.

Liebmann, B. and Smith, C.A. (1996) Description of a complete (interpolated) outgoing longwave radiation dataset. *Bulletin of the American Meteorological Society*, 77, 1275–1277.

Lucas, A.J., Shroyer, E.L., Wijesekera, H.W., Fernando, H.J.S., D'Asaro, E., Ravichandran, M., Jinadasa, S.U.P., MacKinnon, J.A., Nash, J.D., Sharma, R., Centurioni, L., Farrar, J.T., Weller, R., Pinkel, R., Mahadevan, A., Sengupta, D. and Tandon, A. (2014) Mixing to monsoons: air–sea interactions in the Bay of Bengal. *Eos*, 95, 269–270.

Mahadevan, A., Paluszkiewicz, T., Ravichandran, M., Sengupta, D. and Tandon, A. (2016) Introduction to the special issue on the Bay of Bengal: from monsoons to mixing. *Oceanography*, 29, 14–17.

Mallick, S.K., Agarwal, N., Sharma, R., Prasad, K.V.S.R. and Ramakrishna, S.S.V.S. (2020) Thermodynamic response of a high-resolution tropical Indian Ocean model to TOGA COARE bulk air–sea flux parameterization: case study for the Bay of Bengal (BoB). *Pure and Applied Geophysics*, 177, 4025–4044. https://doi.org/10.1007/s00024-020-02448-6.

McPhaden, M.J., Meyers, G., Ando, K., Masumoto, Y., Murty, V.S.N., Ravichandran, M., Syamsudin, F., Vialard, J., Yu, L. and Yu, W. (2009) RAMA: the research moored array for African-Asian-Australian monsoon analysis and prediction. *Bulletin of the American Meteorological Society*, 90, 459–480.

Pearson, K.J., Lister, G.M.S., Birch, C.E., Allan, R.P., Hogan, R.J. and Woolnough, S.J. (2014) Modelling the diurnal cycle of tropical convection across the "grey zone". *Quarterly Journal of the Royal Meteorological Society*, 140, 491–499.

Peatman, S.C. and Klingaman, N.P. (2018) The Indian Summer Monsoon in MetUM-GOML2.0: effects of air–sea coupling and resolution. *Geoscientific Model Development*, 11, 4693–4709.

Rai, D. and Pattnaik, S. (2019) Evaluation of WRF planetary boundary layer parameterization schemes for simulation of monsoon depressions over India. *Meteorology and Atmospheric Physics*, 131, 1529–1548.

Rajeevan, M., Gadgil, S. and Bhate, J. (2010) Active and break spells of the Indian summer monsoon. *Journal of Earth System Science*, 119, 229–247.

Rajeevan, M., Unnikrishnan, C.K. and Preethi, B. (2012) Evaluation of the ENSEMBLES multi-model seasonal forecasts of Indian summer monsoon variability. *Climate Dynamics*, 38, 2257–2274.

Rajesh, P., Pattnaik, S., Mohanty, U., Rai, D., Baisya, H. and Pandey, P. (2017) Land surface sensitivity of monsoon depressions formed over Bay of Bengal using improved high-resolution land state. *Dynamics of Atmospheres and Oceans*, 80, 155–172.

Rao, S.A., Saha, S.K., Pokhrel, S., Sundar, D., Dhakate, A.R., Mahapatra, S., Ali, S., Chaudhari, H.S., Shreeram, P., Vasimalla, S., Srikanth, A.S. and Suresh, R.R.V. (2011) Modulation of SST, SSS over northern Bay of Bengal on ISO time-scale. *Journal of Geophysical Research*, 116(C9). https://doi.org/10.1029/2010JC006804.

Roman-Stork, H.L., Subrahmanyam, B. and Murty, V. (2019) Quasi-biweekly oscillations in the Bay of Bengal in observations and model simulations. *Deep Sea Research Part II: Topical Studies in Oceanography*, 168(104609). https://doi.org/10.1016/j.dsr2.2019.06.017.

Routray, A., Mohanty, U.C., Niyogi, D., Rizvi, S.R.H. and Osuri, K.K. (2010) Simulation of heavy rainfall events over Indian monsoon region using WRF-3DVAR data assimilation system. *Meteorology and Atmospheric Physics*, 106, 107–125.

Rutan, D.A., Kato, S., Doelling, D.R., Rose, F.G., Nguyen, L.T., Caldwell, T.E. and Loeb, N.G. (2015) CERES synoptic product: methodology and validation of surface radiant flux. *Journal of Atmospheric and Oceanic Technology*, 32, 1121–1143.

Saha, S.K., Pokhrel, S., Salunke, K., Dhakate, A., Chaudhari, H.S., Rahaman, H., Sujith, K., Hazra, A. and Sikka, D.R. (2016) Potential predictability Indian summer monsoon rainfall in NCEP CFSv2. *Journal of Advances in Modeling Earth Systems*, 8, 96–120.

Sanchez-Franks, A., Kent, E.C., Matthews, A.J., Webber, B.G.M., Peatman, S.C. and Vinayachandran, P.N. (2018) Intraseasonal Variability of air–sea fluxes over the Bay of Bengal during the Southwest Monsoon. *Journal of Climate*, 31, 7087–7109.

Sengupta, D., Raj, G.N.B., Ravichandran, M., Lekha, J.S. and Papa, F. (2016) Near-surface salinity and stratification in the north Bay of Bengal from moored observations. *Geophysical Research Letters*, 43, 4448–4456.

Shankar, D., Shetye, S.R. and Joseph, P.V. (2007) Link between convection and meridional gradient of sea surface temperature in the Bay of Bengal. *Journal of Earth System Science*, 116, 385–406.

Sharma, R., Agarwal, N., Chakraborty, A., Mallick, S., Buckley, J., Shesu, V. and Tandon, A. (2016) Large-scale air–sea coupling processes in the Bay of Bengal using space-borne observations. *Oceanography*, 29, 192–201.

Shelly, A., Xavier, P., Copsey, D., Johns, T., Rodriguez, J.M., Milton, S. and Klingaman, N.P. (2014) Coupled versus uncoupled hindcast simulations of the Madden–Julian oscillation in the year of tropical convection. *Geophysical Research Letters*, 41, 5670–5677.

Shenoi, S.S.C. (2002) Differences in heat budgets of the near-surface Arabian Sea and Bay of Bengal: implications for the summer monsoon. *Journal of Geophysical Research*, 107(C6). https://doi.org/10.1029/2000JC000679.

Sikka, D.R. (1977) Some aspects of the life history, structure and movement of monsoon depressions. *Pure and Applied Geophysics*, 115, 1501–1529.

Storkey, D., Blockley, E.W., Furner, R., Guiavarc'h, C., Lea, D., Martin, M.J., Barciela, R.M., Hines, A., Hyder, P. and Siddorn, J.R. (2010) Forecasting the ocean state using NEMO: the new FOAM system. *Journal of Operational Oceanography*, 3, 3–15.

Vecchi, G.A. and Harrison, D.E. (2002) Monsoon breaks and subseasonal sea surface temperature variability in the Bay of Bengal. *Journal of Climate*, 15, 1485–1493.

Vinayachandran, P.N., Matthews, A.J., Vijay Kumar, K., Sanchez-Franks, A., Thushara, V., George, J., Vijith, V., Webber,

B.G.M., Queste, B.Y., Roy, R., Sarkar, A., Baranowski, D.B., Bhat, G.S., Klingaman, N.P., Peatman, S.C., Parida, C., Heywood, K.J., Hall, R., King, B., Kent, E.C., Nayak, A.A., Neema, C.P., Amol, P., Lotliker, A., Kankonkar, A., Gracias, D.G., Vernekar, S., Souza, A.C.D., Valluvan, G., Pargaonkar, S.M., Dinesh, K., Giddings, J. and Joshi, M.M. (2018) BoBBLE (Bay of Bengal Boundary Layer Experiment): ocean–atmosphere interaction and its impact on the South Asian monsoon. *Bulletin of the American Meteorological Society*, 99, 1569–1587.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., Hendon, H., Hodgson, J., Kang, H.-S., Kumar, A., Lin, H., Liu, G., Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mastrangelo, D., MacLachlan, C., McLean, P., Minami, A., Mladek, R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A.W., Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser, D., Woolnough, S., Wu, T., Won, D.-J., Xiao, H., Zaripov, R. and Zhang, L. (2017) The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, 98, 163–173.

Walters, D.N., Boutle, I.A., Brooks, M., Melvin, T., Stratton, R.A., Vosper, S., Wells, H., Williams, K.D., Wood, N., Allen, T., Bushell, A.C., Copsey, D., Earnshaw, P.D., Edwards, J.M., Gross, M., Hardiman, S., Harris, C.M., Heming, J., Klingaman, N.P., Levine, R.C., Manners, J.C., Martin, G.M., Milton, S.F., Mittermaier, M., Morcrette, C.J., Riddick, T., Roberts, M.J., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P.L., Wilkinson, J.M., Willett, M.R., Woolnough, S.J. and Xavier, P.K. (2017) The Met Office Unified Model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations. *Geoscientific Model Development*, 10, 1487–1520.

Webster, P.J., Magaña, V.O., Palmer, T.N., Shukla, J., Tomas, R.A., Yanai, M. and Yasunari, T. (1998) Monsoons: processes, predictability, and the prospects for prediction. *Journal of Geophysical Research*, 103, 14451–14510.

Weller, R., Farrar, J.T., Buckley, J., Matthew, S., Venkatesan, R., Lekha, J.S., Chaudhuri, D., Kumar, N.S. and Kuman, B.P. (2016) Air–sea interaction in the Bay of Bengal. *Oceanography*, 29, 28–37.

Wijesekera, H.W., Shroyer, E., Tandon, A., Ravichandran, M., Sengupta, D., Jinadasa, S.U.P., Fernando, H.J.S., Agrawal, N., Arulananthan, K., Bhat, G.S., Baumgartner, M., Buckley, J., Centurioni, L., Conry, P., Farrar, J.T., Gordon, A.L., Hormann, V., Jarosz, E., Jensen, T.G., Johnston, S., Lankhorst, M., Lee, C.M., Leo, L.S., Lozovatsky, I., Lucas, A.J., Mackinnon, J., Mahadevan, A., Nash, J., Omand, M.M., Pham, H., Pinkel, R., Rainville, L., Ramachandran, S., Rudnick, D.L., Sarkar, S., Send, U., Sharma, R., Simmons, H., Stafford, K.M., Laurent, L.S., Venayagamoorthy, K., Venkatesan, R., Teague, W.J., Wang, D.W., Waterhouse, A.F., Weller, R. and Whalen, C.B. (2016) ASIRI: an ocean–atmosphere initiative for Bay of Bengal. *Bulletin of the American Meteorological Society*, 97, 1859–1884.

Williams, K.D., Harris, C.M., Bodas-Salcedo, A., Camp, J., Comer, R.E., Copsey, D., Fereday, D., Graham, T., Hill, R., Hinton, T.J., 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. (2015) The Met Office global coupled model 2.0 (GC2) configuration. *Geoscientific Model Development*, 8, 1509–1524.

Yang, S. and Lau, W.K.-M. (2006). Interannual variability of the Asian monsoon, Chapter 6 in The Asian monsoon, Wang, B.(ed.) Praxis, Chichester, UK.

Yu, L., Jin, X. and Weller, R.A. (2008). Multidecade global flux datasets from the Objectively Analyzed air-sea Fluxes (OAFlux) Project: Latent and sensible heat fluxes, ocean evaporation, and related surface meteorological variables. OAFlux Project Technical Report OA-2008-01, Woods Hole Oceanographic Institution, Woods Hole, MA.