AMERICAN METEOROLOGICAL SOCIETY

Journal of Climate

The DOI for this manuscript is doi: 10.1175/JCLI-D-15-0862.1

MJO propagation across the Maritime Continent in the ECMWF ensemble prediction system

Hye-Mi Kim[1*], Daehyun Kim[2], Frederic Vitart[3], Violeta E. Toma[4], Jong-Seong Kug[5] and Peter J. Webster[4]

[1]*School of Marine and Atmospheric Sciences, Stony Brook University, USA*

[2]*Department of Atmospheric Science, University of Washington, USA*

[3]*ECMWF, UK*

[4]*School of Earth and Atmospheric Science, Georgia Institute of Technology, USA*

[5]*School of Environmental Science and Engineering, POSTECH, South Korea*

Revision submitted to *J. Climate* (March 2016)

---

* *Corresponding author address*:

Hye-Mi Kim, School of Marine and Atmospheric Sciences,

Stony Brook University, Stony Brook, NY, 11794

E-Mail: hyemi.kim@stonybrook.edu

Abstract

The characteristics of the MJO propagation across the Maritime Continent are investigated using a 20-year reforecast dataset from the ECMWF ensemble prediction system. Analysis of the MJO events initialized over the Indian Ocean (phase-2) shows that the initial MJO amplitude and prediction skill relationship is not linear, particularly, when the predictions start in moderate (between strong and weak) MJO amplitude category. To examine the key factors that determine the prediction skill, reforecasts in the moderate category are grouped into high- and low-skill events, and the differences in their ocean-atmospheric conditions as well as the physical processes during reforecast period are examined. The initial distribution of OLR anomalies in high-skill events show a clear dipole pattern of convection with an enhanced convective anomalies over the Indian Ocean and strongly suppressed convective anomalies in the western Pacific. This dipole mode may support the MJO propagation across the Maritime Continent via the Rossby wave response and associated meridional moisture advection. Prominent ocean-atmosphere coupled processes are also simulated during the propagation of high-skill events. However, in low-skill events, the convective signal over the western Pacific is almost absent, less-organized, and the ocean-atmosphere coupled processes are not simulated correctly. We find that in both high- and low-skill events, the amplitude of the convective anomaly decreases significantly after about day-15 possibly due to the systematic mean model bias. A strong wet bias in the vicinity of the Maritime Continent, a cold SST bias in the equatorial Pacific, and associated circulation biases make the West Pacific area unfavorable for MJO propagation, thus limiting its prediction skill.

1. Introduction

There has been a growing interest of scientific, operational and application communities in providing forecasts in the sub-seasonal time range; a range that lies in between the conventional medium-range weather (lead times up to 2 weeks) and seasonal (3-12 months) forecasts. The subseasonal forecast is particularly important since many management decisions, such as in agriculture and food security, fall into this band. Skillful predictions of anomalous weather events, such as extreme precipitation events, in subseasonal timescale could provide policy makers, emergency management, and stakeholders advanced warnings to prepare for mitigatory measures.

Realizing the crucial role of subseasonal prediction on bridging the gap between weather forecasting and climate prediction, there is a current international effort focused on the subseasonal prediction problem. The World Climate Research Program/World Weather Research Program (WCRP/WWRP) Subseasonal to Seasonal (S2S) Prediction Project has recently been launched with the goal of advancing predictive capability at this time scale (Vitart et al. 2012). Accordingly, much attention has been paid from the NOAA/NWS and the White House to producing operational multi-model ensemble forecasts at weeks 3 and 4[1]. The North American Multi-model Ensemble (NMME, Kirtman et al. 2014) project is launching subseasonal inter-comparison assessment of contributing climate model forecasts (Robertson et al. 2015). Additionally, the newly formed National Research Council (NRC) Committee is charged with the goal of developing a 10-year U.S. research agenda to advance subseasonal prediction[2]. Such efforts provide unprecedented

[1] http://www.whitehouse.gov/the-press-office/2014/09/23/fact-sheet-president-obama-announces-new-actions-strengthen-global-resil

[2] http://dels.nas.edu/Upcoming-Event/Developing-Research-Agenda/DELS-BASCPR-13-05/

scientific opportunity to make significant strides in process-level understanding for subseasonal phenomena.

The Madden-Julian Oscillation (MJO; Madden and Julian 1971, 1972) is the dominant mode of subseasonal variability in the tropical atmosphere and ocean. The MJO is a special type of organized tropical convection that is distinct from other forms of tropical phenomena by its vast horizontal scale (wavenumber 1-5), subseasonal time scale (30-60 days), and eastward propagation over the Indo-Pacific warm pool. Enhanced or suppressed convection associated with the MJO affects the global weather and climate system (e.g., Zhang et al. 2013), thereby providing a source of potential predictability in subseasonal time scale. It is hoped that realizing uncapped predictability of the MJO is a key toward improving the subseasonal forecast.

The last two decades have witnessed substantial changes in the mode of MJO prediction. Until the early 2000s, dynamical models were inferior to statistical models in terms of MJO prediction, mainly due to their lack of MJO simulation capability. In most centers, MJO prediction was not a part of operational forecast suite of products until about 2010. Furthermore, the number of ensemble members was limited and forecasts, when they were made, were not issued frequently enough due to the expense of computing resource. In the most recent decade, MJO prediction has benefitted from the significant strides in the ability of models to represent the MJO (e.g., Zhang et al. 2013). Various operational forecast centers are now releasing MJO forecasts and are continuously upgrading their systems by improving model configurations (e.g., improved physics), new ensemble generation methods, better initialization schemes, and increased resolution. Current operational forecasting systems now show useful MJO prediction skill up to 3-4 weeks (Vitart and Molteni 2010, Vitart et al. 2010, Rashid et al. 2011, Vitart et al. 2012, Zhang and Van Den Dool 2012, Zhang et al. 2013, Kim et al. 2014b, Neena et al. 2014, Vitart 2014, Wang et al. 2014, Xiang

et al. 2015). However, this achievement is still below the theoretical estimate of predictability, which may be 6-7 weeks (e.g., Waliser et al. 2003). Thus, there would seem to be room for further enhancing MJO prediction by improving various aspects of the prediction system based on a better understanding the MJO phenomena. Understanding strengths and weaknesses of the current prediction systems would be the first step toward enhancing the MJO prediction skill.

There is consensus emerging from recent studies regarding the factors affecting the MJO predictability and prediction skill. The main factors, besides the ability of the model, are the geographic location of the MJO convection center (i.e. MJO phase) and the magnitude of the MJO signal (i.e. MJO amplitude) in the initial conditions. Earlier, ECMWF and NCEP forecast systems showed a high sensitivity of their MJO prediction skill to the initial MJO phase. They showed relatively low skill when the MJO convection is initially located over the Indian Ocean and propagates over the Maritime Continent, known as "Maritime Continent MJO prediction barrier" (Vitart et al. 2007, Seo et al. 2009, Vitart and Molteni 2010, Wang et al. 2014). However, recent studies showed that the predictability of the MJO is not sensitive to MJO phases, suggesting that the Maritime Continent prediction barrier is a modeling problem, rather than a predictability issue (Kim et al. 2014b, Neena et al. 2014).

Another crucial factor that affects the MJO prediction is its initial amplitude. It has been documented that the MJO prediction skill is higher when the MJO amplitude is initially relatively large (Rashid et al. 2011, Zhang and Van Den Dool 2012, Kim et al. 2014b, Neena et al. 2014, Vitart 2014, Wang et al. 2014, Xiang et al. 2015). In previous studies, initial amplitudes are classified into two or three different categories (e.g., weak, moderate, and strong) and the averaged MJO prediction skill of each category are compared. In all models, prediction skill shows a monotonic increase with the initial MJO amplitude, meaning that initially stronger MJOs possess

116 higher prediction skill. However, the linearity of the relationship between initial MJO amplitude
117 and MJO prediction skill has not yet been fully explored.

118 Although it has been shown that MJO prediction strongly depends on initial MJO phase and
119 amplitude, the physical mechanisms behind these constraints are not well understood. Based on
120 previous studies, there is one simple question that can be asked: Do MJO forecasts with similar
121 initial amplitude and phase always result in similar skill? For example, if two forecasts are
122 initialized with similar amplitude of MJO convective anomaly in the Indian Ocean, do they always
123 result in comparable skill? If not, what ocean-atmosphere conditions and physical processes drive
124 some MJO event to be more predictable than others? If a prediction system systematically fails or
125 succeeds in forecasting the MJO that meets a certain set of conditions, understanding the processes
126 behind those relationships may provide a key for improving the prediction system. In particular,
127 realizing a growing interest in the processes involved in the MJO propagation over the Maritime
128 Continent[3], this study aims to advance our understanding of the Maritime Continent prediction
129 barrier issue. In addition, we investigate the possible influence of the mean biases on MJO
130 propagation and prediction.

131 Section 2 introduces details of the data and methodology. In section 3, we examine the
132 relationship between the initial MJO amplitude and the prediction skill and identify the high- and
133 low-skill MJO events. The characteristics of high-skill versus low-skill events, in terms of initial
134 condition and propagation will be compared in section 4. In section 5, the relationship between the
135 mean biases and MJO prediction is investigated. Summary and discussion will be followed in
136 Section 6.

[3]http://www.bmkg.go.id/ymc/

2. Data and methodology

a. Data

We will use the reforecast data from the ECMWF monthly forecasting system (e.g., Vitart 2014) that has been used for real-time forecast in the current operations. In many studies, the ECMWF system has been shown as superior compared to other systems, particularly for MJO prediction (e.g., Kim et al. 2014b, Neena et al. 2014, Vitart 2014). MJO prediction skill in ECMWF system has gradually improved since 2002, benefited by the substantial improvement of physical parameterization, ocean-atmosphere coupling strategy, ensemble generation methods, data assimilation and etc. (Vitart 2014). The ECMWF Variable Resolution Ensemble Prediction System (hereafter, EC) is a fully coupled model system with a large set of reforecasts generated with the purpose of evaluating and calibrating the model forecast. In the WCRP/WWRP S2S Prediction Project, two publicly available reforecasts of the ECMWF system are version CY40R1 and CY41R1. Version CY40R1 has been used for operation before May-14 2015 and version CY41R1 is current operational version with the reforecasts being produced in real time. Therefore, the reforecast covering all seasons for CY41R1 will not be available until May 2016. In this study, we analyze the reforecast output of version CY40R1. Details of this version can be found at the ECMWF website[4].

The reforecast of CY40R1 consists of a 5-member ensemble generated at the same date as the real-time forecasts for the past 20 years starting at January 1994. The reforecast used in this study is projected to 32-day horizons every Thursday with the first 10 days at about 32 km (T639,

[4] https://software.ecmwf.int/wiki/display/IFS/CY40R1+Official+IFS+Documentation

158 0.28125°) horizontal resolution and changes to about 64 km (T319, 0.5625°) from day-11.
159 Orography field (represented by surface geopotential) prescribed in the EC reforecasts for each
160 T639 and T319 are obtained from ECMWF website[5]. There are 91 vertical levels extending to 0.01
161 hPa. Reforecasts have been initialized from ERA-Interim (Dee et al. 2011) and ORA-S4
162 (Balmaseda et al. 2013). The atmospheric model is coupled to the ocean from day-0. The reforecast
163 data covers 20-years from 1994 to 2013, resulting in total 5,200 sets of 32-day integrations (52
164 sets/year x 20 years x 5 ensembles). Daily mean fields of outgoing longwave radiation (OLR),
165 zonal wind of 200 hPa (U200) and 850 hPa (U850), sea surface temperature (SST), and
166 precipitation (Precip) are extracted from the reforecast. We use the zonal wind field from ERA-
167 Interim products, OLR from the Advanced Very High Resolution Radiometer (Liebmann and
168 Smith 1996), NOAA Optimum Interpolation sea surface temperature (Reynolds et al. 2007), and
169 precipitation from the Global Precipitation Climatology Project version 2.2 (Adler et al. 2003).
170 These datasets will be referred to as "observation" for convenience. The observational period
171 extends from 1981 to 2013. A daily climatology is calculated over the period from 1981 to 2013
172 for observations, and from 1994 to 2013 for EC reforecast.

173 *b. MJO index*

174 The Wheeler and Hendon (2004) Real-time Multivariate MJO index (RMM) is calculated
175 following Gottshalck et al. (2010). RMM indices, RMM1 and RMM2, are the timeseries of the
176 principle component of the 1st and 2nd Empirical Orthogonal Functions (EOF) of the OLR, U200,
177 and U850 averaged between 15°N and 15°S. RMM1 represents the anti-correlation in convection
178 between the Indian Ocean and western Pacific and RMM2 represents convective activity over the

[5] http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/cycle-41r1

Maritime Continent. The circulation patterns (U200 and U850) represent a dynamically coherent baroclinic structure associated with the convective anomaly. These two modes explain about 27 % of the total observed variance. Predicted RMMs are obtained by projecting the reforecast anomalies of zonal winds and OLR onto the observed eigenvectors. More details of the methodology can be found in Kim et al. (2014b). Figure 1 represents the observed MJO life cycle in eight different phases by compositing the OLR and U850 anomalies without discrimination for amplitude or season.

c. Measures of MJO prediction skill: Collective and segment prediction skill

The conventional measures of MJO prediction skill use a collection of MJO forecasts and calculate scalar metrics (e.g., bivariate anomaly correlation coefficient, ACC) as a function of forecast lead time (e.g., Lin et al. 2008). For example, when the RMM is employed, for each lead time, the RMM1 and RMM2 for all cases considered are used to calculate the scalar metrics that represent a general prediction skill of a forecast system. The same method can be applied to subsets of forecasts, for example, those starting from different MJO phases or amplitude. By comparing the prediction skill of each subset, one could examine the sensitivity of prediction skill to various factors (e.g., RMM prediction skill in various phase or amplitude as a function of lead time). However, this conventional method, which is a collective skill measure, cannot be applied to an individual forecast set that has only one predicted value at different lead times. Hence, using a collective skill measure, one cannot group individual forecasts by their prediction skill.

To mitigate the limitation of the conventional skill measurement, we develop a new metric for measuring predictive skill that can be applied to individual forecasts. We use this method to

identify forecast that exhibits relatively higher skill than others. To measure the skill of individual forecasts, ACC is calculated between the predicted and observed RMMs as follows:

$$\mathrm{ACC} = \frac{\sum_{t=1}^{t=N}[a_1(t)b_1(t) + a_2(t)b_2(t)]}{\sqrt{\sum_{t=1}^{t=N}[a_1^2(t) + a_2^2(t)]}\sqrt{\sum_{t=1}^{t=N}[b_1^2(t) + b_2^2(t)]}}$$

where $a_1(t)$ and $a_2(t)$ are the predicted RMM1 and RMM2 values at lead time t, and $b_1(t)$ and $b_2(t)$ are the corresponding observations. Therefore, each 32-day forecast segment has one value of ACC while the conventional skill measure, $ACC(\tau)$, has values as a function of lead days. We will refer to this 32-day segment ACC as 'segment prediction skill' and the conventional prediction skill measure as 'collective prediction skill'.

3. Identification of the high-skill and low-skill MJO events

In this section, we will identify the high- and low-skill MJO events in order to compare their characteristics. Before separating the MJO events, we will briefly review the overall MJO prediction skill in the current EC forecast system. The general RMM prediction skill of the current EC forecast is measured by the collective prediction skill first. The prediction skill of ensemble mean for all season is about 30-32 days as measured by the bivariate anomaly correlation coefficient exceeding 0.5. This is about 3-5 days improvement compared to its previous operational version (Kim et al. 2014b, Vitart 2014). Particularly, skill increases significantly when the predictions start at phase-2 (Fig.1) with approximately 0.2 improvement of ACC at days 25-30 mean compared to the previous version used in Kim et al. (2014b). Experiments with the coupling from day-0 (rather than from day-10 as in the previous version in Kim et al. 2014b) produced a slight improvement in MJO skill scores (a gain of about 1 day of predictive skill, see

Fig. 13 in Janssen et al. 2013). While there is still a gap between the predictability and prediction skill, the gap is getting narrower by continuously improving the system.

We classify individual reforecasts into a few groups depending on the initial MJO amplitude and the segment prediction skill of them. We focus on a specific initial MJO phase – phase 2, in which convection center is located over the Indian Ocean (Fig. 1) and where the MJO is expected to propagate across the Maritime Continent during the forecast period. The selection of a particular initial MJO phase is motivated by the fact that there has been a growing interest of understanding the critical processes involved in the MJO propagation, particularly from the Indian Ocean to the western Pacific, and that the Maritime Continent prediction barrier still exists in many operational forecasting systems.

a. MJO initial amplitude and prediction skill relationship

As a first step toward identifying the high- and low-skill events, the relationship between initial MJO amplitude and the segment prediction skill is examined. A total of 670 reforecasts (134 events x 5 ensembles) with initial phase-2 are selected. To categorize the skill by its initial amplitude, we define the amplitude as the square root of $RMM1^2$ plus $RMM2^2$ based on the observation. The events of the initial amplitude greater than 1.5 are selected as strong MJO events (total 42 cases out of 134 cases), events with initial amplitude between 0.7 and 1.5 as moderate MJO events, and the rest as weak/non-MJO events. Two amplitude thresholds, 0.7 and 1.5, are chosen arbitrarily to categorize about 50 % of reforecasts as moderate MJO events (total 68 events). It needs to be emphasized that events are selected based on the observed initial amplitude and phase. Figure 2 shows the segment prediction skill as a function of initial amplitude. It is clear that the segment prediction skills for initially strong MJO events are skewed to relatively high

values with having only one case lower than 0.5, while the skills in the moderate and weak/non groups are widely distributed. If we take the average of skill for each amplitude category, skill increases monotonically with amplitude from 0.62 in the weak/non group, to 0.70 in the moderate group, and to 0.82 in the strong group, consistent with previous studies. However, when the skills of reforecasts are considered individually, the initial amplitude-prediction skill relationship is not linear. Particularly, the skills of the moderate category are broadly scattered and not clearly skewed in any direction. Identifying factors in ocean-atmospheric condition that lead a reforecast to having a relatively high skill or low skill, would help improving the prediction system. For further analysis, reforecasts in the moderate category (total 68 events) will be grouped into high- and low-skill cases, and the difference in their ocean-atmospheric condition as well as the physical processes during forecast period will be examined in the subsequent sections. To clarify that high- and low-skill events do not depend strongly on season, we further distinguished these phase-2 MJO events by season (not shown). Most of the strong MJOs occurred in boreal winter (Nov-Apr) and weak/non MJOs in summer (May-Oct). However, the frequency of moderate MJO events is less sensitive to seasons. Because the seasonality of MJO prediction is an important topic, this will be investigated in future work.

b. High-skill versus low-skill events

To examine the characteristics of high-skill versus low-skill events starting at phase 2 with moderate amplitude (hereafter, simply "high-" and "low-skill" events), we classify reforecasts into high- and low-skill events based on their segment prediction skill. If the segment skill is equal or above 0.74, the event is considered as high-skill event and the rest as low-skill events. The threshold value 0.74 is determined arbitrarily to split the high- and low-skill events equally (each has 34 events). Using each 34 (170) events of 32-day ensemble mean (individual ensemble

members) forecast sets, the collective prediction skill is calculated as a function of lead days (Fig. 3). The skill of ensemble mean is clearly higher than that of the individual ensembles in high-skill MJO, and there is a large inter-ensemble spread. In low-skill events, although the prediction skill stays much the same as high-skill events up to day-05, the skill drops rapidly as lead time increases. The skill in the ensemble mean reaches correlation of 0.5 at day-16 in low-skill events, while for high-skill events the correlation stays above 0.7 during the entire period. This suggests that, even starting at similar amplitude and phase, some events are better predicted with an additional two weeks of skill.

To condense the characteristics of the amplitude and propagation change by forecast lead days, we compare RMMs composites in a phase-space diagram (Fig. 4). Since the initial phases and amplitudes are computed based on the observation and not prediction, slight differences exist between observed and predicted RMMs even on day-01 due to the model error. Predicted RMMs of individual ensembles on day-01 are plotted together with the ensemble mean. While all observed MJO events start on phase-2 (not shown), some predicted MJOs start in slightly different phases. Because of the limited number of samples, we do not investigate the details of each case but focus on the mean statistics of high- and low-skill events. In the composite of high-skill events (Fig. 4), while the predicted amplitude of the ensemble mean is weaker than the observed during the entire forecast period, the predicted events propagate over the Indian Ocean to western Pacific with comparable phase speed to observation leading to a relatively high prediction skill. In the low-skill events, both observed and predicted RMMs propagate faster than the high-skill events and lose their amplitude rapidly.

4. Favorable initial conditions and key physical processes for MJO propagation across the Maritime Continent

a. Initial condition and propagation characteristics

We have shown that the initially moderate MJO events can result in either high or low skill forecasts. Are there clear differences in initial conditions that translate into changes in prediction skill? To answer this question, we examine the initial ocean-atmospheric conditions and propagation characteristics. Figure 5 compares the composite maps of the predicted OLR and SST anomalies at day-01 for both high- and low-skill events. The predictions shown in Figure 5 are similar to the observations (not shown). The initial distribution of OLR anomalies in high-skill events (Fig. 5a) show a clear dipole pattern of convection with enhanced convective anomalies over the Indian Ocean and strongly suppressed convective anomalies in the western Pacific. The convective anomalies over the western Pacific are associated with positive SST anomalies (maximum > 0.3 K). In contrast, for the low-skill events (Fig. 5b), even though the convective anomalies have similar amplitude and phase in the Indian Ocean, the signal over the western Pacific is almost absent, resulting in a monopole convective anomaly structure. Given that the initial amplitude and phase of the RMM-defined MJOs are similar in the two categories (high- and low-skill), the drastic contrast in the convective pattern is surprising. This issue will be addressed later.

To examine the propagation characteristics, the longitude-time composites of OLR and U850 are employed (Fig. 6). The observed counterparts for the high-skill events have relatively well-organized eastward propagation with anomalous low-level easterly wind preceding the convection anomaly (Fig. 6a). An opposite phase of MJO convection anomaly develops over the western

Indian Ocean at around day-15 associated with the westerly wind anomaly to the east of the suppressed convection. These characteristics are captured to some extent in the prediction but with weaker amplitude than observed (Fig. 6b). The predicted low-level wind anomaly shows eastward propagation to the eastern Pacific until day-32. For the low-skill events (Fig. 6c, d), both the observations and forecasts exhibit a propagation of convective anomaly which is much faster than that in the high-skill events, a phase speed resembling that of the convectively coupled equatorial Kelvin wave (e.g., Wheeler and Kiladis 1999). In short, the high- and low-skill events show clear differences in their initial conditions, especially the distribution of convective anomaly, and in their representation of eastward propagation. In the high- (low-) skill events, the initial condition of convection is characterized by a dipole (monopole) structure, and the phase speed of the eastward propagation is about 5.0 m/s (9.0 m/s), calculated by predicted U850 change from day-01 to day-10. Interestingly, in both high- and low-skill events, predicted OLR signal becomes weak after around day-15.

Recent studies have emphasized the importance of the dipole structure in the convective anomaly for the propagation of the MJO (Kim et al. 2014a; Adames and Wallace 2015). Kim et al. (2014a) showed that a suppressed convective anomaly over the west Pacific is a necessary condition for an enhance convection over the Indian Ocean to propagate across the Maritime Continent. By analyzing the moist static energy of the MJO, Kim et al. (2014a) suggested that the suppressed anomaly over the west Pacific plays a critical role in inducing a Rossby wave response to its west. The circulation anomalies and resulting anomalous meridional moisture advection associated with the Rossby wave response provide favorable conditions for the propagation of the MJO across the Maritime Continent. The moisture budget analysis of Adames and Wallace (2015)

showed similar results. That is, an initial state with a dipole convective anomaly pattern favors the MJO propagation across the Maritime Continent.

Collectively, these results suggest that the horizontal pattern of convective anomaly in the initial conditions could be a useful indicator of how skillful a prediction will be for the cases in which there is initial intermediate MJO amplitude. To test this argument, the relationship between the observed western Pacific (120°E-180°E, 15°S-15°N) OLR anomaly (OLRwp) at day-01 and the segment prediction skills for all initially strong and moderate events starting at phase 2 are sought (Fig. 7). Although the prediction skills are not clearly separated as a function of the OLRwp, most of the relatively high skill events have initially drier western Pacific to the east of the Maritime Continent. The frequency of a dry western Pacific occurring is higher in high-skill than in low-skill events. It seems that the drier western Pacific is a necessary condition but not sufficient condition for high-skill MJO prediction events.

b. Ocean-atmosphere interaction

Previous studies revealed the importance of the ocean feedback to the MJO propagation (e.g., review in DeMott et al. 2015). The warm SST east of the MJO convection in the Pacific Ocean supports the MJO propagation by affecting SST-modulated heat fluxes, thus enhancing the boundary layer moisture convergence which maintains the convective anomalies and fuel the eastward propagation (e.g., Lindzen and Nigam 1987, Maloney and Sobel 2004, Back and Bretherton 2009, Hsu and Li 2012, Hirata et al. 2013). Figure 5a suggests that the ocean feedback process is associated with the high-skill events. The positive western Pacific SST anomaly at day-01 indicates that the suppressed convection allows a steady warming of the SST due to the weak surface wind and enhanced downward solar radiation. To examine the temporal changes of the

convection and underlying SST, the area averaged OLR and SST anomaly over the western Pacific (110°E-160°E, 15°S-15°N) are compared as a function of lead days (Fig. 8). Since the SST does not change significantly during the short forecast period, we have removed the 32-day SST average of each 32-day segment in both observation and prediction, respectively. During observed high-skill MJO (Fig. 8a), positive SST and positive OLR anomaly (suppressed convection) over the western Pacific is present at day-01. During the initial suppressed phase in the western Pacific, and prior to maximum convection, low-level easterly wind anomalies become significant over the warm western Pacific (Fig. 6a) where the climatological westerly wind dominates. Decreased evaporative cooling and enhanced downward solar radiation raise SST anomalies to the east of the convection (e.g., Hirata et al. 2013). After about day-05, the suppressed convective anomaly decreases its amplitude and changes to a convectively active phase (Fig. 8a). During the convection phase, reduced solar heating and evaporative surface cooling and mixing via strong low-level westerly induce SST change to negative anomalies. For both observation and predictions, this convection-SST phase relationship is well captured during the eastward propagation for high-skill MJO events. However, in low-skill events, no significant SST-convection phase relationship can be found in the western Pacific in both observation and prediction (Fig. 8b).

c. Drawbacks of using the RMM index

Before we move on the investigation of the mean bias and prediction skill relationship, we will discuss some drawbacks in using the RMM index as a measure of the MJO. In the RMM space-phase diagram (Fig. 4), we showed that in the high-skill events, the RMM-defined MJO signal shows propagation even after day-15 (Fig. 4) and prediction maintains high-skill throughout (Fig. 3). However, even in high-skill events, the convective signal becomes almost absent after lead day-15 (Fig. 6). If the forecast convective signal is marginal after day-15, then what causes

such a high prediction skill until day-32? It has become apparent that the fractional contribution of winds to the variance of RMMs is relatively larger than the contribution of the convective fields. This is a major weakness of the RMM index (Straub 2013, Ventrice et al. 2013, Liu et al. 2015, Wolding and Maloney 2015). To address this issue in our results, Figure 9 shows the observed and predicted anomalies (averaged over 10°S-10°N) of three variables used in defining the RMM index. On day-01, while overall predicted amplitude is slightly weaker than the observed, the convection and wind anomalies are well captured in the predictions (Fig. 9a, b), resulting in high skill at the beginning of the forecast (Fig. 3) in both high- and low-skill events. However, for day20-25 average in high-skill events (Fig. 9c), while the predicted convective anomaly shows negligible strength over the Indo-Pacific region compared to the observation, the wind anomalies still have comparable magnitudes (Fig. 9c). Therefore, although the reforecasts do not predict the convective anomaly correctly, the prediction skill could result in high-skill after day-15 since the fractional contribution of winds to the RMM is larger than the convection signal (Straub 2013).

Another issue is that both high- and low-skill events are selected as initially moderate "phase-2" category, although the initial convective anomalies in the western Pacific have different structure (dipole versus monopole mode, Fig. 4). In both observation and prediction, the western Pacific dry anomaly is apparent in the high-skill events (Fig. 9a), but not in the low-skill events (Fig. 9b). However, the anomalous low and upper level wind patterns at day-01 are similar between high- and low-skill events. These similar wind patterns and amplitude at day-01 force these events to be categorized as moderate phase-2 events. Therefore, when the RMM index is used for the prediction of MJO convective signal, it may potentially misinform the skill of the MJO. To compensate the weakness of RMM index in representing the convection, several alternative methods have been introduced (e.g., Kiladis et al. 2014, Liu et al. 2015). However, the use of

400 proper MJO index depends on how one defines the MJO phenomena. In this study, we will
401 continue to use RMMs as an index, bearing in mind the drawbacks, since most of MJO prediction
402 studies and operational predictions still rely heavily on the RMM index.

403

404 **5. Systematic mean biases and the possible influence on MJO prediction**

405 To explain the possible barrier for MJO prediction, we examine the relationship between the
406 MJO prediction skill and the systematic mean bias in the EC reforecasts. Figure 6 showed that the
407 convective signal in the reforecasts almost disappears after about day-15 when the MJO convection
408 center reaches the western Pacific. We assumed that the mean bias could be an important factor
409 that deteriorates the MJO propagation further east. Figure 10 shows the climatological annual
410 mean bias of precipitation, SST and zonal winds (U850 and U200) averaged over the entire lead
411 days (day01-32) for the entire reforecast period (20 years) for all events starting at phase 2,
412 regardless of the initial amplitude. Excessive precipitation bias is found in the vicinity of the
413 Maritime Continent in agreement to other studies of various state-of-the-art coupled climate
414 models (e.g., DeMott et al. 2014). A warm SST bias is found in the equatorial Indian Ocean
415 extending to the western Pacific with maximum over the Maritime Continent oceans (Fig. 10b).
416 The warm SST biases could increase low-level moist static energy thus inducing favorable
417 conditions for convection and the related excessive precipitation. Over the equatorial central to
418 eastern Pacific, cold SST biases are dominant with a significantly large maximum value of about
419 -1.0 K. The strong SST gradient between the western and eastern Pacific induces a strong pressure
420 gradient and associated low-level easterly (Fig. 10a) and upper-level westerly bias (Fig. 10b)
421 similar to those found in climate models (e.g., Lin 2007).

To examine the temporal change of these biases, the initial bias (day-01) is compared (Fig. 10c, d). Wet biases in the vicinity of the Maritime Continent and the associated low-level easterly wind in the western Pacific appear at day-01 (Fig. 10c). At day-01, strong cold SST bias, which is likely an Ocean initialization error, and associated upper level wind bias are already occurring in the equatorial East Pacific (Fig. 10d). The excessive precipitation bias over the Maritime Continent regions and the east-west SST gradient bias may induce a pressure gradient bias that, in turn, could enhance the wind biases after day-01. Such an enhanced easterly wind bias over the tropical Pacific can induce steep eastward shoaling of the thermocline thus exaggerates positive ocean-atmospheric Bjerknes feedback (1966). The enhanced Walker circulation is likely to enhance the ocean upwelling thus contributing to the cold SST bias in the regions of easterly biases. Moreover, the strong wind bias may cause an excessive latent heat flux and contribute to the cold SST bias as well. While the cold SST bias is confined to the East Pacific at the beginning of the forecast, it continuously shifts to the west as lead time increases (not shown). In addition to this basin-wide feedback, the interaction between local SST and shortwave solar flux could also play a role on enhancing the mean bias. Over the cold tongue region, cold SST bias could increase the static stability of the boundary layer, thus increasing the low cloud amount that, in turn, decreases the surface downward shortwave flux thus cooling the SST (Klein and Hartmann 1993, Peter and Bretherton 2005). This local positive feedback in the cold tongue region and the basin-wide Bjerknes feedback could jointly enhance the Walker circulation until the bias fields reach their equilibrium state.

In short, the initial wet precipitation bias in the Maritime Continent and the cold SST bias in equatorial Pacific lead to an enhanced Walker circulation and amplify the positive ocean-atmosphere feedback to some extent. However, it is unclear from where the bias originates and

what causes their development. Since the atmospheric and oceanic fields are closely coupled, bias in one field could impact other fields. More work is needed to understand the origin of the biases in the reforecasts.

Another factor that amplifies the bias could be the model resolution. A wet bias is found over the Maritime Continent vicinity at day-01 (Fig. 10c). Figure 11 compares the change of precipitation and U850 biases. Values are averaged over 5°S and 5°N as a function of forecast lead days. The easterly wind bias increases as forecast lead day increases, but the wet bias maintains a comparable magnitude until day-10 (Fig. 11). At day-11, the time when the model changes its horizontal resolution from about 32 km to 64 km, the wet bias almost doubles in magnitude in the center of the Maritime Continent (Fig. 11). The low and upper level wind biases rapidly amplify in consistency with the growing wet bias.

Schiemann et al. (2014) showed that increasing horizontal resolution improves the precipitation simulation particularly in Maritime Continent region due to better representation of complex orography. Figure 12 compares the sub-grid orography (represented in surface geopotential) used as boundary conditions up to day-10 (Fig. 12a) and after day-11 (Fig. 12b). It represents the height of the terrain with respect to the model defined Earth. Up to day-10, the orography is at finer resolution (of 32 km) and represents complex and heterogeneous regions. After day-11, since the resolution decreases (of 64 km), there are relatively fewer grid points that resolve the complexity of land regions. The wet bias and low-level convergence east of the convective Maritime Continent region could be enhanced immediately when the resolution changes.

How do those biases hinder the MJO eastward propagation and thus prediction? The wet bias over the Maritime Continent could induce a local Walker circulation that, in turn, would promote a descending motion over the west Pacific. Also, the easterly anomaly over the west Pacific would advect a drier air to the west Pacific where the mean moisture gradient is negative. Both horizontal and vertical circulation biases would make the west Pacific area unfavorable for the development of anomalous convection, thereby weakening MJO propagation to the west Pacific. Moreover, the strong wind biases over the equatorial Pacific may inhibit the MJO propagation after day-15 when the MJO enters the western Pacific and the MJO associated wind has the opposite sign to the wind bias. The relationship between the mean bias and MJO prediction deserves further investigation.

6. Summary and discussion

We examined the characteristics of MJO propagation across the Maritime Continent in the ECMWF ensemble prediction system using a 20-year reforecast dataset produced by the recent operational version (CY40R1). We focused on MJO events initialized over the Indian Ocean (phase-2) and investigate the key dynamical factors attributed to the MJO propagation and prediction. Analysis of the initial MJO amplitude and prediction skill relationship shows that the initial amplitude-prediction skill relationship is not linear. Particularly, when the prediction starts with moderate amplitude, predictions could have either high- or low- skills depending on the initial ocean-atmospheric condition. In the high-skill events, the initial condition is characterized by a clear dipole pattern of convection with an enhanced convective anomaly over the Indian Ocean and strongly suppressed convective anomalies in the western Pacific. This suppressed convective anomaly is a necessary condition for enhanced convection over the Indian Ocean to propagate across the Maritime Continent. Therefore, high-skill events result in a relatively well-organized

eastward propagation with anomalous low-level easterly wind preceding the convection anomaly. According to the results found in this study, the horizontal pattern of convective anomaly in the initial conditions could be a useful indicator of predicting how skillful the prediction will be for the cases in which the initial MJO amplitude is intermediate. During the propagation, the convection-SST phase relationship is well captured in the high-skill events. In low-skill events, the west Pacific suppressed convection is almost absent, the propagation is poorly organized, and a convection-SST relationship is not obvious.

However, even in high-skill events, the amplitude of the predicted convection anomaly decreases significantly after about day-15. We found that the systematic mean biases partially influence extended MJO prediction. The initial wet bias in the Maritime Continent region, the cold bias in the Pacific SST, and the associated wind biases grow as lead day increases. These strong biases over the Maritime Continent to Pacific Ocean make the west Pacific area unfavorable for the MJO propagation, thus limit the prediction. In spite of continuous improvement, the biases over the Maritime Continent and the neighboring oceans are known as a challenge for the state-of-the-art coupled GCM (e.g., DeMott et al. 2014).

The Maritime Continent is the region of the largest precipitation and tropical heating, and modulates the global weather and climate through teleconnection. To improve the MJO prediction, in addition to model development for better MJO simulation, the correct representation of the mean climate particularly over the Maritime Continent is crucial. It has to be accompanied by solving the MJO-mean state trade-off issue which is the fact that improving the MJO simulation tends to degrade other aspects such as the mean state (Kim et al. 2011, Boyle et al. 2015). The need for such model improvements is also necessary to improve predictions for both the tropics and

511 extratropics and for subseasonal to longer timescales. An ongoing international effort[3] will be of
512 great significance to foster our understanding of the Maritime Continent prediction barrier.

513 While our current study provides some new findings in the area of MJO prediction, there are
514 non-negligible weaknesses, which are mainly due to the lack of sufficient reforecast sample size
515 and variables. First, the proposed mechanism of the suppressed convection anomaly in the western
516 Pacific playing a role on the eastward propagation is based on existing studies (Hirata et al. 2013,
517 Kim et al. 2014a) but the limitations in reforecast data availability presented a challenge in
518 exploring detailed mechanisms. Second, we analyzed the MJO prediction irrespective to the
519 season. Generally, the MJO is expected to be better predicted during the boreal winter compared
520 to the summer due to the more pronounced MJO signal (e.g., Wang et al. 2014). The boreal summer
521 intraseasonal oscillation (BSISO) has distinct characteristics that distinguish it from the MJO (e.g.,
522 Lawrence and Webster 2002, Kikuchi et al. 2012). Therefore, to accurately evaluate the current
523 state of the MJO (or BSISO) prediction, an appropriate method should be applied to specific
524 season. Third, we used mixed cases of primary and secondary MJO in our analysis, although
525 primary and secondary MJOs have distinct characteristics as well as different predictability (e.g.,
526 Matthew 2008). Neena et al. (2014) showed that the predictability of the secondary MJO is about
527 5 days higher than that of the primary MJO in ECMWF system. Fourth, we have shown (Fig. 2)
528 that initially strong MJOs generally have high skill, and weak/non MJOs have lower skill, so we
529 focused on the moderate category. However, looking closely at Fig. 2, one can find that some
530 ensemble members have extremely high skill (about 0.98) even in weak MJO category and
531 extremely low skill even in strong MJO category. Besides the initially moderate MJO category,
532 investigating the characteristics of high- and low-skill MJOs in each strong and weak category will
533 be of great significance to foster better MJO prediction. Fifth, we only focused on the MJO events

that are initialized at phase 2. This was done intentionally as our aim is to understand the characteristics of MJOs propagating over the Maritime Continent. We plan to extend our study to MJOs starting at various phases in further study. Lastly, the MJO activity exhibits a significant interannual change through the modulation of the tropical mean state, such as the El Nino Southern Oscillation (ENSO) or Indian Ocean Dipole (IOD) mode (review in DeMott et al. 2015), but we have not distinguished MJO events by these climate modes in this paper. The limitations listed above cannot be overcome without having sufficient number of variables, frequently initialized reforecasts, and large ensembles. Moreover, the results of the present study may depend on the forecast system. Whether the conclusion of this study is hold for other forecasting systems is questionable. Subsequent studies may be necessary with multi-models that have large sets of ensemble members to help identify common weakness and strengths in the current MJO forecasts. For future work, we plan to examine our hypothesis in multi-models from WCRP/WWRP S2S Prediction Project[1] (Vitart et al. 2012) and the NMME reforecasts (Kirtman et al. 2014).

Acknowledgements

The constructive and valuable comments from anonymous reviewers are greatly appreciated. We would like to thank ECMWF for providing the data and Dr. Hai-Ru Chang for helping data downloading. The National Science Foundation under grant NSF-AGS 0965610 and the KMA R&D Program under Grant APCC 2013-3141 provided funding support.

References

Adames, A .F. and J. M. Wallace, 2015: Three-dimensional structure and evolution of the moisture field in the MJO , J. Atmos. Sci., 72, 3733–3754.

Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). J. Hydrometeor., 4, 1147–1167.

Balmaseda, M. A., Mogensen, K. and Weaver, A. T. (2013), Evaluation of the ECMWF ocean reanalysis system ORAS4. Q.J.R. Meteorol. Soc., 139: 1132–1161. doi: 10.1002/qj.2063

Back, L. E., and C. S. Bretherton, 2009: On the relationship between SST gradients, boundary layer winds, and convergence over the tropical oceans, J. Clim., 22, 4182–4196.

Bjerknes, J. 1966: A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature. Tellus, 4, 820–829.

Boyle, J. S., S. A. Klein, D. D. Lucas, H.-Y. Ma, J. Tannahill, and S. Xie, 2015: The parametric sensitivity of CAM5's MJO. J. Geophys. Res. Atmos., 120, 1424–1444.

Dee, D. P., and coauthors 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc., 137, 553-597.

DeMott, C. A., C. Stan, M. Branson, and D. A. Randall, 2014: Intraseasonal variability in coupled GCMs: The role of ocean feedbacks and model physics. J. Climate, 27, 4970-4995, doi: 10.1175/JCLI-D-13-00760.1.

DeMott, C. A., N. P. Klingaman, and S. J. Woolnough, 2015: Atmosphere-ocean coupled processes in the Madden-Julian oscillation, Rev. Geophys.,53, doi:10.1002/2014RG000478.

575 Fu, X., B. Wang, J. Y. Lee, W. Q. Wang, and L. Gao, 2011: Sensitivity of dynamical intraseasonal
576 prediction skills to different initial conditions, Mon. Wea. Rev., 139, 2572–2592.

577 Fu, X., J. Y. Lee, P. C. Hsu, H. Taniguchi, B. Wang, W. Wang, S. Weaver, 2013: Multi-model
578 MJO forecasting during DYNAMO/CINDY period, Clim. Dyn., 41, 1067–1081.

579 Gottschalck, J., and coauthors, 2010: A framework for assessing operational Madden-Julian
580 Oscillation forecasts: A CLIVAR MJO Working Group project, Bull. Amer. Meteor. Soc., 91,
581 1247–1258.

582 Hirata, F. E., P. J. Webster, and V. E. Toma, 2013: Distinct manifestations of austral summer
583 tropical intraseasonal oscillations, Geophys. Res. Lett., 40 (12), 3337-3341.

584 Hsu, P.-C., and T. Li, 2012: Role of the boundary layer moisture asymmetry in causing the
585 eastward propagation of the Madden-Julian oscillation, J. Clim., 25, 4914–4931.

586 Janssen, P., Ø. Breivik, K. Mogensen, F. Vitart, M. Balmaseda, J.-R. Bidlot, S. Keeley, M.
587 eutbecher, L. Magnusson and F. Molteni, 2013: Air-Sea Interaction and Surface Waves,
588 ECMWF Technical Memorandum 712.

589 Kikuchi, K., B. Wang, and Y. Kajikawa, 2012: Bimodal representation of the tropical intraseasonal
590 oscillation. Climate Dyn., 38, 1989–2000, doi:10.1007/s00382-011-1159-1

591 Kiladis, G. N., J. Dias, K. H. Straub, M. C. Wheeler, S. N. Tulich, K. Kikuchi, K. M. Weickmann,
592 and M. J. Ventrice, 2014: A comparison of OLR and circulation-based indices for tracking
593 the MJO. Mon. Wea. Rev., 142, 1697-1715.

594 Kim, D., and coauthors, 2009: Application of MJO Simulation Diagnostics to Climate Models, J.
595 Climate, 22, 6413–6436.

596 Kim, D., A. H. Sobel, E. D. Maloney, D. M. W. Frierson, I. S. Kang, 2011: A systematic
597 relationship between intraseasonal variability and mean state bias in AGCM simulations, J.
598 Climate, 24 (21), 5506-5520.

599 Kim, D., J. S. Kug, and A. H. Sobel, 2014a: Propagating versus Nonpropagating Madden–Julian
600 Oscillation Events. J. Climate, 27, 111–125.

601 Kim, H. M., P. J. Webster, V. E. Toma, and D. Kim, 2014b: Predictability and Prediction Skill of
602 the MJO in Two Operational Forecasting Systems. J. Climate, 27, 5364–5378.

603 Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1
604 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction.
605 Bull. Amer. Meteor. Soc., 95, 585–601.

606 Lawrence, D. and P. J. Webster, 2002: The boreal summer intraseasonal oscillation and the South
607 Asian monsoon. J. Atmos. Sci., 59, 1593-1606.

608 Lau, W. K. M., and D. E. Waliser, Eds., 2011: Intraseasonal Variability of the Atmosphere–ocean
609 Climate System, 2nd Ed. Springer, Heidelberg, Germany, pp 613.

610 Liebmann B. and C. A. Smith, 1996: Description of a complete (interpolated) outgoing longwave
611 radiation dataset. Bull Am Meteorol Soc 77:1275–1277

612 Lin, H., G. Brunet, and J. Derome, 2008: Forecast skill of the Madden-Julian Oscillation in two
613 Canadian atmospheric models, Mon. Wea. Rev., 136, 4130–4149.

614 Lin, J. L., 2007: The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean-atmosphere
615 feedback analysis. J. Climate, 20, 4497-4525.

Liu, P. Q. Zhang, C. Zhang, Y. Zhu, M. Khairoutdinov, H. M. Kim, C. Schumacher, M. Zhang, 2015: A Revised Real-Time Multivariate MJO Index, Mon. Wea. Rev. (accepted)

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific, J. Atmos. Sci., 28, 702–708.

Madden, R. A., and P. R. Julian, 1972: Description of global-scale circulation cells in the tropics with a 40–50 day period, J. Atmos. Sci., 29, 1109–1123.

Maloney, E. D., and A. H. Sobel, 2004: Surface fluxes and ocean coupling in the tropical intraseasonal oscillation, J. Clim., 17, 4368–4386.

Matthews, A. J., 2008: Primary and successive events in the Madden–Julian Oscillation. Quart. J. Roy. Meteor. Soc., 134, 439–453.

Neena, J. M., J. Y. Lee, D. Waliser, B. Wang, X. Jiang, 2014: 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.

Rashid, H. A., H. H. Hendon, M. C. Wheeler, and O. Alves, 2011: Prediction of the Madden-Julian oscillation with the POAMA dynamical prediction system, Clim. Dyn., 36, 649–661.

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

Robertson, A. W., A. Kumar, M. Peña, and F. Vitart, 2015: Improving and Promoting Subseasonal to Seasonal Prediction. Bull. Amer. Meteor. Soc., 96, ES49–ES53.

Saha, S, and coauthors, 2014: The NCEP Climate Forecast System Version 2. J. Climate, 27, 2185–2208.

637 Seo, K. H., W. Wang, J. Gottschalck, Q. Zhang, J-K. Schemm, W. Higgins, and A. Kumar, 2009:
638 Evaluation of MJO forecast skill from several statistical and dynamical forecast models, J.
639 Climate, 22, 2372–2388.

640 Schiemann, R., M.-E. Demory, M. S. Mizielinski, M. J. Roberts, L. C. Shaffrey, J. Strachan, P. L.
641 Vidale, 2014: The sensitivity of the tropical circulation and Maritime Continent precipitation
642 to climate model resolution. Climate Dynamics 42, 2455–2468.

643 Straub, K. H., 2013: MJO initiation in the real-time multivariate MJO index. J. Climate, 26, 1130–
644 1151, doi:10.1175/JCLI-D-12-00074.1

645 Ventrice M. J., M. C. Wheeler, H. H. Hendon, C. J. Schreck III, C. D. Thorncroft, G. N. Kiladis,
646 2013: A Modified Multivariate Madden-Julian Oscillation Index Using Velocity Potential,
647 Mon. Wea. Rev., 141, 4197-4120.

648 Vitart, F., S. Woolnough, M. A. Balmaseda, and A. Tompkins, 2007: Monthly forecast of the
649 Madden-Julian Oscillation using a coupled GCM, Mon. Wea. Rev., 135, 2700–2715.

650 Vitart F., A. Leroy, M. C. Wheeler, 2010: A comparison of dynamical and statistical predictions
651 of weekly tropical cyclone activity in the southern hemisphere. Mon. Wea. Rev., 138, 3671–
652 3682.

653 Vitart, F. and F. Molteni, 2010: Simulation of the Madden– Julian Oscillation and its
654 teleconnections in the ECMWF forecast system, Q. J. R. Meteorol. Soc., 136, 842–855.

655 Vitart, F., A. Robertson and D. Anderson, 2012: Subseasonal to Seasonal Prediction Project:
656 bridging the gap between weather and climate. WMO Bulletin, 61(2), 23-28.

Vitart, F., 2014: Evolution of ECMWF sub-seasonal forecast skill, Q. J. R. Meteorol. Soc., DOI: 10.1002/qj.2256

Waliser, D. E., K. M. Lau, W. Stern, and C. Jones, 2003: Potential Predictability of the Madden-Julian Oscillation, Bull. Amer. Meteor. Soc., 84, 33–50.

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

Weaver, S. J., W. Wang, M. Chen and A. Kumar, 2011: Representation of the MJO variability in the NCEP Climate Forecast System, J. Climate, 24, 4676–4694.

Wheeler, M., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber-frequency domain, J. Atmos. Sci., 56, 374 – 399.

Wheeler, M. C. and H. 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.

Wolding, B. O., and E. D. Maloney, 2015: Objective diagnostics and the Madden–Julian oscillation. Part I: Methodology. J. Climate, 28, 4127–4140, doi:10.1175/JCLI-D-14-00688.1

Xiang, B., M. Zhao, X. Jiang, S.-J. Lin, X. Fu, T. Li, G. Vecchi, 2015: The 3-4 week MJO Prediction skill in a GFDL Coupled Model, J. Climate. doi: http://dx.doi.org/10.1175/JCLI-D-15-0102.1.

Zhang, C., 2005: Madden-Julian Oscillation, Rev. of Geophysics, 43, RG2003, doi:10.1029/2004RG000158, 2005.

Zhang Q. and H. van den Dool, 2012: Relative merit of model improvement versus availability of retrospective forecasts: the case of Climate Forecast System MJO prediction. Weather Forecast 27, 1045–1051.

Zhang, C., J. Gottschalck, E. D. Maloney, M. Moncrieff, F. Vitart, D. E. Waliser, B. Wang, and M. C. Wheeler, 2013: Cracking the MJO nut, Geophys. Res. Lett., 40, 1223–1230.

682 **Figure list**

683 Figure 1. MJO life cycle composite maps for OLR (W/m^2, shading) and 850hPa zonal wind (m/s,
684 contour interval is 0.5) anomalies calculated for each of the eight MJO phase for all season
685 from 1981 to 2013.

686 Figure 2. Relationship between the initial MJO amplitude (x-axis) and ensemble mean segment
687 prediction skill (y-axis) for the forecasts starting at phase 2. Two gray vertical lines indicate
688 initial amplitude of 0.7 and 1.5.

689 Figure 3. Collective prediction skill (bivariate anomaly correlation coefficient) as a function of
690 forecast lead days for high-skill (red) and low-skill (blue) events initialized at phase 2 in
691 ensemble mean (thick lines) and individual ensembles (thin lines).

692 Figure 4. RMM composite phase-space diagram for high-skill (red) and low-skill (blue) events
693 starting at phase 2 in observation (solid lines with open circle) and prediction (dotted lines
694 with closed circle). Circles represent every 5 day from day-01 (square). Red and blue
695 (orange and light blue) dots represent the ensemble mean (individual ensembles).

696 Figure 5. The composite maps of predicted SST (K, shading) and OLR anomaly (W/m^2, contour)
697 at day-01 for (a) high-skill and (b) low-skill MJO events. The SST anomalies are multiplied
698 by 100. Negative OLR anomalies are in dashed contour with 4 W/m^2 interval and omitting
699 the values from -5 to 5 W/m^2.

700 Figure 6. Longitude-time composites of the OLR (W/m^2, shading) and U850 (m/s, contour)
701 anomaly averaged over 10°S-10°N for high-skill (top) and low-skill (bottom) MJO events
702 in observation (left) and prediction (right). Contour interval is 0.4 m/s and zero lines are
703 omitted.

Figure 7. Segment prediction skill (y-axis) and observed OLR anomaly over the west Pacific (WP) (x-axis, 120°E-180°E, 15°S-15°N) at day-01 for all initially strong and moderate events.

Figure 8. Area averaged [110°E-160°E,15°S-15°N] values for SST (K, red) and OLR (W/m^2, black) anomalies as a function of forecast lead days for (a) high-skill and (b) low-skill MJO. Observed values are in solid line and predicted values are marked with open circle.

Figure 9. Latitudinal average [10°S-10°N] of OLR (W/m^2, gray and black), U850 (m/s, blue), and U200 (m/s, red) anomalies for observation (bold lines and gray shading) and prediction (dashed line) at day-01 for (a) high-skill and (b) low-skill, and for (c) day20-25 average for high-skill events.

Figure 10. Climatological annual mean bias for phase-2 events in (a) precipitation (mm/day, shading) and U850 (contour interval 1 m/s) and (b) SST (K, shading, multiplied by 10) and U200 (contour interval 2 m/s). (c, d) same as (a, b) except for initial bias. Dashed lines indicate negative values.

Figure 11. Mean bias for precipitation (shading) and U850 (contour interval 1 m/s) averaged over 5°S-5°N as a function of forecast lead days.

Figure 12. Orography represented by surface geopotential (m^2/s^2, divided by 1,000) in the land surface boundary condition prescribed (a) before day-10 and (b) after day-11.

Figures

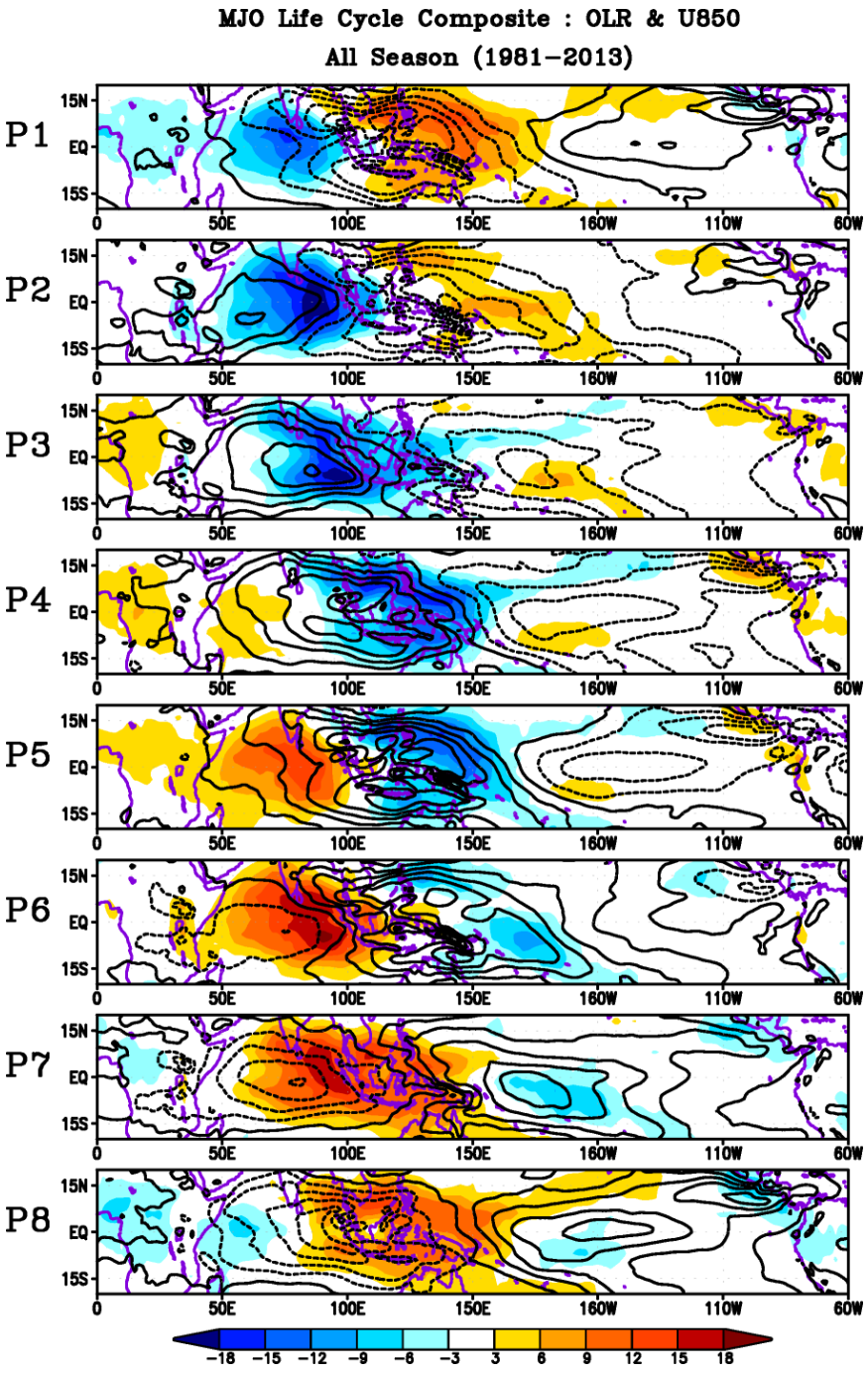
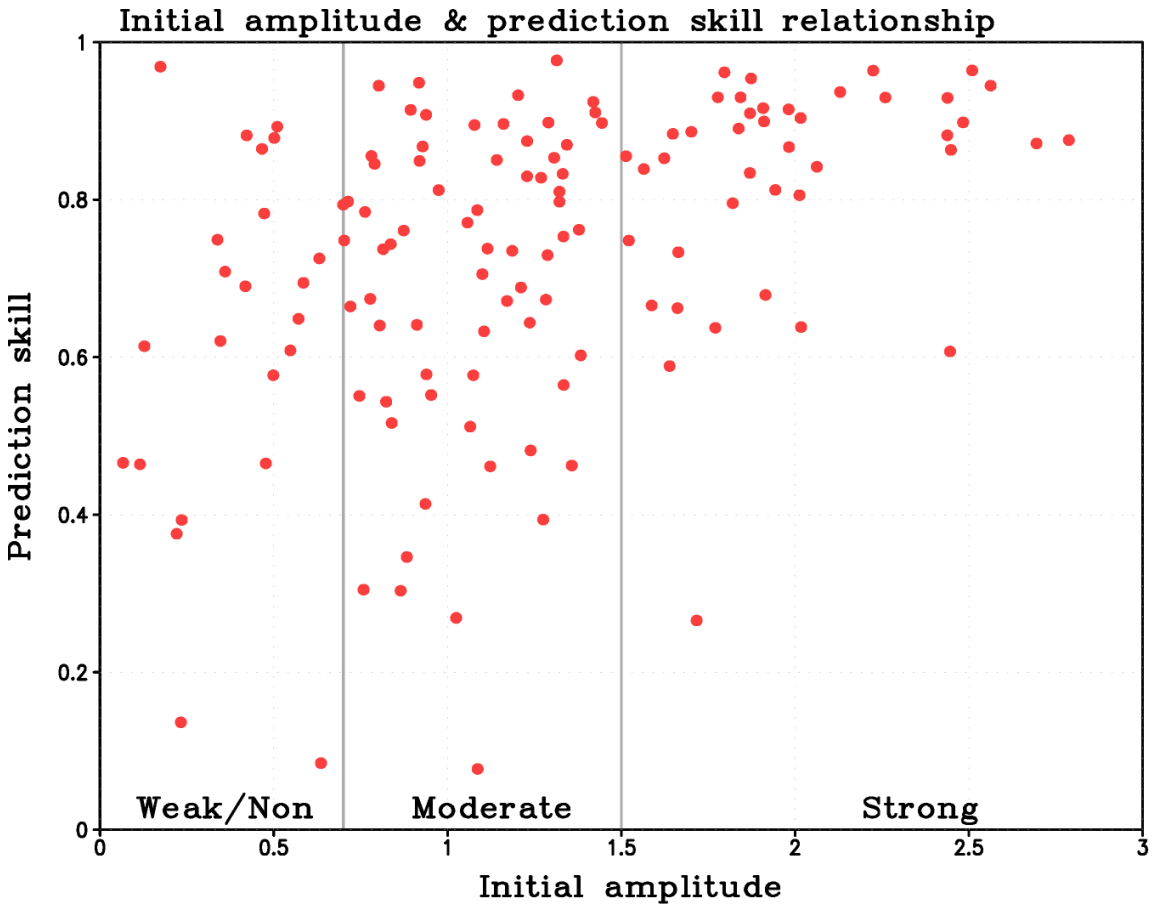


Figure 1. MJO life cycle composite maps for OLR (W/m^2, shading) and 850hPa zonal wind (m/s, contour interval is 0.5) anomalies calculated for each of the eight MJO phase for all season from 1981 to 2013.



728

729 Figure 2. Relationship between the initial MJO amplitude (x-axis) and ensemble mean segment
730 prediction skill (y-axis) for the forecasts starting at phase 2. Two gray vertical lines indicate initial
731 amplitude of 0.7 and 1.5.

732

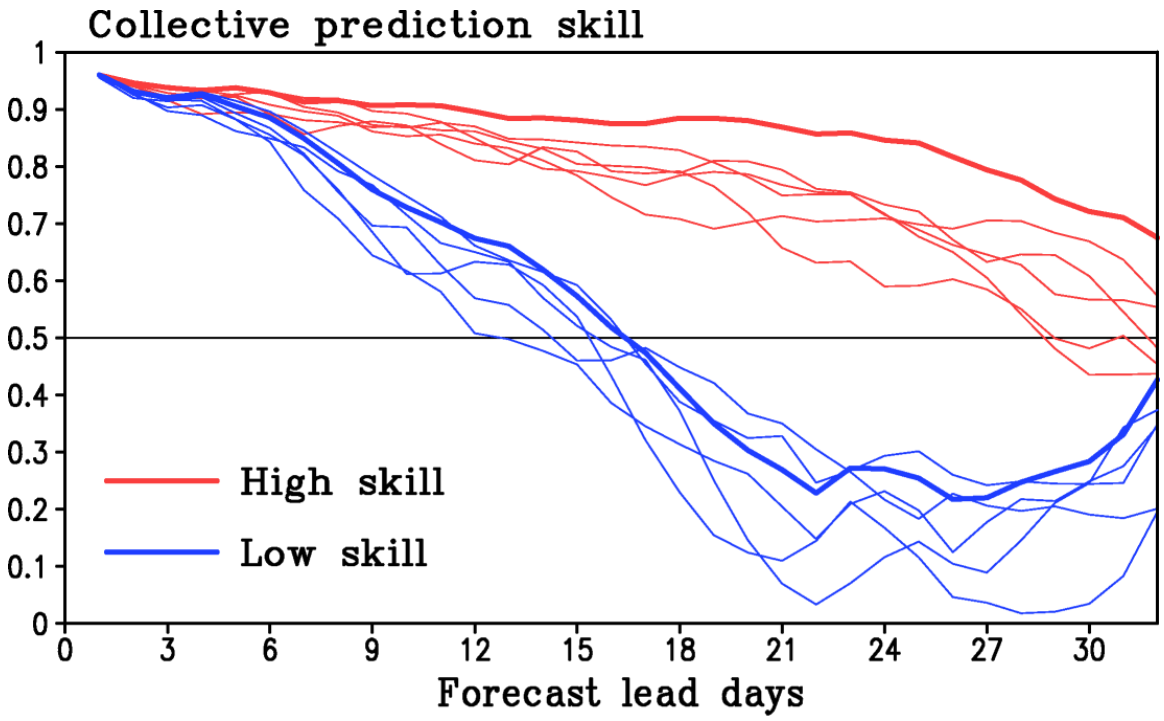


Figure 3. Collective prediction skill (bivariate anomaly correlation coefficient) as a function of forecast lead days for high-skill (red) and low-skill (blue) events initialized at phase 2 in ensemble mean (thick lines) and individual ensembles (thin lines).

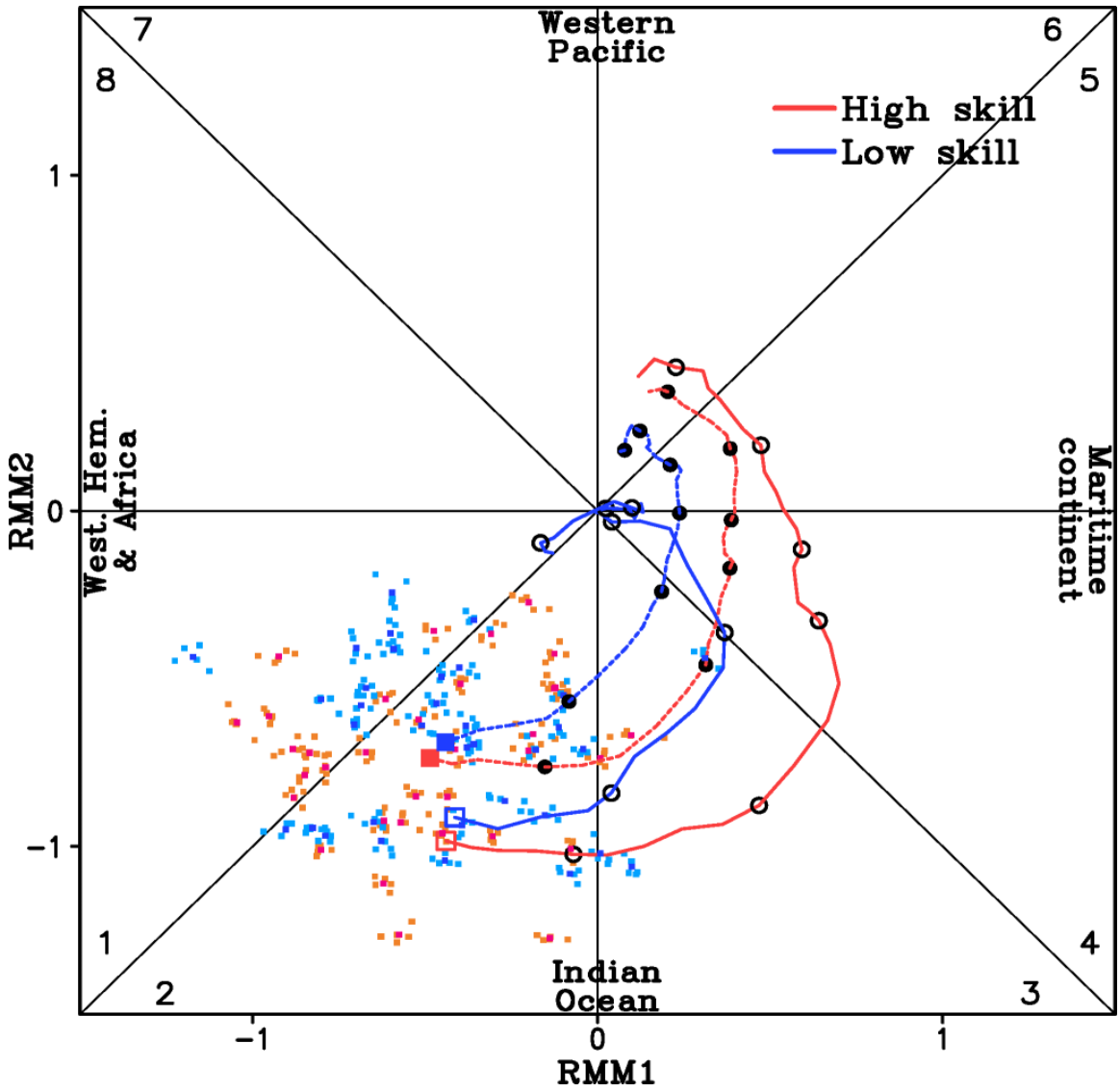
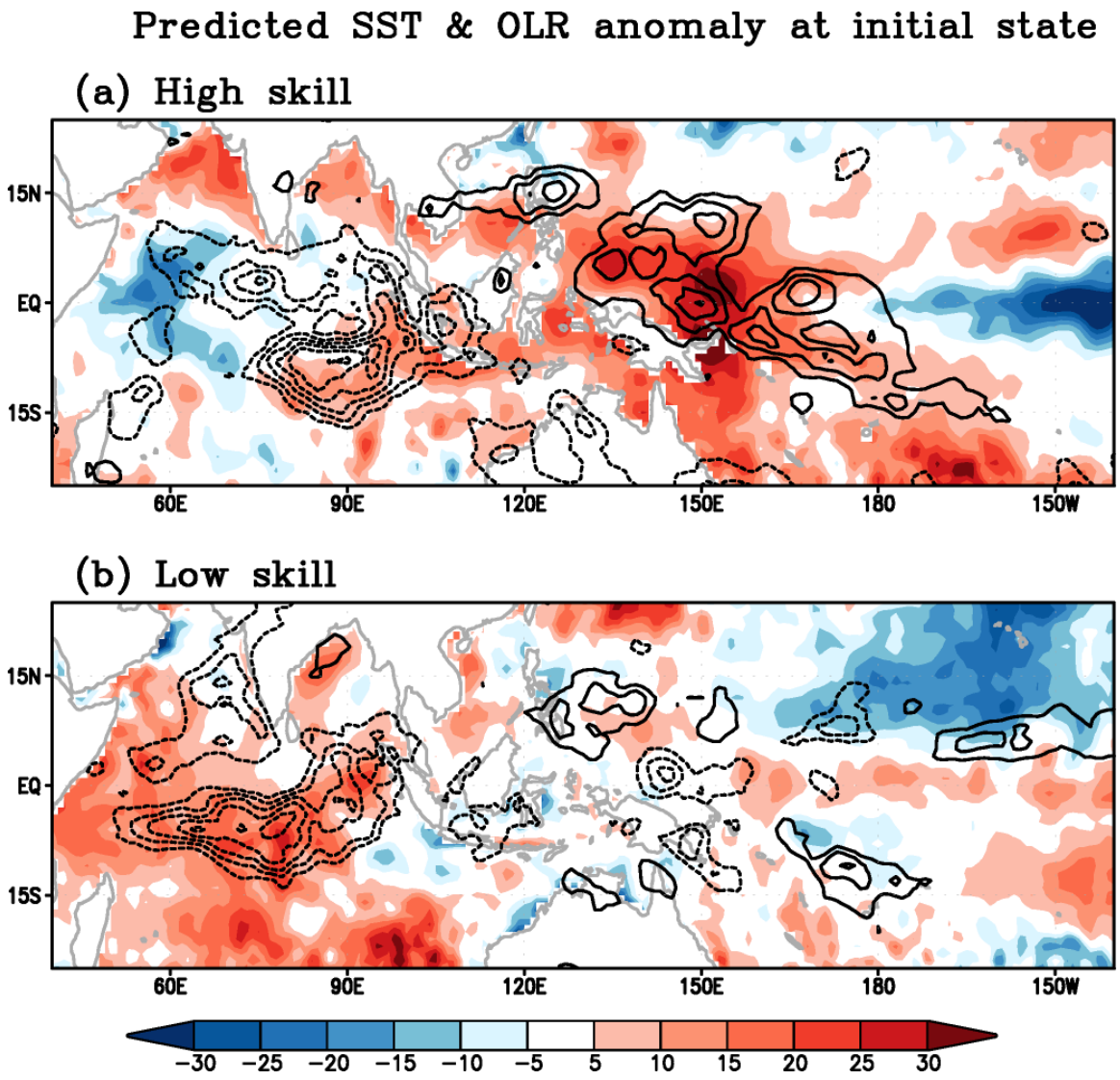
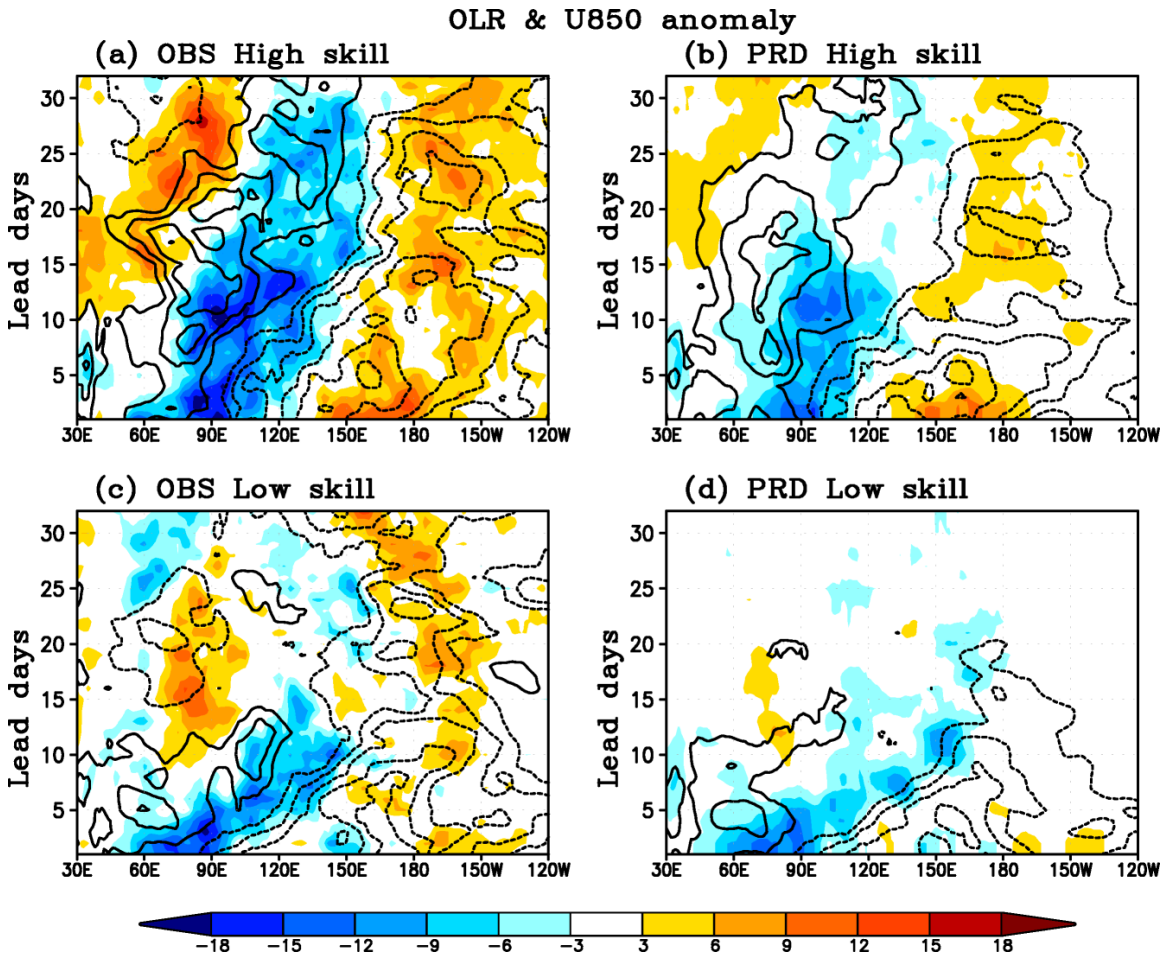


Figure 4. RMM composite phase-space diagram for high-skill (red) and low-skill (blue) events starting at phase 2 in observation (solid lines with open circle) and prediction (dotted lines with closed circle). Circles represent every 5 day from day-01 (square). Red and blue (orange and light blue) dots represent the ensemble mean (individual ensembles).



744

745 Figure 5. The composite maps of predicted SST (K, shading) and OLR anomaly (W/m^2, contour)
746 at day-01 for (a) high-skill and (b) low-skill MJO events. The SST anomalies are multiplied by
747 100. Negative OLR anomalies are in dashed contour with 4 W/m^2 interval and omitting the values
748 from -5 to 5 W/m^2.



749

750 Figure 6. Longitude-time composites of the OLR (W/m^2, shading) and U850 (m/s, contour)
751 anomaly averaged over 10°S-10°N for high-skill (top) and low-skill (bottom) MJO events in
752 observation (left) and prediction (right). Contour interval is 0.4 m/s and zero lines are omitted.

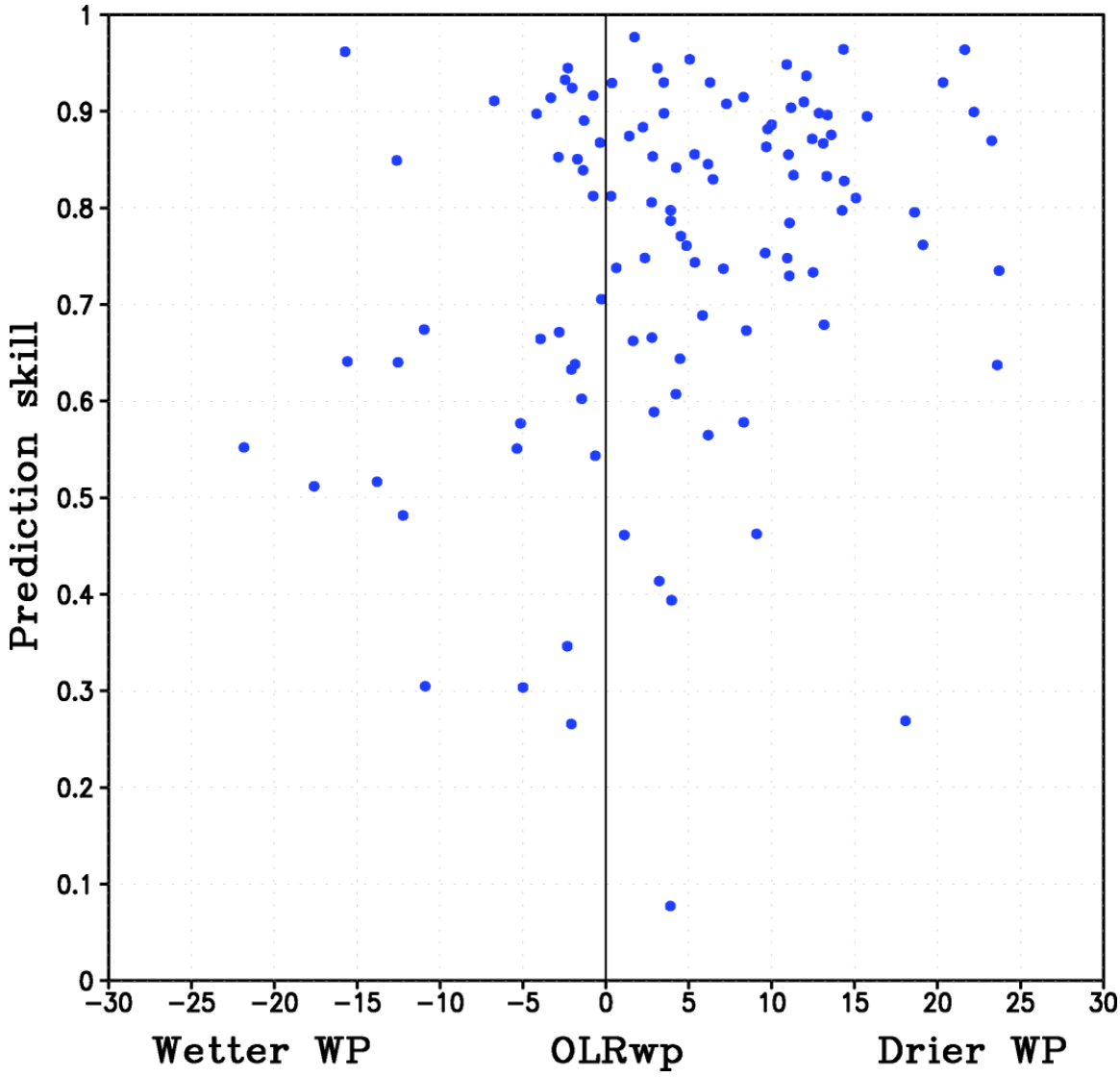


Figure 7. Segment prediction skill (y-axis) and observed OLR anomaly over the west Pacific (WP) (x-axis, 120°E-180°E, 15°S-15°N) at day-01 for all initially strong and moderate events.

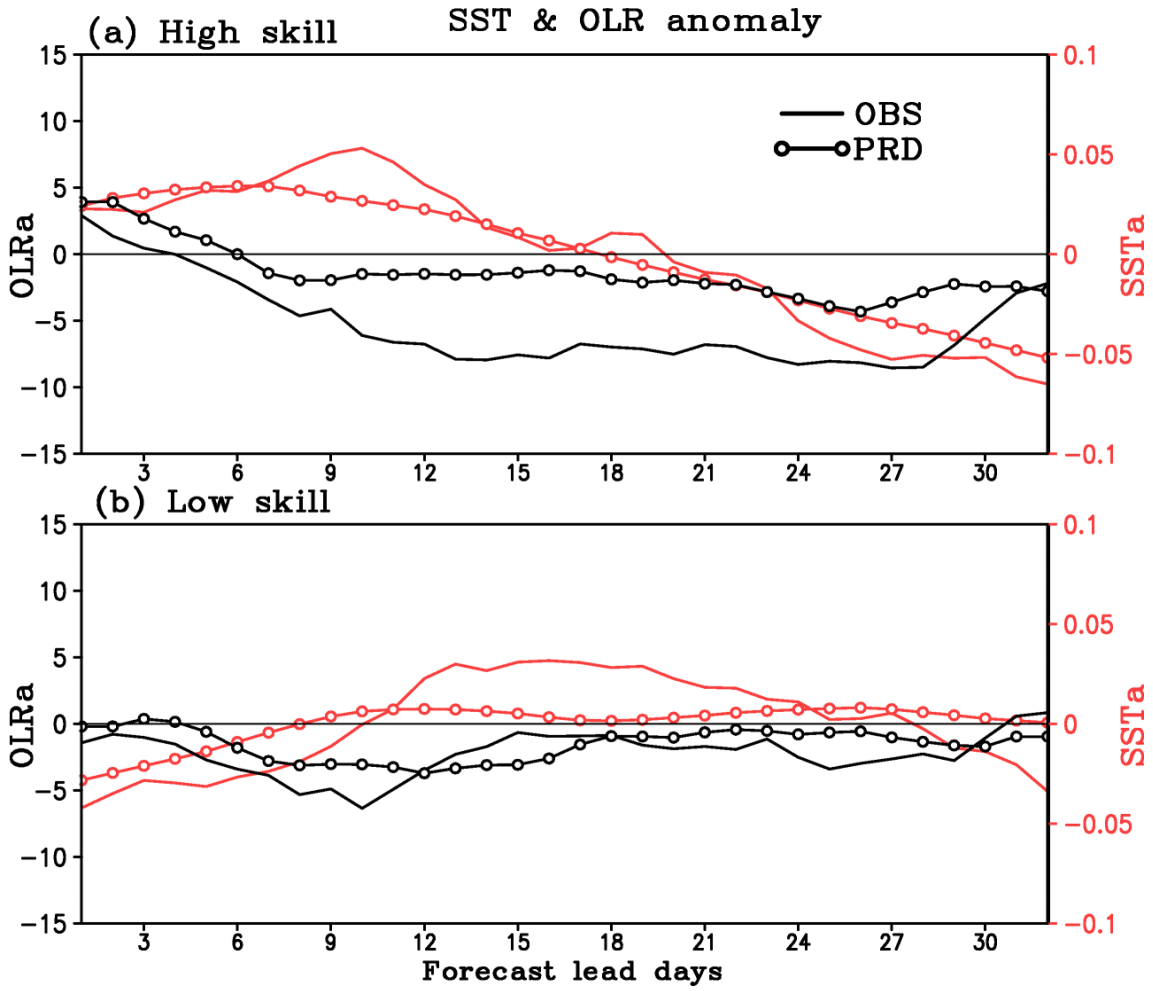


Figure 8. Area averaged [110°E-160°E,15°S-15°N] values for SST (K, red) and OLR (W/m^2, black) anomalies as a function of forecast lead days for (a) high-skill and (b) low-skill MJO. Observed values are in solid line and predicted values are marked with open circle.

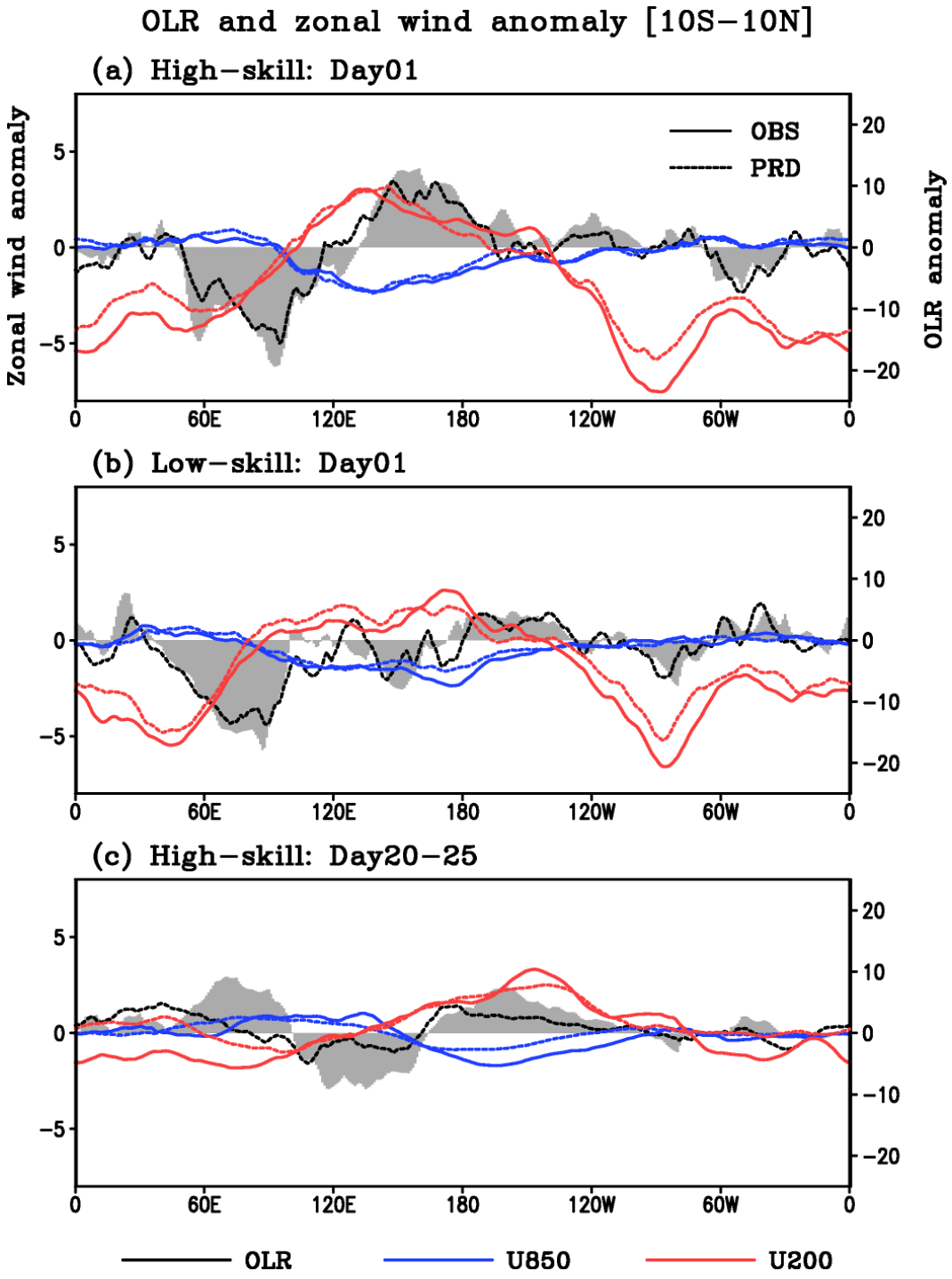
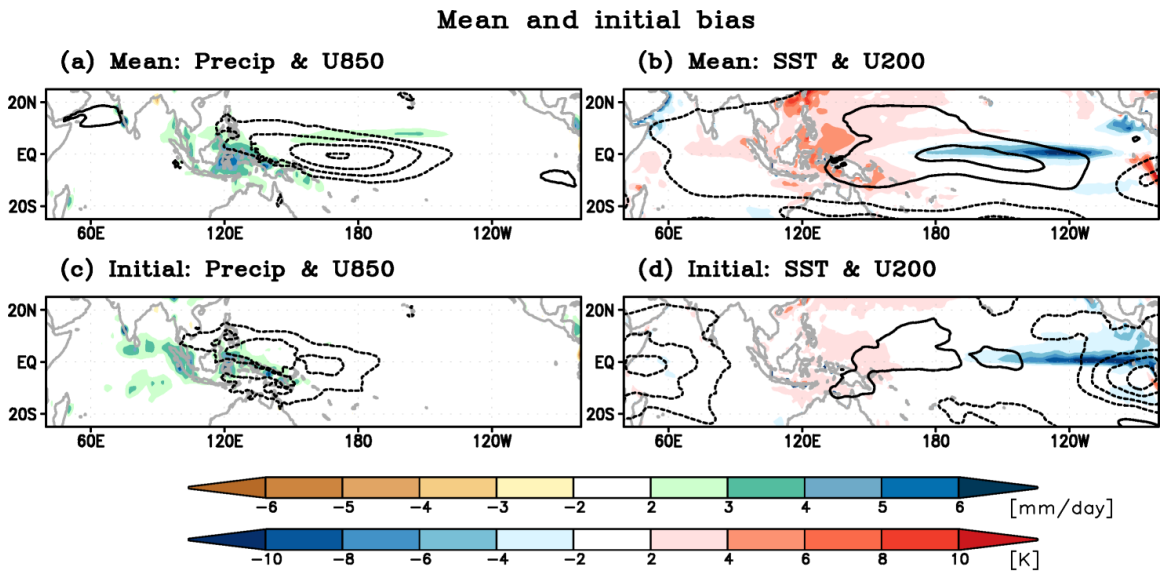


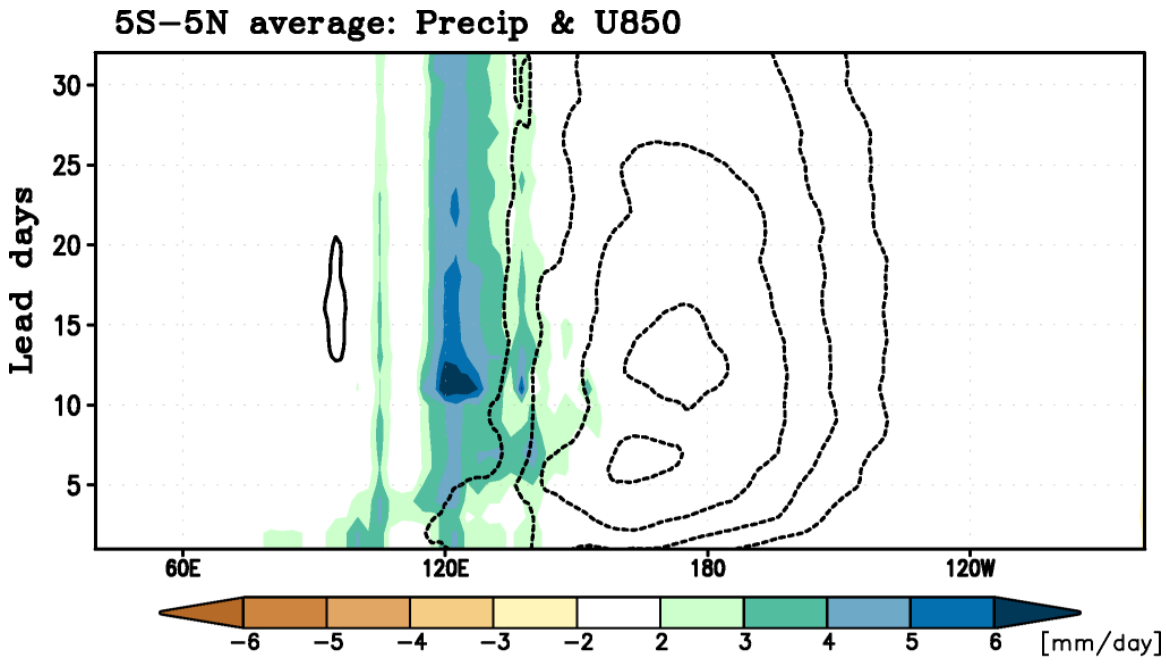
Figure 9. Latitudinal average [10°S-10°N] of OLR (W/m^2, gray and black), U850 (m/s, blue), and U200 (m/s, red) anomalies for observation (bold lines and gray shading) and prediction (dashed line) at day-01 for (a) high-skill and (b) low-skill, and for (c) day20-25 average for high-skill events.

765



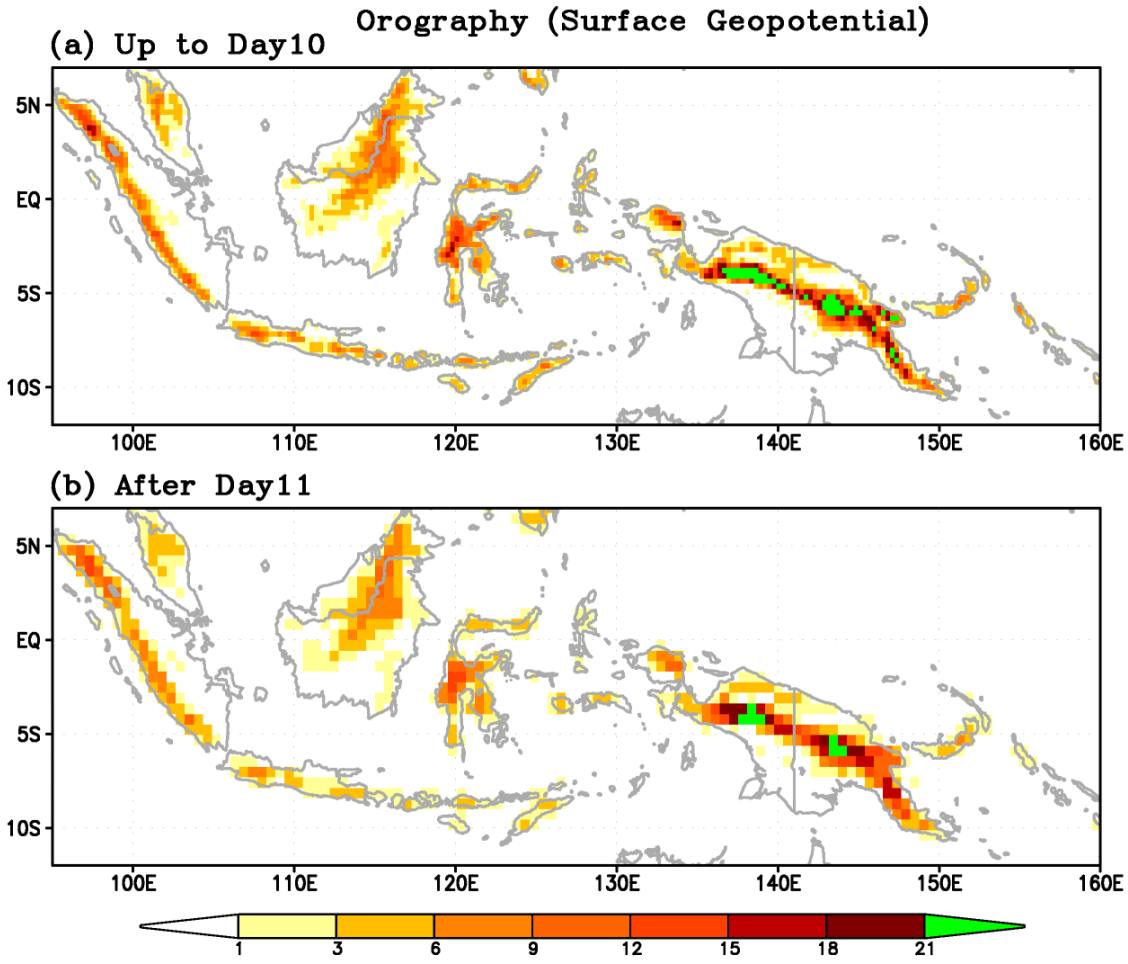
766

767 Figure 10. Climatological annual mean bias for phase-2 events in (a) precipitation (mm/day,
768 shading) and U850 (contour interval 1 m/s) and (b) SST (K, shading, multiplied by 10) and U200
769 (contour interval 2 m/s). (c, d) same as (a, b) except for initial bias. Dashed lines indicate negative
770 values.



771

772 Figure 11. Mean bias for precipitation (shading) and U850 (contour interval 1 m/s) averaged over
773 5°S-5°N as a function of forecast lead days.



774

775 Figure 12. Orography represented by surface geopotential (m^2/s^2, divided by 1,000) in the land
776 surface boundary condition prescribed (a) before day-10 and (b) after day-11.

777