JOBNAME: MWR 00#00 2015 PAGE: 1 SESS: 8 OUTPUT: Wed May 13 11:56:24 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

Evaluation of Submonthly Precipitation Forecast Skill from Global Ensemble Prediction Systems

SHUHUA LI AND ANDREW W. ROBERTSON

International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, New York

(Manuscript received 29 August 2014, in final form 16 March 2015)

ABSTRACT

The prediction skill of precipitation at submonthly time scales during the boreal summer season is investigated based on hindcasts from three global ensemble prediction systems (EPSs). The results, analyzed for lead times up to 4 weeks, indicate encouraging correlation skill over some regions, particularly over the Maritime Continent and the equatorial Pacific and Atlantic Oceans. The hindcasts from all three models correspond to high prediction skill over the first week compared to the following three weeks. The ECMWF forecast system tends to yield higher prediction skill than the other two systems, in terms of both correlation and mean squared skill score. However, all three systems are found to exhibit large conditional biases in the tropics, highlighted using the mean squared skill score.

The sources of submonthly predictability are examined in the ECMWF hindcasts over the Maritime Continent in three typical years of contrasting ENSO phase, with a focus on the combined impact of the intraseasonal MJO and interannual ENSO. Rainfall variations over Borneo in the ENSO-neutral year are found to correspond well with the dominant MJO phase. The contribution of ENSO becomes substantial in the two ENSO years, but the MJO impact can become dominant when the MJO occurs in phases 2–3 during El Niño or in phases 5–6 during the La Niña year. These results support the concept that ''windows of opportunity'' of high forecast skill exist as a function of ENSO and the MJO in certain locations and seasons, which may lead to subseasonal-to-seasonal forecasts of substantial societal value in the future.

1. Introduction

There has been a recent surge of interest in prediction on subseasonal time scales between the realm of medium-range weather forecasts (up to 7–10 days) and seasonal climate forecasts (3–6 months) (National Academy of Sciences 2010; World Meteorological Organization 2013). The subseasonal realm lies beyond the estimated limit of deterministic weather predictability in atmospheric models (Lorenz 1975), but falls short of the seasonal averages where anomalies in the atmosphere's surface boundary conditions are felt, principally from sea surface temperatures (SSTs) (Charney and Shukla 1981). This interest has been driven by recent advances in modeling (Vitart 2014) and understanding of climate phenomena on subseasonal time scales, and in particular the Madden–Julian oscillation [MJO; see recent review by Zhang (2013)]. The concept of seamless prediction, in which there is potential predictive power on all time scales associated with phenomena occurring on that time scale (Hoskins 2013), is particularly attractive in the subseasonal time range, where there is evidence that in addition to the inertia of SST anomalies, the MJO (Waliser et al. 2003X; AU1 Waliser 2011; Neena et al. 2014X) and processes involving AU2 the stratosphere (Baldwin and Dunkerton 2001; Scaife and Knight 2008), soil moisture (Koster et al. 2010), snow cover (Lin and Wu 2011), surface, and sea ice (Holland et al. 2011) can play important roles. Besides the growing scientific basis for subseasonal prediction, the weekly-to-seasonal time range is one in which early warning systems for high-impact weather events could be of substantial societal benefit (World Meteorological Organization 2013). This is particularly urgent in the context of increasing societal exposure to extreme weather threats, be they caused by growing populations or due to decadal climate variability or anthropogenic climate change.

A number of previous studies, such as Waliser et al. (2003X) and Neena et al. (2014X), have investigated AU3 AU4

Corresponding author address: Andrew W. Robertson, IRI—Monell 230, 61 Route 9W, Palisades, NY 10964.
E-mail: awr@iri.columbia.edu

DOI: 10.1175/MWR-D-14-00277.1

JOBNAME: MWR 00#00 2015 PAGE: 2 SESS: 8 OUTPUT: Wed May 13 11:56:24 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

TABLE 1. Three global EPS dynamical models. The JMA and ECMWF hindcasts start from 20 May and then are updated every 10 days or so (i.e., 20, 31 May; 10, 20, 30 Jun; and so on), and every 7 days, respectively. The CFSv2 hindcast starts from 21 May, and then updated every 5 days. Note that the grid resolution denotes the postprocessed regridded data for the JMA and ECMWF models, whose original resolutions of model integrations are much higher, as described in the text.

Model	Grid resolution	Ensemble	Frequency	No. of starts	Period
JMA	144 × 73	5	3 month^{-1}	13	1979–2008
CFSv2	384 × 190	4	5 day	25	1982–2010
ECMWF	360 × 181	5	Weekly	18	1992–2009

intraseasonal variability over the eastern Pacific and predictability associated with the Asian summer monsoon. The aim of this paper is to examine submonthly precipitation prediction skill for the June–September boreal summer season using three ensemble prediction systems (EPSs) that are currently used for operational subseasonal prediction. While weather forecasts are usually evaluated on a daily basis, seasonal forecasts are typically issued as 3-month averages, so as to enhance the predictable climate signal by averaging out the daily weather noise (Palmer and Anderson 1994; Mason et al. 1999). In keeping with the intermediate time scale, we focus here on weekly averages at forecast leads of 1–4 weeks. The subseasonal range is beyond deterministic weather prediction so that some time aggregation is required, and the weekly scale makes a natural choice. Weekly time averages are chosen for simplicity and convenience of comparison with observations, although predictability may be higher (perhaps considerably) for daily rainfall occurrence frequency (i.e., the number of rain days), especially in the tropics (Moron et al. 2006). Because of the limited ensemble size (4–5) for all the three models, the skill evaluation will be confined to deterministic (i.e., nonprobabilistic) skill metrics only.

In section 2 a description of models and precipitation data, along with the analysis methods, is provided. The results are presented in section 3, with a summary and discussion given in section 4.

2. Models and data

a. Description of models

The model forecast data analyzed in this study are hindcasts from the National Centers for Environmental Prediction (NCEP) Climate Forecast System, version 2 (CFSv2) (Saha et al. 2010), the Japan Meteorological Agency (JMA) one-month EPS (Japan Meteorological Agency 2013), and the European Centre for Medium-Range Weather Forecasts (ECMWF) Variable Resolution Ensemble Prediction System monthly forecast system (VarEPS-monthly) used for operational monthly forecasts at ECWMF (Vitart et al. 2008). All the three sets of hindcasts are produced based on global EPS forecast systems (see Table 1 for detail). Since these T1
forecast data have 4–5 ensemble members only, our analysis uses the ensemble mean for each of the three model hindcast sets.

The NCEP CFSv2 (Saha et al. 2014) is a fully coupled dynamical prediction system. The atmospheric component of CFSv2 is a 2007 version of the NCEP Global Forecast System (GFS) (Saha et al. 2010), with a spectral truncation of 126 waves (T126) in the horizontal and 64 levels in the vertical. The oceanic component is the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model version 4 (MOM4) (Griffies et al. 2004). The CFSv2 system increases the length of skillful MJO forecasts [bivariate anomaly correlation (BAC) above 0.5] from 6 to 17 days compared to its predecessor, CFSv1 (Saha et al. 2014). The skill is highest during boreal winter and for target phases when enhanced convection is in the central Indian Ocean and the central Pacific (Wang et al. 2014). Also, CFSv2 makes more skillful ENSO predictions overall than CFSv1, especially for long-lead predictions from early in the calendar year through the northern spring predictability barrier (Barnston and Tippett 2013).

The ECMWF forecasting system is referred to as VarEPS-monthly, which is a merged 32-day ensemble system updated once a week that has been run since 2002 (Vitart 2014). We use the model configuration, which was operational in 2010. The EPS integration uses an atmospheric model at T399 resolution forced by persisted SST anomalies from days 0 to 10, and then coupled ocean–atmosphere integrations at T255 resolution from days 9 to 32 (Vitart et al. 2008; Hirons et al. 2013). The ocean component used for the coupled system is the Hamburg Ocean Primitive Equation Model (HOPE). The oceanic initial conditions include one control and four perturbed ocean analysis. Thus, the VarEPS-monthly reforecasts (or hindcasts) consist of a five-member ensemble of 32-day merged integrations for the 1992–2009 period. The model's skill at predicting the MJO has improved significantly since 2002, with an average gain of about 1 day of prediction skill per year, with BAC = 0.5 skill beyond 30 days in the version considered in the current study (Vitart 2014). This is the highest of the three models considered.

JOBNAME: MWR 00#00 2015 PAGE: 3 SESS: 8 OUTPUT: Wed May 13 11:56:24 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

JMA's one-month EPS consists of an atmospheric GCM (AGCM) run here at T159 resolution with atmospheric initial conditions obtained from the JMA global analysis and initial land surface conditions obtained from the JMA Land Surface Analysis System. Persisted SST anomalies are used as the lower boundary condition, together with a climatological distribution of sea ice (http://ds.data.jma.go.jp/tcc/tcc/products/model/outline/index.html). The MJO hindcasts are skillful up to a lead time of 13 days, but the model fails to reproduce the realistic eastward and northward propagation of active convection (Matsueda and Takaya 2012).

As seen in Table 1, the three forecast systems have different model resolution, different frequency of production (ranging from 5 to 10 days or so), and different physical configuration. However, they all provide daily (6-hourly for CFSv2) forecasts for up to about one month or beyond. Since our analysis is focused on targets from 1 to 4 weeks, all the hindcast data of precipitation were truncated at day 28, but spanning over all the available start dates from late May through mid-September. The common period for the three datasets is 1992–2008, and the postprocessed spatial resolution is regridded to the resolution of the observed precipitation dataset (i.e., 144 × 72 grid points at 2.5° longitude/latitude; Table 1). Note that the original horizontal resolution of the model integrations for the ECMWF hindcasts is T399 to day 10 and then T255, equivalent to approximately 80 km; while the JMA hindcast originally has T159 resolution, or about 125-km grid. The datasets for these two hindcast sets were provided to us as postprocessed data regridded at 1° and 2.5°, respectively (Table 1). The five ensemble members for the JMA and ECMWF hindcasts consist of one control run and four perturbed runs, while for CFSv2 the four ensemble members are made up with four runs each day, implemented at 0000, 0600, 1200, and 1800 UTC, respectively.

The hindcasts used in this study were obtained in early 2012, and correspond to the versions of the operational systems used at that time. Both the ECMWF and JMA systems have since been updated, while the CFSv2 is fixed. Additional model details of the three systems can be found in annex 2 of World Meteorological Organization (2013).

b. Precipitation data

The NCEP Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP) provides global precipitation data from 1979 onward at 2.5° resolution, using gauge observations, satellite estimates, and numerical model predictions (Xie and Arkin 1996, 1997). We use the CMAP pentad (5-day average) precipitation data during 1992–2008. This dataset is used to obtain daily estimates first using linear interpolation, from which the weekly averages are then calculated to match the model hindcast data.

c. Skill metrics

The most basic and commonly used skill metric for verification of subseasonal or seasonal forecasting is the correlation of anomalies (CORA), calculated between ensemble mean of model products against observed quantities (e.g., Zhu et al. 2014). The model forecast and observed anomalies are computed with respect to their own seasonally varying climatologies, respectively (Stockdale 1997). More importantly, the forecast anomalies are computed with respect to a lead-dependent hindcast climatology. In this study, we use the ensemble means of the model precipitation forecasts, averaged over all available start dates (ranging from 13 to 25 during the boreal summer season, varying among the three models) and valid for weeks 1–4. The corresponding weekly averages of the observations are computed from CMAP pentad mean rainfall data. Thus, the CORA is computed over all the available start dates during late May through mid-September over the 1992–2008 period. The four weekly periods correspond to averages over the forecast days 1–7 (week 1), days 8–14 (week 2), days 15–21 (week 3), and days 22–28 (week 4). Note that in order to reduce the impact of the annual cycle, the seasonality over the entire period is removed for the model hindcasts and observed rainfall, respectively, simply by subtracting their respective 17-yr climatology.

A second skill score is the mean square skill score (MSSS), which is based on mean squared error and often used in verification of deterministic forecasts (e.g., Goddard et al. 2013). The MSSS is a function of both the prediction to be evaluated, a reference prediction (here climatology), and the corresponding observations (Murphy 1988). Thus, the overall bulk MSSS value provides a comparison of forecast performance relative to the reference climatological prediction, together with an evaluation of the contribution from the conditional biases inherent in the forecasts (see Goddard et al. 2013).

3. Results

a. Correlation of anomalies (CORA) skill

1) Spatial distribution over the globe

The CORA maps between JMA precipitation hindcast and CMAP observation over the globe (50°S–60°N) are shown in Fig. 1. A dry mask based on total June–September (JJAS) season CMAP rainfall (1992–2008), set as 20 mm for all locations, is shown as white on all the

F1

JOBNAME: MWR 00#00 2015 PAGE: 4 SESS: 8 OUTPUT: Wed May 13 11:56:24 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

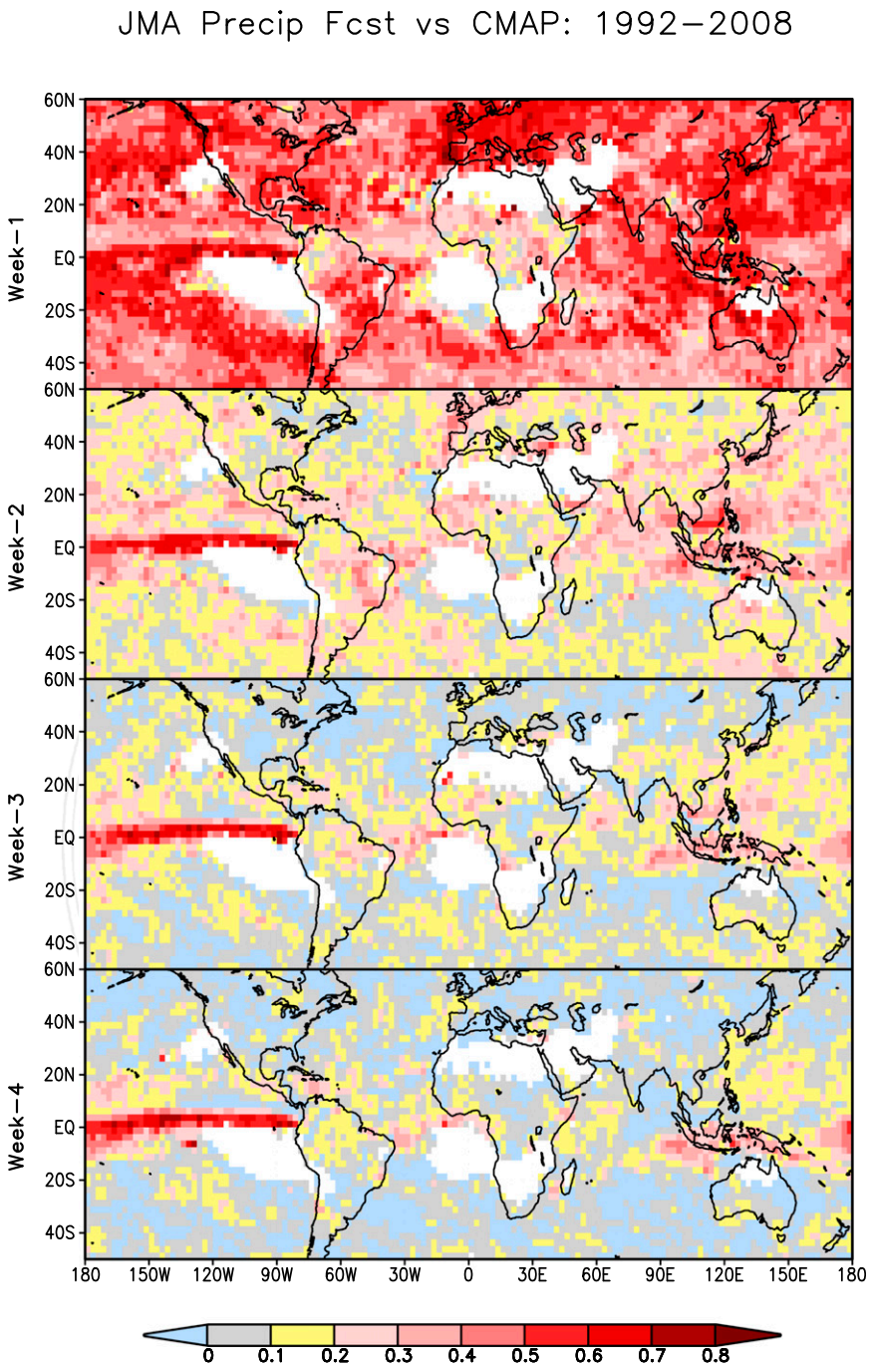

FIG. 1. Correlation of anomalies between JMA model precipitation hindcast and CMAP rainfall data for weeks 1–4 during the period 1992–2008. The white areas denote dry mask during the June–September season, where the total CMAP rainfall over 122 days is less than 20 mm.

maps. The correlation skill is generally very high during the first week and drops rapidly in most regions in the subsequent three weeks. Nevertheless, there is consistently high CORA for all the leads, or weeks 1–4, over the equatorial Pacific, located to the south of the intertropical convergence zone (ITCZ). Weeks 2–4 also exhibit relatively higher skills (0.2–0.3) over the tropical Atlantic, and the Maritime Continent. The 13 hindcasts per year over 17 years yields a very large sample size of 221 hindcasts. Using a two-sided Student's *t* test with 220

JOBNAME: MWR 00#00 2015 PAGE: 5 SESS: 8 OUTPUT: Wed May 13 11:56:25 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

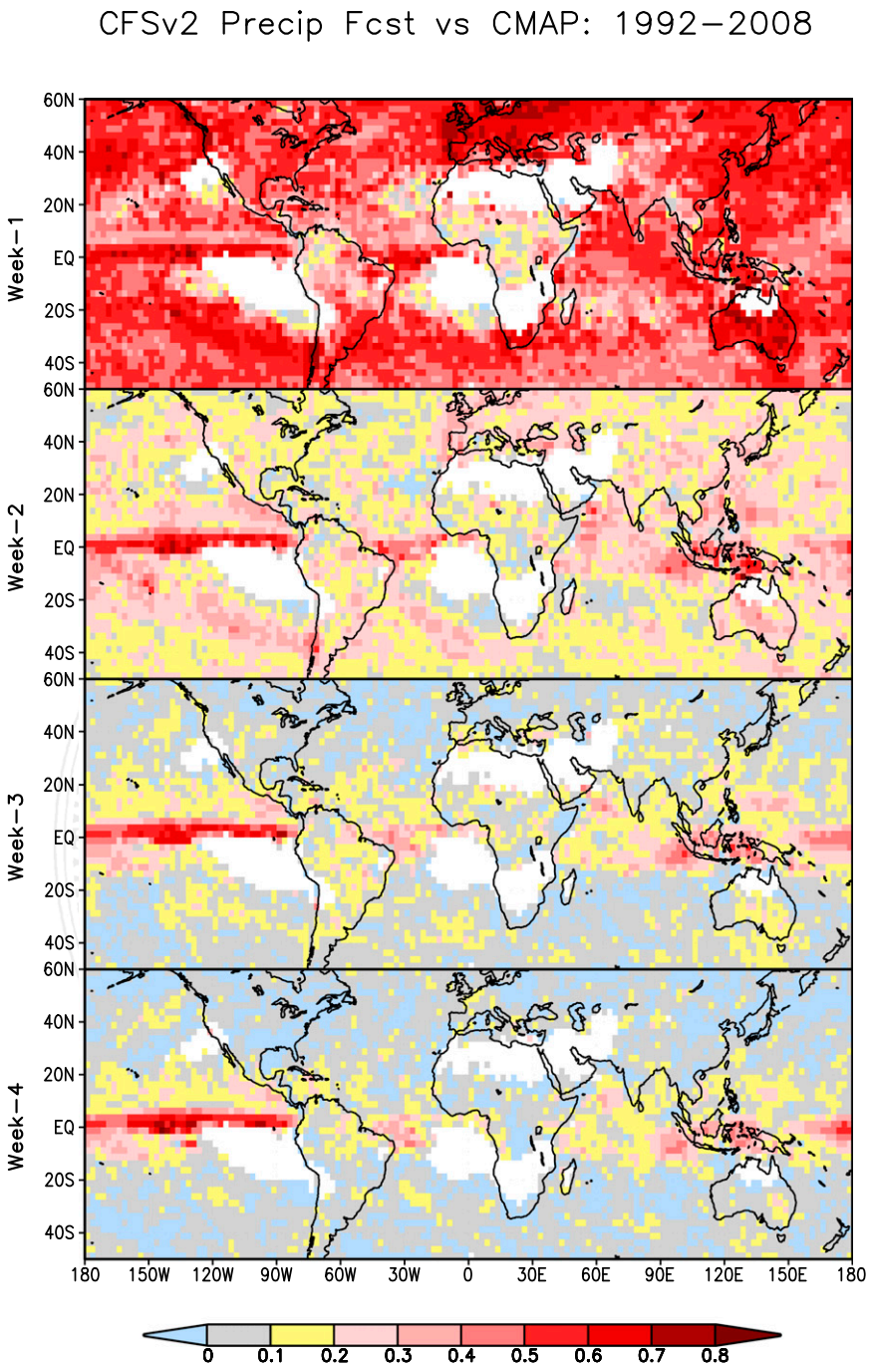

FIG. 2. As in Fig. 1, but for the CFSv2 hindcast.

degrees of freedom, a CORA value of 0.2 is highly statistically significant, with a p value of 0.003 shown as pink areas in Figs. 1–3.

The CORA distribution for the CFSv2 and ECMWF precipitation forecasts exhibit most of the above characteristics (Figs. 2 and 3). However, the CORA skill of ECMWF is notably higher than the other two models, especially over the tropics for weeks 2–4. Indeed, relatively high CORA values over the tropical Pacific resemble a slightly broader ITCZ-like characteristic than that from JMA or CFSv2. The high CORA feature from the ECMWF precipitation hindcast is also seen over the equatorial Atlantic and tropical Indian Ocean, as well as the Maritime Continent (Fig. 3), where higher CORA

JOBNAME: MWR 00#00 2015 PAGE: 6 SESS: 8 OUTPUT: Wed May 13 11:56:25 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

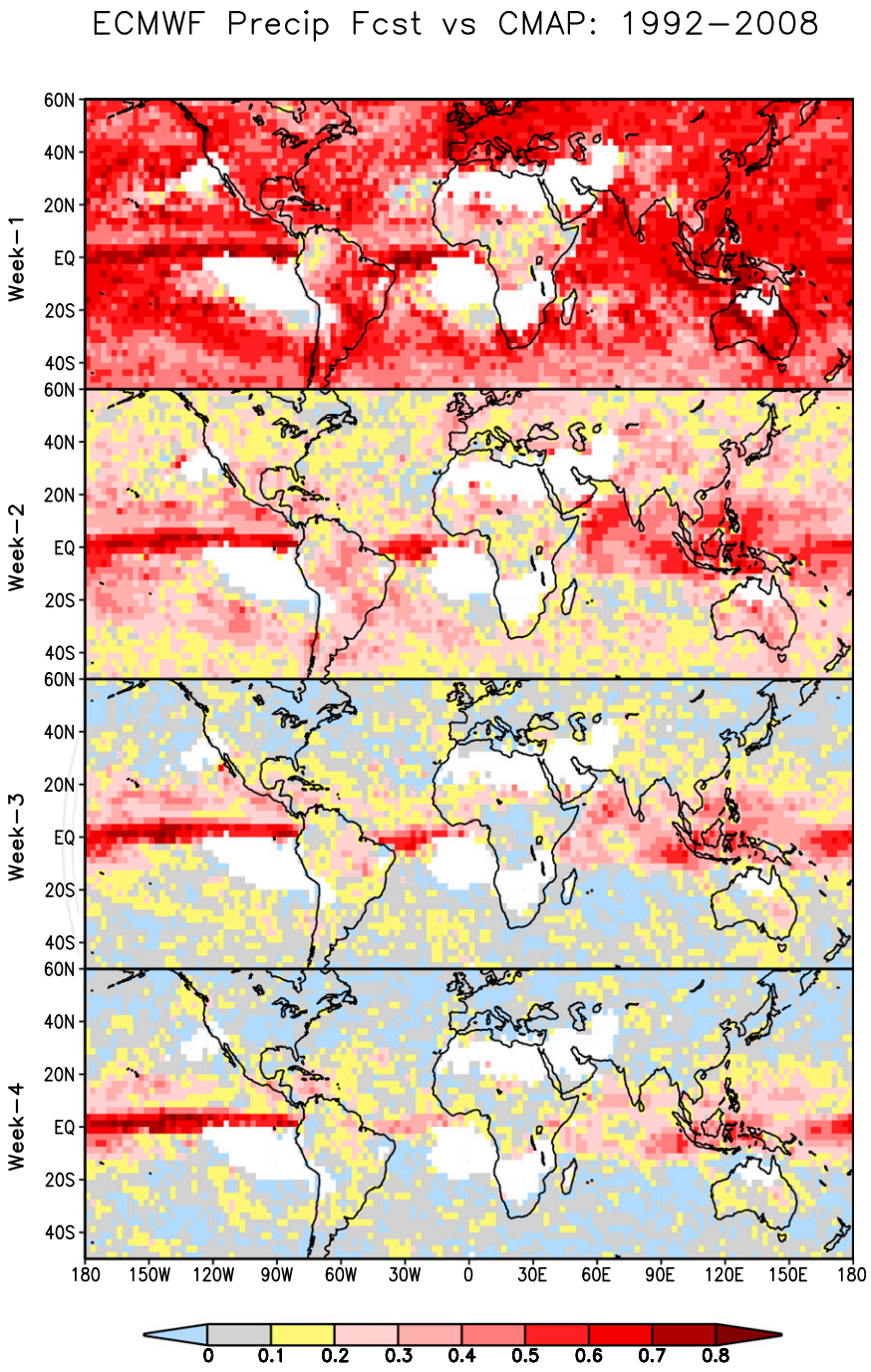

FIG. 3. As in Fig. 1, but for the ECMWF hindcast.

exists compared to the JMA and CFSv2 hindcasts. These dominant characteristics will be further discussed in the next section based on aggregate CORA skill over the tropical land area only and a specific region of interest, the Maritime Continent.

Skill levels over the continents are generally disappointing especially in weeks 3–4, although the ECMWF model does exhibit substantial skill at week 2 over South America, Eurasia, and Australia. The rapidly decreased skill level beyond one week in the extratropics is generally consistent with the skill analysis using the Predictive Ocean Atmosphere Model for Australia (POAMA) coupled system (Zhu et al. 2014). Skill levels over Africa are poor even in week 1, which suggests that the observational data quality may be poor on the weekly time scale. This could also be due to less predictability in that

JOBNAME: MWR 00#00 2015 PAGE: 7 SESS: 8 OUTPUT: Wed May 13 11:56:25 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

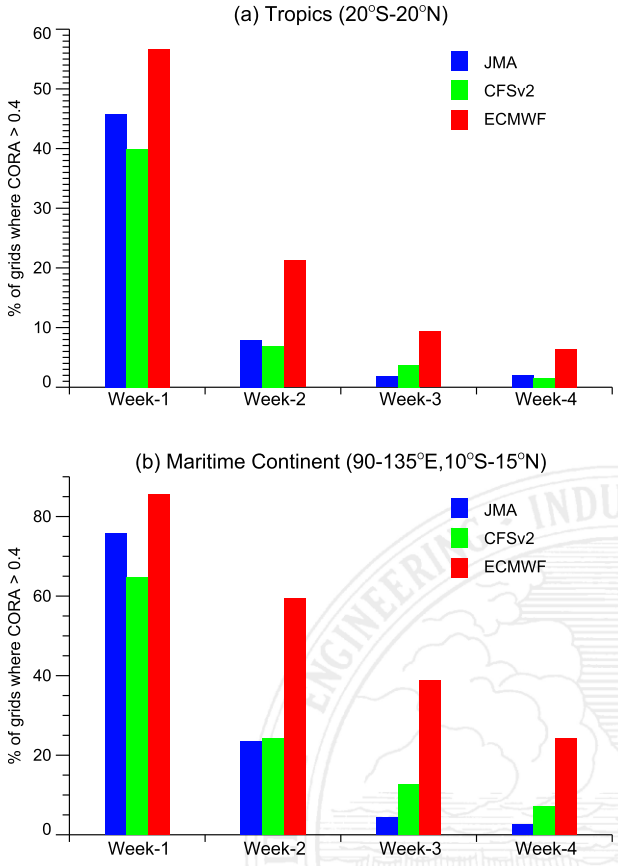

FIG. 4. Aggregate correlation of anomalies between model precipitation hindcast and CMAP rainfall data for weeks 1–4, represented by percentage of CORA exceeding the threshold of 0.4. (a) Tropical land areas only (20°S–20°N) and (b) Maritime Continent (10°S–15°N, 90°–135°E) land areas only. Areas excluded by the dry mask (see text) are excluded here also, but the difference is minimal.

region, with less impact from large-scale variability of ENSO, MJO, etc., than in other tropical regions.

The sharp skill drop-off after week 1, particularly in the extratropics, together with the persistence of skill over the equatorial Pacific and Atlantic Oceans in weeks 2–4 are consistent with the studies of Kumar et al. (2011) and Zhu et al. (2014). The roles of ENSO and the MJO in the latter are further investigated in section 3b below.

2) AGGREGATE CORA OVER SPECIFIC REGIONS

F4 Figure 4 shows the spatially aggregated CORA values for all the three models over tropical land areas within 20° latitudes, and Maritime Continent land points (10°S–15°N, 90°–135°E), represented by percentage of grid points where the CORA exceeds a threshold value of 0.4 (Li et al. 2008). The near exponential drop-off of skill after week 1 is clear for both regions, with the striking exception of the ECMWF over the Maritime Continent where the drop-off is rather uniform. The aggregate skill from ECMWF hindcast outperforms the other two models in all cases, and particularly for weeks 2–4 over the Maritime Continent. The near-exponential drop-off of skill over tropical land areas contrasts with the notable areas of skill in weeks 2–4 over the tropical oceans seen in all models in Figs. 1–3. When considering topical ocean areas, Kumar et al. (2011) also found much more persistence in skill levels of monthly averaged precipitation at lead times from 10–80 days (especially over the tropical Pacific), attributed to ENSO.

The much slower drop-off of skill in the ECMWF result over the Maritime Continent stands in contrast to the general expectation of loss of deterministic predictability from atmospheric initial conditions after week 1. A small area over Borneo Island will be further investigated in section 3b, along with exploring possible sources of submonthly predictability with a focus on this area as a case study.

b. *Sources of submonthly predictability*

While submonthly predictability may arise from various phenomena as discussed in the introduction, our attention here is devoted to the impact of the MJO and ENSO. We first examine the boreal summer teleconnections of each phenomenon in our CMAP dataset globally, and then focus on a small region of the Maritime Continent where the skill is particularly high (i.e., Borneo Island).

1) ENSO AND MJO TELECONNECTIONS IN CMAP PRECIPITATION

F5 Figure 5 shows the correlation between precipitation from CMAP and an ENSO index [defined as the SST anomaly over the Niño-3.4 region (5°S–5°N, 120°–170°W)] for June–July–August (JJA) during the 1992–2008 period. The correlation map for the three-month season, with a 95% significance mask, indicates a pattern of high positive correlations over the equatorial Pacific, in general correspondence to the relatively high CORA of the model hindcasts seen there in Figs. 1–3. Also evident is the negative, but strong correlations over the Maritime Continent, as well as the eastern Indian Ocean and the tropical Atlantic. These statistically significant features imply possible important relationships between ENSO and the submonthly predictability of precipitation. Across-the-season modulation of ENSO teleconnections are quite pronounced in some regions such as the tropical North Atlantic (Enfield and Mayer 1997), which suggest AU5 that there may be seasonal modulation of model skill as well. This issue is left for a future study.

To further examine the impact of ENSO on the skill of submonthly precipitation forecasts, we compute the CORA skill of the ECMWF model separately for five

JOBNAME: MWR 00#00 2015 PAGE: 8 SESS: 8 OUTPUT: Wed May 13 11:56:25 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

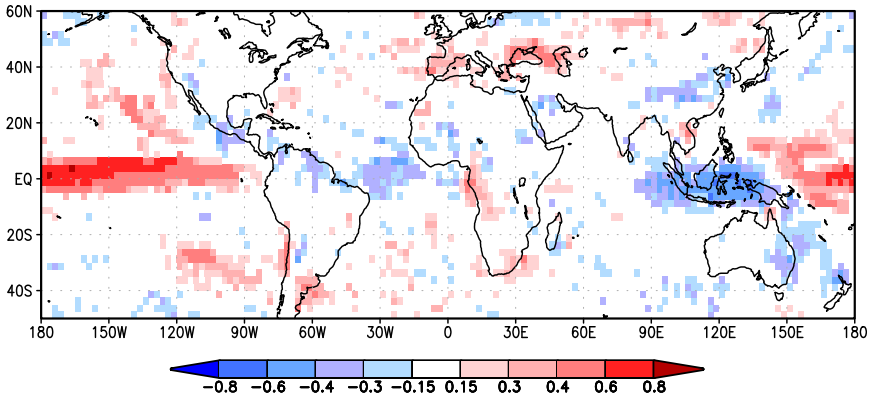

FIG. 5. Correlation between CMAP precipitation and Niño-3.4 index for June–August, 1992–2008, with a 95% significance mask.

F6 ENSO years versus five neutral years. Figure 6 shows CORA maps for the lead-3 ECMWF hindcasts for the two sets of years. The p value for a CORA of 0.2 is 0.059 for a sample size of 90 (5 years × 18 hindcasts per year), so the pink shaded areas in Fig. 6 are still highly significant. The CORA skill for the ENSO years is indeed generally much higher than for the neutral years, except over the equatorial Atlantic where modes of tropical

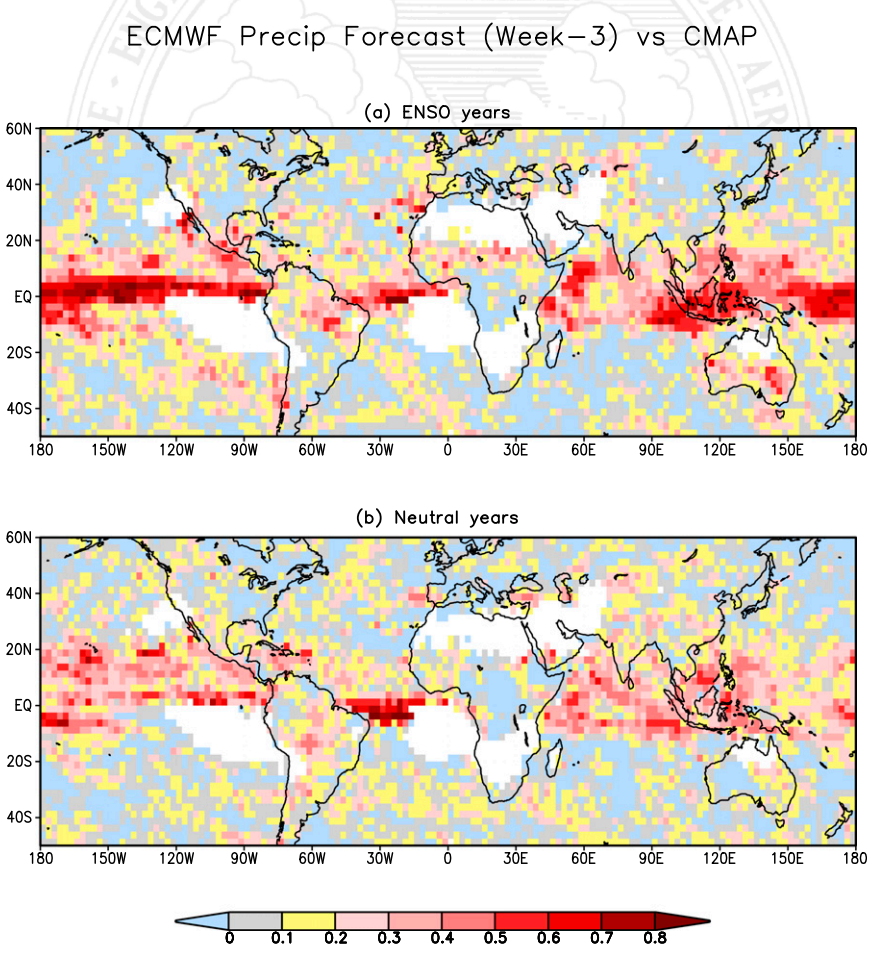

FIG. 6. Correlation of anomalies between ECMWF lead-3 precipitation hindcasts (over 18 start dates) and CMAP observation during (a) five ENSO years: 1997–2000, 2002; and (b) five neutral years: 1992, 2001, 2003, 2005, and 2008.

JOBNAME: MWR 00#00 2015 PAGE: 9 SESS: 8 OUTPUT: Wed May 13 11:56:26 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

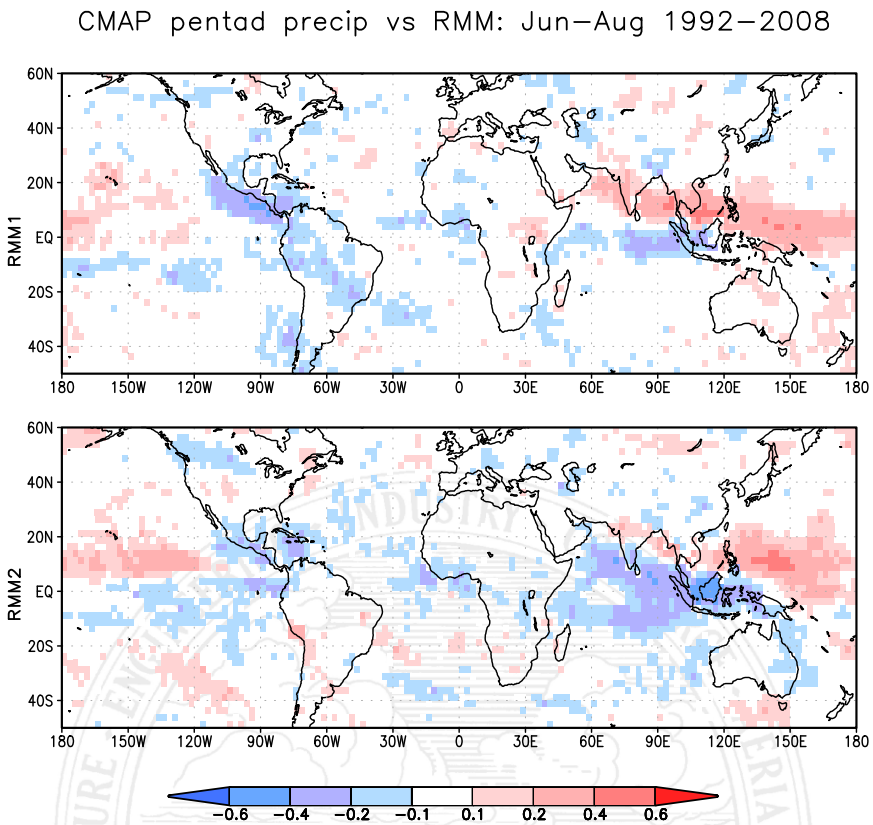

FIG. 7. Correlation between CMAP pentad precipitation and 5-day mean Real-time Multivariate MJO (RMM) components, RMM1 and RMM2, during June–August, 1992–2008, with a 95% significance mask.

Atlantic variability may be important contributors to the submonthly skill. There is slightly higher CORA skill over landmasses such as south-southeastern Australia, Sahelian Africa, and northern Brazil during ENSO years, plausibly associated with known ENSO teleconnections in those regions. The western Indian Ocean is seen to be a region of skill independent of ENSO, possibly due to persistent SST anomalies there; this is particularly marked in the ECMWF model (cf. Figs. 1–3).

F7 Figure 7 shows the correlation between CMAP pentad precipitation and the Real-time Multivariate MJO (RMM) indices defined by Wheeler and Hendon (2004): RMM1 and RMM2. These represent the two leading principal components of intraseasonal near-equatorially averaged outgoing longwave radiation (OLR) and zonal winds. A correlation value of 0.1 has a p value of 0.081 for a sample size of 18 samples per JJA season $\times$ 17 years. Strong negative correlations exist for both RMM1 and RMM2 over the Maritime Continent and equatorial Indian Ocean, consistent with the pattern of the summertime MJO (Yoo et al. 2010), also called the boreal summer intraseasonal oscillation (BSISO; Lee et al. 2013). Both RMM1 and RMM2 also exhibit significant positive correlations over the western equatorial Pacific. Of these regions of significant MJO teleconnections in CMAP precipitation, only the Maritime Continent is a region of significant CORA skill of the three EPS systems beyond week 2 (e.g., Fig. 3). This MJO influence may also partly account for the skill of the ECMWF system in week 2 over Central and South America, but other sources of predictability such as ENSO may be at play as well. Vitart and Molteni (2010) showed that a similar set of ECMWF reforecasts produced realistic MJO–precipitation teleconnections in the tropics, in particular over Central America, South America, and central Africa.

2) BORNEO ISLAND

F8 Turning to the Maritime Continent, Fig. 8 shows a scatterplot of the CMAP precipitation anomalies, over a portion of Borneo Island (four grid boxes 1.25°S–1.25°N, 111.25°–113.75°E), versus the ECMWF lead-3 hindcasts, valid for 18 weeks from 4–10 June to 1–7 October during the 1992–2008 period. There is a near-linear relationship despite large scatter with a correlation of 0.52. The week-2 hindcasts (valid for weeks 28 May–3 June

JOBNAME: MWR 00#00 2015 PAGE: 10 SESS: 8 OUTPUT: Wed May 13 11:56:26 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

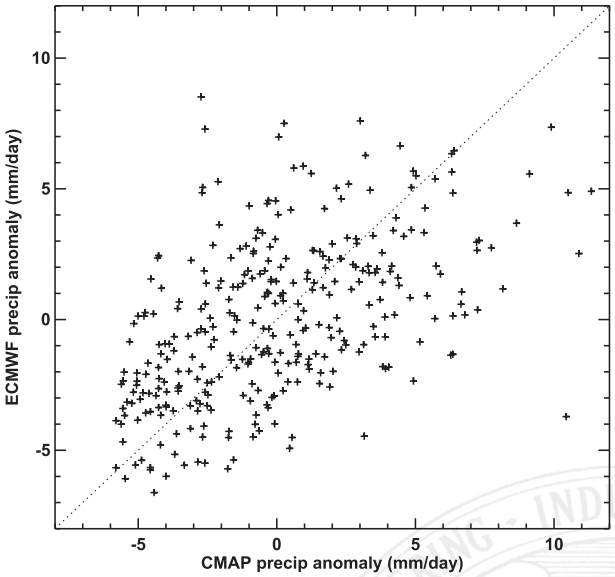

FIG. 8. Scatterplot of CMAP weekly precipitation anomaly vs ECMWF hindcast (lead 3) valid for 18 weeks (from 4–10 Jun through 1–7 Oct), 1992–2008, over the Borneo Island. The dotted line is a 45° line.

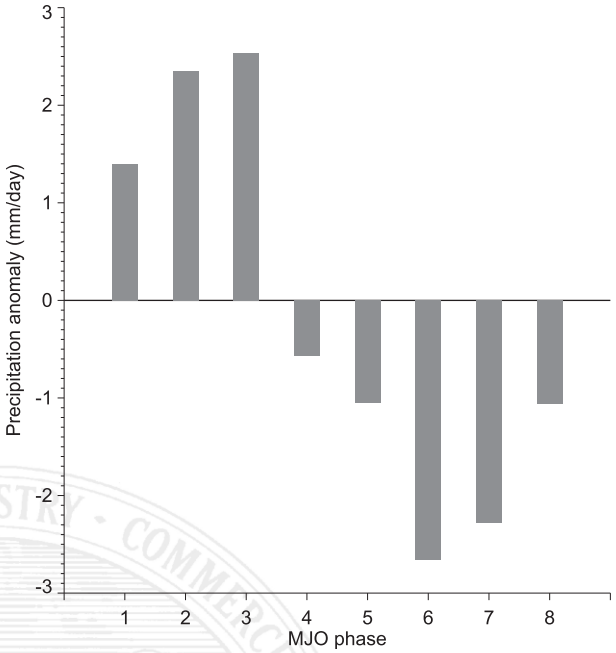

FIG. 9. Composite Borneo-averaged pentad CMAP precipitation anomaly in relation to pentad-dominant MJO phase during June–August, 1992–2008.

through 24–30 September) exhibit reduced scatter (not shown), with a correlation of 0.62.

By compositing the pentad CMAP rainfall anomaly over Borneo (averaged over four grid boxes), for each MJO phase (1–8) during boreal summer, 1992–2008, we make the MJO phase composites of Borneo-averaged rainfall anomalies, shown in Fig. 9. Phases 2–3 of MJO correspond to enhanced Borneo rainfall, while phases 6–7 are associated with greatly reduced precipitation over Borneo. This agrees with the correlation between Borneo rainfall and the RMM indices (Fig. 7), in particular RMM2. However, perhaps surprisingly, in terms of the Wheeler–Hendon MJO phase-portrait diagrams, the active phase of the MJO over Borneo corresponds predominantly to the "Indian Ocean" sector of the diagram, while the suppressed rainfall phases over Borneo correspond to the "western Pacific" sector (phases 6–7). A sample MJO phase portrait is given in Fig. 10, which shows the observed MJO RMM1–RMM2 phase and strength for June–September 2002 as a typical El Niño year. The MJO was very active with an amplitude greater than 1.0 standard deviation during most of the June–August period. A strong MJO event propagated across the Indian Ocean and Maritime Continent (mostly phases 2–4) in the middle of June which, according to Fig. 9, would already be felt in rainfall over Borneo in phases 1–3. This result is consistent with Peatman et al. (2014), who interpret this accelerated MJO influence as a vanguard of precipitation that jumps ahead of the main body of MJO convection. Moron et al. (2015) recently found that a quiescent large-scale circulation type is prevalent over the Maritime Continent in MJO phases 1–3, when the MJO convective envelope is over the Indian Ocean, and that this allows the diurnal cycle

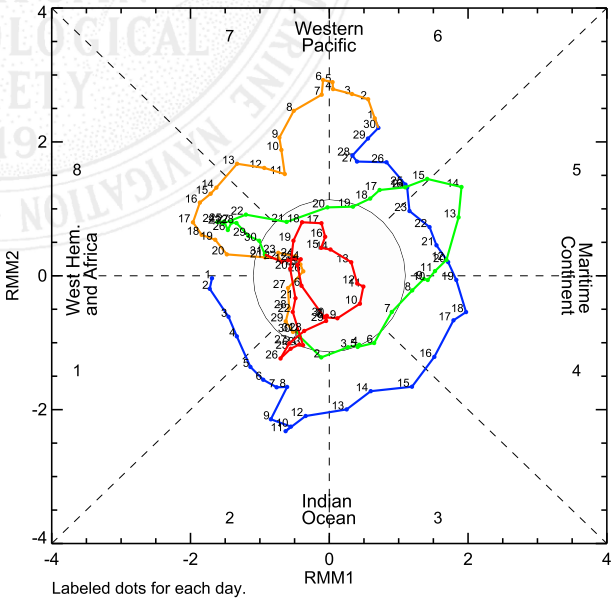

FIG. 10. Real-time Multivariate MJO (RMM) phase space (Wheeler and Hendon 2004) from 1 Jun through 30 Sep 2002, a typical year of El Niño.

JOBNAME: MWR 00#00 2015 PAGE: 11 SESS: 8 OUTPUT: Wed May 13 11:56:26 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

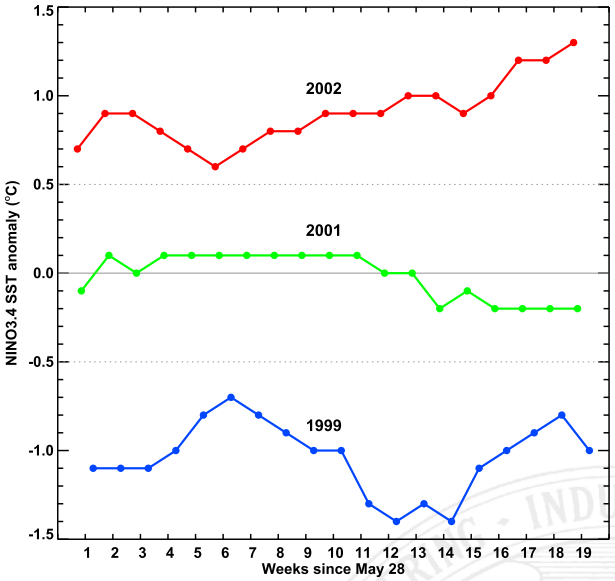

FIG. 11. Time series of observed weekly Niño-3.4 index during late May through late September for 2002, 1999, and 2001 as sampling years of El Niño, La Niña, and neutral-ENSO conditions, respectively. The *x* axis denotes week numbers since 28 May, but the ENSO index has a minor shift of −2, −1, and +2 days for those three years.

in rainfall to amplify over the islands, leading to the positive rainfall anomalies at that time, seen in Fig. 9.

We next select three typical ENSO years during boreal summer 1992–2008 to represent El Niño (2002), La Niña (1999), and neutral ENSO (2001), respectively, and examine the evolution of CMAP versus the ECMWF F11 hindcasts averaged over Borneo. Figure 11 shows weekly time series of the observed Niño-3.4 index for the three years, indicating moderate warm ENSO conditions increasing through the boreal summer of 2002, strong cold ENSO conditions in 1999 with some subseasonal modulation, and near-zero Niño-3.4 anomaly values in 2001.

F12 Figure 12a shows time series of weekly precipitation anomalies from weeks 2 and 3 of the ECMWF hindcasts and observed rainfall from CMAP, averaged for the four grid boxes over Borneo Island in 2002. Note that no bias correction of these anomalies was done beyond subtraction of the respective climatology. The verification period is from 28 May–3 June through 24–30 September (i.e., 18 weeks in total, for week 2, and from 4–10 June through 1–7 October for week 3). The uppercase letters, along with the MJO phase number, at the bottom of each panel denote the dominant MJO phase (amplitude ≥ 1.0) for the corresponding 7-day periods. That is, "A" for the Western Hemisphere and Africa (phases 1 and 8), "I" for the Indian Ocean (phases 2 and 3), "M" for the Maritime Continent (phases 4 and 5), and "P" for the western Pacific (phases 6 and 7). As shown in Fig. 12a, there is an

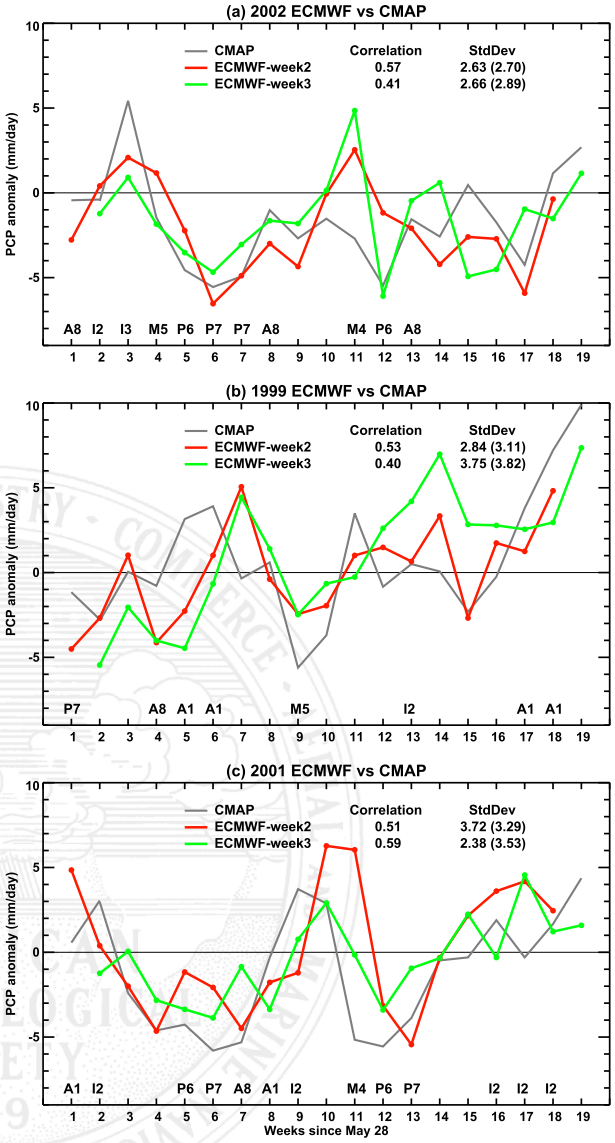

FIG. 12. Time series of precipitation anomalies (mm day^{-1}) over a portion of Borneo Island, from the CMAP data and ECMWF hindcast, valid for weeks 2 and 3 during June–September for (a) 2002, (b) 1999, and (c) 2001. The uppercase letter along with the single digit at the bottom of each panel denote the dominant MJO phase sector and phase number (see the text for details), with amplitude greater than 1.0 (Wheeler and Hendon 2004). Along the abscissa, weeks 1–5 correspond approximately to June; weeks 6–10 correspond to July; weeks 11–14 correspond to August; weeks 15–18 correspond to September; and week 19 corresponds to 1–7 Oct. Correlation values with CMAP are given in each panel, together with the respective standard deviations, with the CMAP values in parentheses.

encouragingly good correspondence of rainfall anomalies between CMAP and the ECMWF hindcasts. For this El Niño year, the CMAP rainfall anomalies are mostly negative as expected, and this is captured by the model at both leads. However, the observed rainfall becomes far-above

JOBNAME: MWR 00#00 2015 PAGE: 12 SESS: 8 OUTPUT: Wed May 13 11:56:26 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

average for the third week (11–17 June), when the MJO is dominantly over the Indian Ocean, or phases 2 to 3, with a strength of around 2.0 (Fig. 10). This excursion was captured by the model at both leads, although it was too weak, especially at week 3. Note that there are no other instances of strong MJO phases 2–3 during the summer of 2002, and the CMAP anomalies remain negative for most of the season. The model does, however, go positive near week 11. The El Niño strength is evolving from weak to moderate during the season, with the Niño-3.4 index ranging from 0.6° to 1.3°C, and a tendency of a strengthening El Niño signal after early July (Fig. 11).

For the La Niña year 1999, the weekly rainfall anomalies are mostly near zero, and positive late in the season (Fig. 12b), in response to the La Niña impact; this intensifies later toward the fall as the event strengthens. However, during the ninth week (23–29 July), CMAP rainfall was nearly 5.5 mm day^{-1} below average, while the MJO was in phase 5, with a strength of about 1.0, consistent with Fig. 9. This negative excursion was captured in the model hindcasts, although again with a weaker amplitude than observed. The MJO becomes weak by early and mid-August (weeks 11–14) after going into phases 6 and 1. The Niño-3.4 index remains moderate through the season but slightly stronger during mid-August to early September, with a growing La Niña signal from early July to late August, as seen at the bottom of Fig. 11.

For the case of neutral ENSO year 2001 (Fig. 12c), the CMAP rainfall exhibits large intraseasonal fluctuations, quite symmetrically about zero where less precipitation corresponds to MJO active phase over the western Pacific (phase 6 or 7) and above-average rainfall tends to coincide with MJO active phases over the Indian Ocean (e.g., phase 2). These rainfall anomalies are remarkably well captured by the ECMWF model at both weeks 2 and 3, and the attribution of the model's skill to the MJO is clear, although there is some phase delay in the model. The Niño-3.4 anomaly index for the season is mostly near zero, ranging between −0.2° and 0.1°C (Fig. 11). Again, the Borneo rainfall anomalies correspond well with the evolution of MJO (dominant) phases as manifested in Fig. 9. For these three typical ENSO cases, the model hindcasts generally underestimate the larger observed precipitation anomalies. An amplitude bias in the ensemble-mean hindcasts is to be expected, even though the standard deviations (given in the figure panels) are generally only slightly underestimated. As more ensemble members are added, the ensemble mean will start to filter out the unpredictable weather noise and the remaining signal will have a smaller magnitude. As long as there is noise, the variance of the ensemble mean will be less than that of the observations (ensemble member) and the amplitude of the ensemble mean is less than that of the observations (DelSole and Tippett 2007; Tippett et al. 2010). AU6

c. *Mean square skill score*

The correlation of anomalies (CORA) is a scale-invariant measure of the linear association between the predicted anomalies and the observations (e.g., Murphy 1988). Therefore, it can only provide a measure of relative association. The MSSS is another important skill metric used to assess deterministic forecasts, based on the mean squared error; it measures the accuracy of the forecast relative to a reference set of forecasts, taken here to be climatological forecasts, given by the climatological average rainfall for a given weekly average and grid point (Murphy 1988). As in the case of CORA, we subtract the lead and start-date-dependent climatological average from both the forecasts and the observations, so that the mean forecast bias is zero by construction. The maximum MSSS here is 1.0, and negative values imply that the forecast is less accurate than the reference "forecast" (i.e., climatology) indicating no skill or large conditional biases. Figure 13 shows global maps of MSSS from the JMA hindcasts over the same period and starts as for CORA. In general, there is skill for the first week over the western Pacific Ocean, eastern tropical Indian Ocean, and western Europe, but the skill decreases rapidly for weeks 2–4, as seen for the CORA. Similar patterns of MSSS skill can also be seen for the CFSv2 precipitation hindcast (Fig. 14), even though it shows generally lower skill compared to the JMA hindcast, suggesting larger forecast conditional biases. F13 F14

Like the intercomparison of CORA among these three models, the MSSS from ECMWF hindcast (Fig. 15) indicates better skill than the other two models. Substantial skill exists for the first week over most of tropical ocean/land and even the northern midlatitudes, but the skill drops considerably from weeks 2 to 4, as seen for the CORA. In particular, it demonstrates better skill for the first week, persisting over the tropics for week 2 and over the eastern equatorial Indian Ocean through weeks 3 and 4. The ECMWF model is the only one of the three in which the strip of positive skill values in the eastern and central equatorial Pacific ITCZ region seen in the CORA maps for weeks 2–4 is also reflected in the MSSS. However, in common with the other models, the area of positive CORA in the Atlantic ITCZ does not appear in the ECMWF model either. F15

The MSSS can be decomposed into the square of the correlation coefficient, together with the squares of the conditional and mean prediction biases (Murphy 1988). Since the mean biases are subtracted at the outset, the MSSS in Figs. 13–15 consists of only the $(\text{CORA})^2$ and

JOBNAME: MWR 00#00 2015 PAGE: 13 SESS: 8 OUTPUT: Wed May 13 11:56:26 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

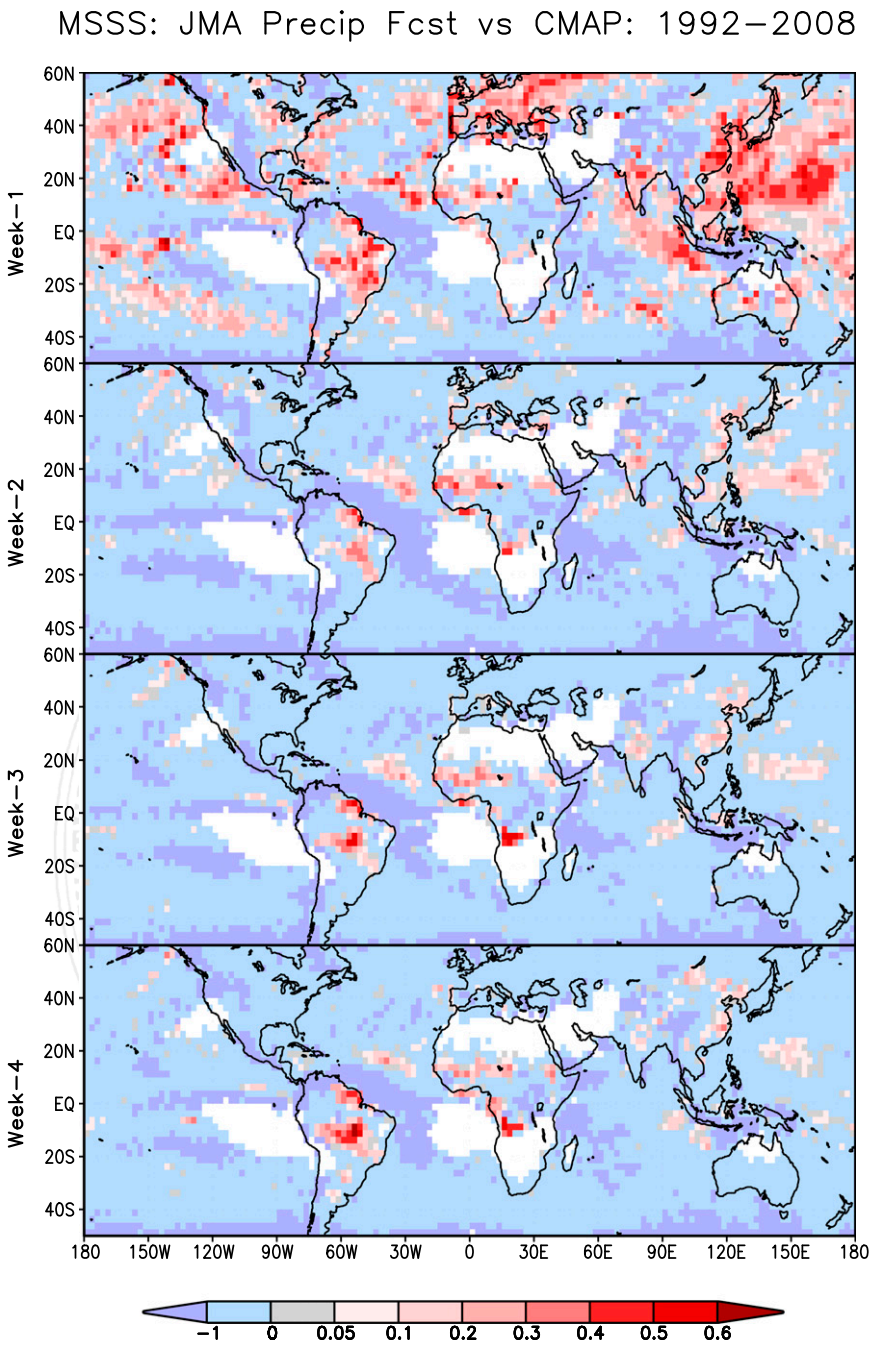

FIG. 13. Mean square skill score (MSSS) between the JMA model precipitation hindcast and CMAP rainfall data over weeks 1–4.

the squared conditional bias. The latter (not squared) is shown for all three prediction systems for week 3 in F16 Fig. 16. This was computed as (CORA − s_h/s_o) following Murphy (1988), where s_h and s_o are the standard deviations of the hindcasts and the observations, respectively. The conditional bias is mostly negative, implying that the variance of the hindcasts is too large relative to the observations and the correlation coefficient. (Goddard et al. 2013). The regions of negative MSSS in Figs. 13–15 correspond to regions of large conditional bias, in which the hindcasts are penalized if s_h does not become small where there is no CORA skill. Overall, the CFSv2 hindcast corresponds to the largest conditional biases, JMA to the second, and

JOBNAME: MWR 00#00 2015 PAGE: 14 SESS: 8 OUTPUT: Wed May 13 11:56:26 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

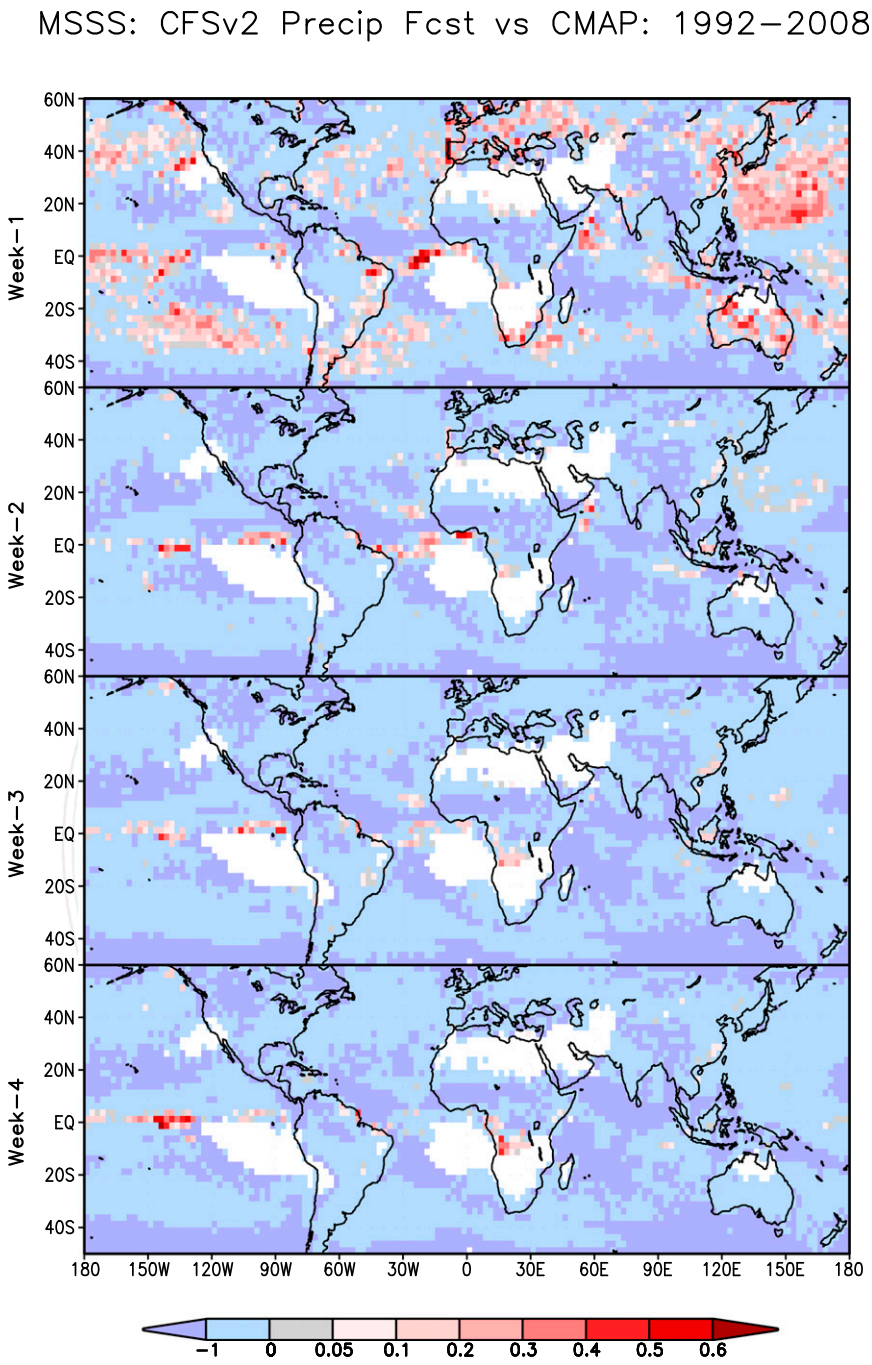

FIG. 14. Mean square skill score (MSSS) between CFSv2 precipitation hindcast and CMAP rainfall data over weeks 1–4.

ECMWF to the least out of the three. The strip of positive CORA skill values in the eastern and central equatorial Pacific ITCZ region for weeks 2–4 that disappears in the MSSS of the JMA and CFSv2 is consistent with larger conditional biases seen in those models there, compared to the ECMWF model (Fig. 16). This is a region of strong persistence in SST anomalies associated with ENSO that may impact the CORA metric more strongly.

The picture that emerges of submonthly accuracy of the three submonthly EPS systems is clearly less rosy in terms of MSSS than CORA, even when mean biases are subtracted. All three systems exhibit large conditional biases (strongly negative MSSS) in the deep convection

JOBNAME: MWR 00#00 2015 PAGE: 15 SESS: 8 OUTPUT: Wed May 13 11:56:27 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

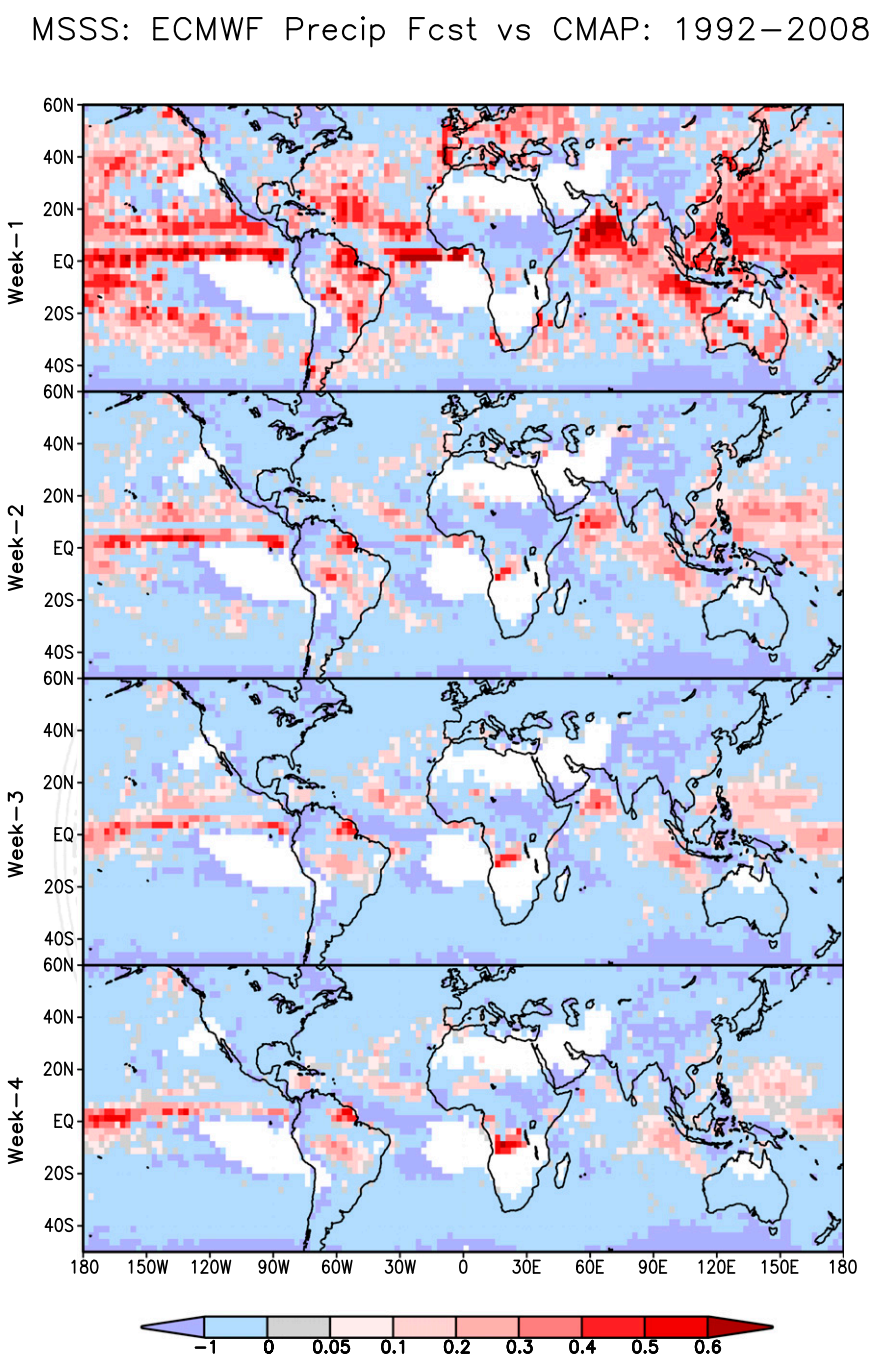

FIG. 15. Mean square skill score (MSSS) between ECMWF precipitation hindcast and CMAP rainfall data over weeks 1–4.

regions over the tropical Atlantic and over the Maritime Continent. There are also large biases over the Sahel in CFSv2, and in ECMWF over the eastern Sahel, while the JMA does better over these convective land regions. The conditional bias is likely to be exacerbated by the small ensemble size (five members for ECMWF and JMA, and four members for CFSv2). For a single ensemble member with realistic variance, (CORA $- s_h/s_o$) will be negative (unless CORA is perfect). As more members are added, the variance of the ensemble mean will get smaller and make the conditional bias less negative (assuming the correlation remains much the same or improves with the use of more members). (Recalculating the MSSS for the

JOBNAME: MWR 00#00 2015 PAGE: 16 SESS: 8 OUTPUT: Wed May 13 11:56:27 2015 Total No. of Pages: 19
/ams/mwr/0/mwrD1400277

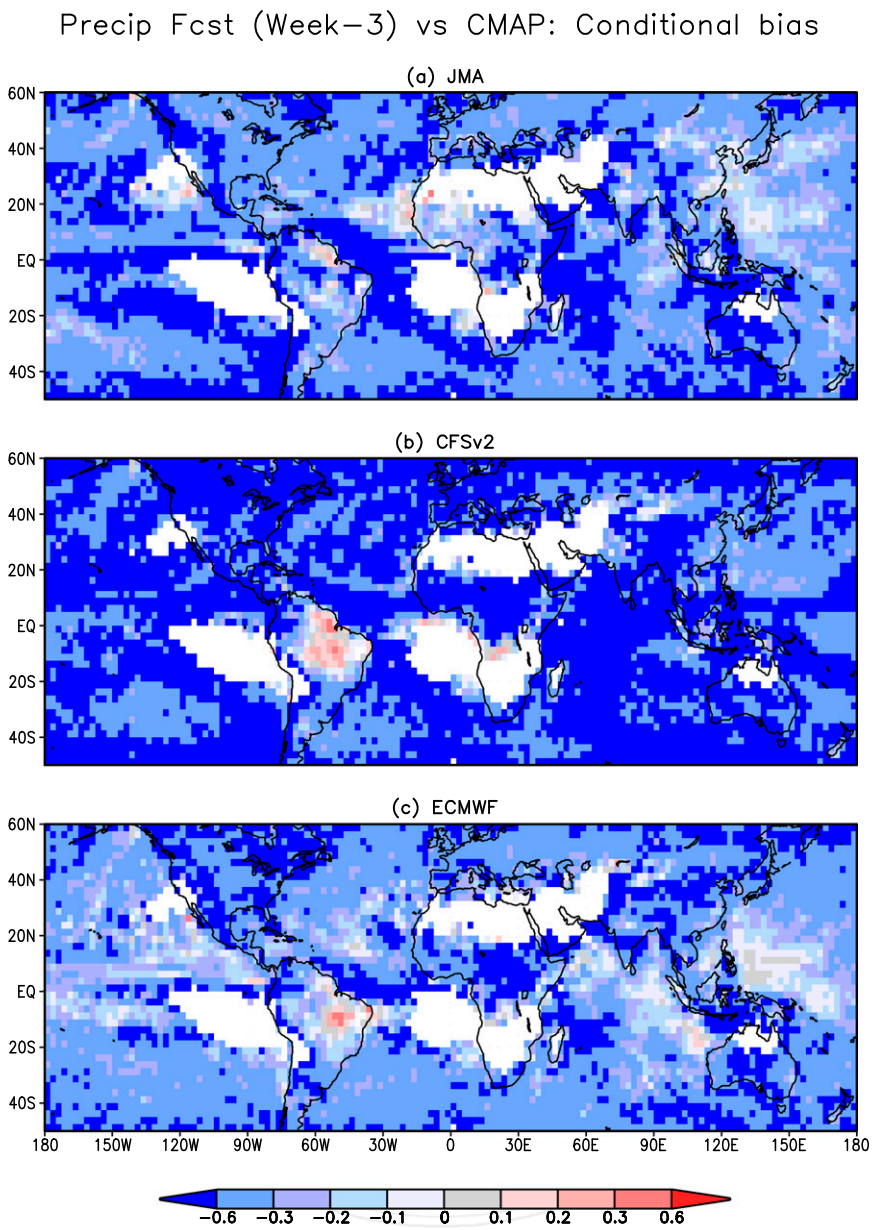

FIG. 16. Conditional bias of precipitation hindcasts vs CMAP rainfall data valid for week 3: (a) JMA, (b) CFSv2, and (c) ECMWF.

ECMWF model with four instead of five members does not make a visible difference to Fig. 15, indicating that the larger conditional biases in the CFSv2 are not likely to be the result of its smaller ensemble size of four versus five members.) Since the ensemble sizes of hindcasts from subseasonal EPS systems are generally much smaller than those of the real-time forecasts [cf. annex 3 in World Meteorological Organization (2013)], the MSSS skills calculated from hindcasts are likely to considerably underestimate the MSSS.

4. Summary and discussion

This study has compared the performance of sub-monthly precipitation hindcasts based on three ensemble prediction systems against objectively analyzed precipitation data. The model predictions valid for four consecutive weeks (weeks 1–4 of the forecasts), composited from 13, 18, and 25 start dates, respectively (Table 1), are evaluated using processed weekly precipitation datasets during late May through mid-September over the 17-yr period 1992–2008. The prediction skill is quantified using

CORA and MSSS deterministic skill metrics over the globe, as well as the tropical land areas and the Maritime Continent.

In general, all three model hindcast sets indicate very good skill for the first week, but dramatically decreased skill for weeks 2–4, with the exception of some tropical ocean areas, particularly over the Pacific and Atlantic ITCZs, but also over the Maritime Continent. Among the three EPS models, the ECMWF forecast system demonstrates noticeably better skill than the other two, especially for weeks 3 and 4. This comparison applies to the prediction skill as manifested by both CORA and MSSS, suggesting that the ECMWF hindcast corresponds to not only the best correlation skill, but also the combined measure of relative (linear) association and the biases inherent in the forecasts. The comparison between the CORA and MSSS metrics indicates that conditional biases in precipitation weekly averages can be large in all the EPS systems, especially over the deep convection regions of the Atlantic ITCZ and the West African monsoon and the Sahelian region, but also over the Maritime Continent. These biases are particularly large in the CFSv2 model, which warrants further investigation.

The predictability of submonthly precipitation was interpreted in connection with the low-frequency ENSO variability and intraseasonal MJO phase/strength. Such associations were examined qualitatively by selecting three typical years of neutral, El Niño, and La Niña events in correspondence with weekly time series of precipitation anomalies over a portion of the Maritime Continent (Borneo Island), along with the corresponding dominant phase of MJO. Our results suggest that the rainfall variation for neutral-ENSO years corresponds well with the dominant MJO phase (e.g., above-average rainfall in response to the phase over the Indian Ocean); whereas for the situation with moderate/strong ENSO events, the relationship of rainfall anomaly with the MJO appears to become weaker, while the contribution of ENSO to the submonthly skill is substantial. However, if a moderate/strong MJO event crosses over the Indian Ocean near the Maritime Continent (e.g., phases 2–3 with strength of greater than 2.0), the rainfall tends to be far above-average for typical El Niño years, such as 2002. For the La Niña year 1999, below-average rainfall is observed when a moderate MJO is over phases 5–6 during late July. In such cases, the MJO impact or modulation becomes significant, regardless of whether the ENSO event is strong or moderate. Neena et al. (2014) have recently documented MJO predictability and skill from eight EPS systems during boreal winter, including variants of the models examined here. They found the ensemble mean hindcasts to have skill up to 15–25 days, while their estimates of the MJO predictability limit from intra-ensemble behavior to be 5–10 days more, implying that future improvements of the EPS's could lead to higher forecast skill.

The advantage of multimodel ensemble (MME) combination forecast has been documented by numerous studies (e.g., Robertson et al. 2004; Li et al. 2008). It has been acknowledged that the overall prediction skill from MME products, at least for seasonal forecasting, outperforms the best individual model. Our study here was limited to model-to-model analysis. The start dates are nonuniform among the three EPS systems, which makes it difficult to verify the performance of MME forecast and the construction of subseasonal multimodel ensembles is left for future work.

The association of rainfall predictability with ENSO and MJO is only analyzed here for a portion of the Maritime Continent and for a few typical ENSO years, and it remains to be confirmed whether our findings extend to other continents and years. Further, it is still challenging to quantify the relative contribution of ENSO and MJO to the predictability of precipitation, since ENSO is confined to a fixed location, while the MJO is a propagating wave with variable amplitude and phase. Nonetheless, these results support the concept that "windows of opportunity" of high forecast skill exist as a function of ENSO and the MJO in certain locations and seasons, which may lead to subseasonal-to-seasonal forecasts of substantial societal value in the future. It is hoped that the results presented here will be of value to the recently begun "Sub-seasonal to Seasonal Prediction Project (S2S)" coordinated jointly by the World Weather Research Program (WWRP) and World Climate Research Program (WCRP) (Vitart et al. 2012).

Acknowledgments. It is a pleasure to acknowledge helpful comments from Y. Takaya and F. Vitart. The paper also benefitted substantially from the constructive reviews of D. Waliser and three anonymous reviewers. We gratefully acknowledge the three modeling centers (ECMWF, JMA, and NCEP) for making their hindcast data available. We thank in particular F. Vitart and F. Molteni for help in accessing the ECMWF data, and to J. del Corral for postprocessing in the IRI data library. This work was funded by grants to the IRI from the U.S. Agency for International Development (AID-OAA-A-11-00011) and the National Oceanic and Atmospheric Administration (NA050AR4311004).

REFERENCES

Baldwin, M. P., and T. J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science,* **294,** 581–584, doi:10.1038/news011025-4.

Barnston, A. G., and M. K. Tippett, 2013: Predictions of Nino3.4 SST in CFSv1 and CFSv2: A diagnostic comparison. *Climate Dyn.,* **41,** 1615–1633, doi:10.1007/s00382-013-1845-2.

Charney, J. G., and J. Shukla, 1981: Predictability of monsoons. *Monsoon Dynamics,* J. Lighthill and R. P. Pearce, Eds., Cambridge University Press, 99–109.

DelSole, T., and M. K. Tippett, 2007: Predictability: Recent insights from information theory. *Rev. Geophys.,* **45,** RG4002, doi:10.1029/2006RG000202.

Enfield, D. B., and D. A. Mayer, 1997: Tropical Atlantic sea surface temperature variability and its relation to El Niño–Southern Oscillation. *J. Geophys. Res.,* **102,** 929–945, doi:10.1029/96JC03296.

Goddard, L., and Coauthors, 2013: A verification framework for interannnual-to-decadal predictions experiments. *Climate Dyn.,* **40,** 245–272, doi:10.1007/s00382-012-1481-2.

Griffies, S. M., M. J. Harrison, R. C. Pacanowski, A. R. Rosati, Z. Liang, M. Schmidt, H. Simmons, and R. Slater, 2004: A technical guide to MOM4. GFDL Ocean Group Tech. Rep. 5, 342 pp. [Available online at www.gfdl.noaa.gov/bibliography/related_files/smg0301.pdf.]

Hirons, L. C., P. Inness, F. Vitart, and P. Bechtold, 2013: Understanding advances in the simulation of intraseasonal variability in the ECMWF model. Part I: The representation of the MJO. *Quart. J. Roy. Meteor. Soc.,* **139,** 1417–1426, doi:10.1002/qj.2060.

Holland, M. M., D. A. Bailey, and S. Vavrus, 2011: Inherent sea ice predictability in the rapidly changing Arctic environment of the Community Climate System Model, version 3. *Climate Dyn.,* **36,** 1239–1253, doi:10.1007/s00382-010-0792-4.

Hoskins, B., 2013: The potential for skill across the range of the seamless weather-climate prediction problem: A stimulus for our science. *Quart. J. Roy. Meteor. Soc.,* **139,** 573–584, doi:10.1002/qj.1991.

Japan Meteorological Agency, 2013: Outline of the Operational Numerical Weather Prediction at the Japan Meteorological Agency. Appendix to WMO Tech. Progress Rep. on the Global Data-Processing and Forecasting System and Numerical Weather Prediction, 188 pp. [Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2013-nwp/index.htm.]

Koster, R. D., and Coauthors, 2010: Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophys. Res. Lett.,* **37,** L02402, doi:10.1029/2009GL041677.

Kumar, A., M. Chen, and W. Wang, 2011: An analysis of prediction skill of monthly mean climate variability. *Climate Dyn.,* **37,** 1119–1131, doi:10.1007/s00382-010-0901-4.

Lee, J. Y., B. Wang, M. C. Wheeler, X. Fu, D. E. Waliser, and I.-S. Kang, 2013: Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Climate Dyn.,* **40,** 493–509, doi:10.1007/s00382-012-1544-4.

Li, S., L. Goddard, and D. G. DeWitt, 2008: Predictive skill of AGCM seasonal climate forecasts subject to different SST prediction methodologies. *J. Climate,* **21,** 2169–2186, doi:10.1175/2007JCLI1660.1.

Lin, H., and Z. Wu, 2011: Contribution of the autumn Tibetan Plateau snow cover to seasonal prediction of North American winter temperature. *J. Climate,* **24,** 2801–2813, doi:10.1175/2010JCLI3889.1.

Lorenz, E. N., 1975: Climate predictability. *The Physical Basis of Climate and Climate Modelling,* GARP Publication Series, Vol. 16, World Meteorological Organization, 133–136.

Mason, S. J., L. Goddard, N. E. Graham, E. Yulaeva, L. Sun, and P. A. Arkin, 1999: The IRI seasonal climate prediction system and the 1997/98 El Niño event. *Bull. Amer. Meteor. Soc.,* **80,** 1853–1873, doi:10.1175/1520-0477(1999)080<1853:TISCPS>2.0.CO;2.

Matsueda, S., and Y. Takaya, 2012: Forecast skill of MJO with the JMAs One-month Ensemble Prediction System. CAS/JSC WGNE Research Activity Atmospheric Oceanic Model, Vol. 42, 6.11–6.12. AU7

Moron, V., A. W. Robertson, and M. N. Ward, 2006: Seasonal predictability and spatial coherence of rainfall characteristics in the tropical setting of Senegal. *Mon. Wea. Rev.,* **134,** 3248–3262, doi:10.1175/MWR3252.1.

——, ——, J.-H. Qian, and M. Ghil, 2015: Weather types across the Maritime Continent: From the diurnal cycle to interannual variations. *Frontiers Atmos. Sci.,* in press.

Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.,* **116,** 2417–2424, doi:10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2.

National Academy of Sciences, 2010: *Assessment of Intraseasonal to Interannual Climate Prediction and Predictability.* National Research Council, 192 pp.

Neena, J. M., J. Y. Lee, D. E. Waliser, B. Wang, and X. Jiang, 2014a: Predictability of the Madden–Julian oscillation in the Intraseasonal Variability Hindcast Experiment (ISVHE). *J. Climate,* **27,** 4531–4543, doi:10.1175/JCLI-D-13-00624.1.

——, X. Jiang, D. E. Waliser, J. Y. Lee, and B. Wang, 2014b: Eastern Pacific intraseasonal variability: A predictability perspective. *J. Climate,* **27,** 8869–8883, doi:10.1175/JCLI-D-14-00336.1.

Palmer, T. N., and D. L. T. Anderson, 1994: The prospects for seasonal forecasting—A review paper. *Quart. J. Roy. Meteor. Soc.,* **120,** 755–793, doi:10.1002/qj.49712051802.

Peatman, S. C., A. J. Matthews, and D. P. Stevens, 2014: Propagation of the Madden–Julian Oscillation through the Maritime Continent and scale interaction with the diurnal cycle of precipitation. *Quart. J. Roy. Meteor. Soc.,* **140,** 814–825, doi:10.1002/qj.2161.

Robertson, A. W., U. Lall, S. E. Zebiak, and L. Goddard, 2004: Improved combination of multiple atmospheric GCM ensembles for seasonal prediction. *Mon. Wea. Rev.,* **132,** 2732–2744, doi:10.1175/MWR2818.1.

Saha, S., and Coauthors, 2010: The NCEP climate forecast system reanalysis. *Bull. Amer. Meteor. Soc.,* **91,** 1015–1057, doi:10.1175/2010BAMS3001.1.

——, and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate,* **27,** 2185–2208, doi:10.1175/JCLI-D-12-00823.1.

Scaife, A. A., and J. R. Knight, 2008: Ensemble simulations of the cold European winter of 2005/6. *Quart. J. Roy. Meteor. Soc.,* **134,** 1647–1659, doi:10.1002/qj.312.

Stockdale, T. N., 1997: Coupled ocean–atmosphere forecasts in the presence of climate drift. *Mon. Wea. Rev.,* **125,** 809–818, doi:10.1175/1520-0493(1997)125<0809:COAFIT>2.0.CO;2.

Tippett, M. K., A. G. Barnston, and T. DelSole, 2010: Comments on ''Finite samples and uncertainty estimates for skill measures for seasonal prediction.'' *Mon. Wea. Rev.,* **138,** 1487–1493, doi:10.1175/2009MWR3214.1.

Vitart, F., 2004: Monthly forecasting at ECMWF. *Mon. Wea. Rev.,* **132,** 2761–2779, doi:10.1175/MWR2826.1. AU8

——, 2014: Evolution of ECMWF sub-seasonal forecast skill scores. *Quart. J. Roy. Meteor. Soc.,* **140,** 1889–1899, doi:10.1002/qj.2256.

——, and F. Molteni, 2010: Simulation of the Madden–Julian Oscillation and its teleconnections in the ECMWF forecast system. *Quart. J. Roy. Meteor. Soc.,* **136,** 842–855, doi:10.1002/qj.623.

——, and Coauthors, 2008: The new VarEPS-monthly forecasting system: A first step towards seamless prediction. *Quart. J. Roy. Meteor. Soc.,* **134,** 1789–1799, doi:10.1002/qj.322.

——, A. W. Robertson, and D. L. T. Anderson, 2012: Sub-seasonal to Seasonal Prediction Project: Bridging the gap between weather and climate. *WMO Bull.,* **61,** 23–28.

Waliser, D. E., 2011: Predictability and forecasting. *Intraseasonal Variability of the Atmosphere-Ocean Climate System,* 2nd ed. W. K. M. Lau and D. E. Waliser, Eds., Springer, 613 pp.

——, K. M. Lau, W. Stern, and C. Jones, 2003a: Potential predictability of the Madden–Julian oscillation. *Bull. Amer. Meteor. Soc.,* **84,** 33–50, doi:10.1175/BAMS-84-1-33.

——, W. Stern, S. Schubert, and K. M. Lau, 2003b: Dynamic predictability of intraseasonal variability associated with the Asian summer monsoon. *Quart. J. Roy. Meteor. Soc.,* **129,** 2897–2925, doi:10.1256/qj.02.51.

Wang, W., M.-P. Hung, S. J. Weaver, A. Kumar, and X. Fu, 2014: MJO prediction in the NCEP Climate Forecast System version 2. *Climate Dyn.,* **42,** 2509–2520, doi:10.1007/s00382-013-1806-9.

Wheeler, M. C., and H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.,* **132,** 1917–1932, doi:10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

World Meteorological Organization, 2013: Sub-seasonal to seasonal prediction: Research implementation plan, December 2013. WMO, 71 pp. [Available online at http://www.wmo.int/pages/prog/arep/wwrp/new/documents/S2S_Implem_plan_en.pdf.]

Xie, P., and P. A. Arkin, 1996: Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate,* **9,** 840–858, doi:10.1175/1520-0442(1996)009<0840:AOGMPU>2.0.CO;2.

——, and ——, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.,* **78,** 2539–2558, doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.

Yoo, J. H., A. W. Robertson, and I. S. Kang, 2010: Analysis of intraseasonal and interannual variability of the Asian summer monsoon using a hidden Markov model. *J. Climate,* **23,** 5498–5516, doi:10.1175/2010JCLI3473.1.

Zhang, C., 2013: Madden–Julian oscillation: Bridging weather and climate. *Bull. Amer. Meteor. Soc.,* **94,** 1849–1870, doi:10.1175/BAMS-D-12-00026.1.

Zhu, H., M. C. Wheeler, A. H. Sobel, and D. Hudson, 2014: Seamless precipitation prediction skill in the tropics and extratropics from a global model. *Mon. Wea. Rev.,* **142,** 1556–1569, doi:10.1175/MWR-D-13-00222.1.