MJO Propagation Processes and Mean Biases in the SubX and S2S Reforecasts

Hyemi Kim¹, Matthew A. Janiga², and Kathy Pegion³

¹School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY
²Naval Research Laboratory, Marine Meteorology Division, Monterey, CA
³Department of Atmospheric, Oceanic and Earth Sciences, George Mason University, Fairfax, VA

J. Geophysical Research - Atmospheres
(Special section: Bridging Weather and Climate: Subseasonal-to-Seasonal (S2S) Prediction)

Key Points:
• SubX and S2S reforecasts show MJO prediction skill out to 4.5 weeks based on the RMM index.
• SubX and S2S models fail to predict the MJO convection, associated circulations, and moisture advection processes beyond 10 days.
• SubX and S2S models have mean biases across the Indo-Pacific: a drier low troposphere, excess of surface precipitation, frequent occurrence of light precipitation.

* Corresponding author:
Hyemi Kim (hyemi.kim@stonybrook.edu)
119 Endeavor Hall, School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, 11794-5000, USA
Abstract

The Madden-Julian Oscillation (MJO) is the leading source of global subseasonal predictability; however, many dynamical forecasting systems struggle to predict MJO propagation through the Maritime Continent (MC). Better understanding the biases in simulated physical processes associated with MJO propagation is the key to improve MJO prediction. In this study, MJO prediction skill, propagation processes, and mean state biases are evaluated in reforecasts from models participating in the Subseasonal Experiment (SubX) and Subseasonal to Seasonal (S2S) prediction projects. SubX and S2S reforecasts show MJO prediction skill out to 4.5 weeks based on the Real-time Multivariate MJO (RMM) index consistent with previous studies. However, a closer examination of these models’ representation of MJO propagation through the MC reveals that they fail to predict the MJO convection, associated circulations, and moisture advection processes beyond 10 days with most of models underestimating MJO amplitude. The biases in the MJO propagation can be partly associated with the following mean biases across the Indo-Pacific: a drier low troposphere, excess surface precipitation, more frequent occurrence of light precipitation rates, and a transition to stronger precipitation rates at lower humidity than in observations. This indicates that deep convection occurs too frequently in models and is not sufficiently inhibited when tropospheric moisture is low, which is likely due to the representation of entrainment.
1. Introduction

The Madden-Julian Oscillation (MJO, Madden and Julian 1971, 1972) is the leading source of global subseasonal predictability. There have been enormous advances in MJO prediction in the last decade; leading dynamical forecasting systems are skillful out to five weeks (e.g., review by Kim et al. 2018). Several internationally coordinated multi-model reforecast experiments have made such advances possible. For example, the World Weather Research Programme/World Climate Research Programme Sub-seasonal to Seasonal Prediction project (hereafter S2S project, Vitart et al. 2017) was launched in 2013 and provides both near real-time and a long record (> 10 years) of extended-range (up to 60 days) forecasts from operational centers with the aim of better understanding global subseasonal predictability and its translation into useful forecast information. Many S2S project models have demonstrated improved MJO prediction compared to their earlier versions (e.g., Vitart 2017, Lim et al. 2018). Another newly launched Subseasonal Experiment (SubX, Kirtman et al. 2017), which is a National Oceanic and Atmospheric Administration (NOAA) Climate Testbed project, consists of multi-models from the current generation prediction systems by multi-agencies from North America, such as the NOAA, National Aeronautics and Space Administration (NASA), and Naval Research Lab (NRL). The SubX project produces real-time forecasts as well as reforecasts going back to 1999. MJO prediction skill in SubX models is comparable to those from the S2S project (Janiga et al. 2018; Pegion et al. 2019). These ongoing internationally coordinated efforts on subseasonal prediction provide an unprecedented opportunity to further understand the predictability of the MJO.

Although studies have shown continuous improvements in MJO prediction in the past decade, simulation of MJO eastward propagation, particularly across the Maritime Continent
(MC), remains a challenge. When the MJO propagates eastward from the Indian Ocean to the western Pacific, it is often disrupted by the MC due to orography and diurnal convection; this is referred to as the MC barrier effect (e.g., review by Kim et al. 2019). This MC barrier effect is exaggerated in dynamical forecasting systems; the percentage of MJO events not crossing the MC is significantly higher in forecasts than in observations indicating the inability of models to maintain MJO propagation through the MC (Neena et al. 2014; Kim et al. 2014, 2018; Wang et al. 2014; Xiang et al. 2015; Liu et al. 2017; Vitart 2017; Wang et al. 2018). Better understanding of the sources of model errors in MJO propagation processes is crucial for improving MJO prediction skill.

Recent theoretical, observational, and modeling studies demonstrate that the propagation of the MJO is tightly linked to the seasonal mean state. Predicated on the moisture mode framework (e.g., Raymond and Fuchs 2009; Sobel and Maloney 2012, 2013; Adames and Kim 2016), it has been argued that the advection of the low-tropospheric seasonal mean moisture by the MJO associated circulation anomalies plays an important role in the eastward propagation of the MJO through the MC (Jiang et al. 2015, 2018; Adames and Kim 2016; Jiang 2017; Kim et al. 2019). Specifically, the horizontal gradient of the mean moisture determines the magnitude of the moistening (drying) to the east (west) of the enhanced MJO, thereby enabling or disabling the eastward propagation. In global climate models (GCMs), mean biases develop quickly within a week in the tropics (e.g., Janiga et al. 2018) and can distort the MJO simulation and prediction by influencing further development of the MJO and its interaction with the mean state (Gonzalez and Jiang 2017; Jiang 2017; Kim 2017; Lim et al. 2018). Kim (2017) found that the predicted MJO amplitude rapidly becomes weaker than the observed amplitude when it propagates across the MC in the European Centre for Medium-Range Weather Forecasts (ECMWF) model. The
weak MJO in the model is mostly attributed to reduced horizontal moisture advection which is partly caused by the dry bias in the low-tropospheric seasonal mean moisture. Consistent results are found with S2S project models by Lim et al. (2018) which shows that models with smaller biases in the mean moisture gradient tend to have higher MJO prediction skill. However, a lack of detailed analysis of MJO propagation processes in Lim et al. (2018) and the use of a single model in Kim (2017) motivate further research on the relationship between MJO prediction skill, MJO propagation processes, and mean state biases in a diverse set of models.

Better understanding the sources of errors in MJO processes is the key to improving MJO prediction. Process-based diagnostics aim to ascertain which physical processes should be targeted for improvement in models to better simulate the MJO and help focus model development efforts (e.g., review by Jiang et al. 2019; Maloney et al. 2019). However, process-based model evaluations of MJO prediction are limited in number and largely limited to case studies of individual MJO events (Ling et al. 2014; Hannah and Maloney 2014; Hannah et al. 2015; Klingaman et al. 2015). The SubX and S2S projects, which make available a large number of ensemble forecasts of numerous fields covering multiple years from current dynamical prediction systems, provide an opportunity to conduct a thorough process-based study of MJO prediction. The process-based approach in this study sheds light on which physical processes in the dynamical models, especially in the cumulus parameterization, needs to be better represented to further improve MJO prediction. The SubX and S2S reforecasts and validation data sets are introduced in section 2. MJO prediction skill, amplitude and phase biases are assessed in section 3. MJO propagation processes and mean state biases are discussed in section 4, followed by summary and discussion in section 5.
2. Data and method

Reforecasts from the SubX project include the National Centers for Environmental Prediction Environmental Modeling Center Global Ensemble Forecast System (NCEP-GEFS), NASA Global Modeling and Assimilation Office Goddard Earth Observing System (NASA-GEOS5), Naval Research Laboratory Navy Earth System Prediction Capability (Navy-ESPC), National Center for Atmospheric Research (NCAR) Community Climate System Model version 4 run at the University of Miami Rosenstiel School for Marine and Atmospheric Science (RSMAS-CCSM4), and NOAA Earth System Research Laboratory Flow-Following Icosahedral Model (ESRL-FIM). In addition to these five SubX models, NCAR Community Earth System Model version 1 (NCAR-CESM1) reforecasts with two different vertical levels (L30 and L46) are compared. Because of the comparable MJO simulation characteristics and prediction skill in L30 and L46, we combine them into a 20-member ensemble of 10 members each. Two models from the S2S project, ECMWF version CY43R3 (ECMWF-CY43R3) and Korea Meteorological Administration-UK Met Office coupled Global Seasonal forecast (KMA-GloSea5), are evaluated. Those two S2S project models are chosen due to their relatively high MJO prediction skill (Lim et al. 2018). The ECMWF-CY43R3 version is the most updated version which consists of 20-year reforecasts. All models have fully coupled atmosphere-ocean-land-sea ice models, except the NCEP-GEFS which is forced by prescribed sea surface temperature (SST) from the operational NCEP Climate Forecast System v2 (Pegion et al. 2019). Ensemble sizes, reforecast periods, and initialization intervals in eight reforecasts are compared in Table 1. Readers are referred to Vitart et al. (2017) and Pegion et al. (2019) for a more detailed description of each model configurations. All models have reforecasts out to a minimum of 32 days. The results shown in this study are from the ensemble mean unless otherwise stated.
ECMWF Reanalysis Interim (ERA-interim, Dee et al. 2011), NOAA Advanced Very High-Resolution Radiometer (Liebmann and Smith 1996) Outgoing Longwave Radiation (OLR) product from 1979 to 2017, and Global Precipitation Climatology Project 1° daily version 1.2 (GPCP-IDD, Huffman et al. 2001) from 1997 to 2015 are compared with the reforecasts. We refer to these datasets as observations for simplicity. All observations and reforecasts are interpolated to a 1° latitude and 1° longitude grid. To define the MJO phase and amplitude, the Real-time Multivariate MJO (RMM, Wheeler and Hendon 2004) indices are calculated with the OLR and zonal winds at 850 and 200hPa. Predicted RMMs are obtained by projecting the predicted anomalies onto the observed combined EOF eigenvectors following Vitart (2017). RMM amplitude is defined as $\sqrt{RMM1^2 + RMM2^2}$ (Wheeler and Hendon 2004). We evaluate MJO predictions during boreal winter (reforecasts initialized between October 1st and March 31st). Climatology and anomalies are calculated over each day of the reforecast period, which ranges from 17 to 20 years among the models (Table 1). Given the long duration of the reforecast periods from each model, the composites and MJO skill presented in this study constitute a representative distribution of MJO events and differences in reforecast period and initialization interval among the models do not alter the overall conclusions.

3. MJO Prediction and Propagation

3.1 MJO prediction skill and biases

We first compare the MJO prediction skill, mean amplitude bias and mean phase biases based on conventional skill metrics described in Kim et al. (2018). Figure 1a shows the RMM bivariate correlation coefficients (RMM skill, hereafter) as a function of forecast lead time during boreal winter for all MJO initial phases and amplitudes. The lead time when models reach
correlation coefficient 0.5 (an indicator of useful skill) is about 20-22-day in RSMAS-CCSM4, Navy-ESPC, and ESRL-FIM, and 25-27-day in KMA-GloSea5, NCEP-GEFS, NASA-GEOS5, NCAR-CESM1, and 33-day in ECMWF-CY43R3. Compared to the previous version (CY40R1) of the ECMWF system (Kim 2017), RMM skill is increased about 2-3 days in CY43R3. Overall, SubX models have skill comparable to S2S models (Vitart 2017; Lim et al. 2018; Janiga et al. 2018; Pegion et al. 2019). As demonstrated in many studies, RMM skills are greater (lower) for forecasts with strong (weak) MJO events at initialization time (not shown). Generally, models with a relatively larger number of ensembles or with more frequent initialization (ECMWF-CY43R3, NCAR-CESM1, NASA-GEOS5) show higher skill, although the sensitivity of ensemble number and initialization frequency to the RMM skill needs to be examined in more detail.

Hereafter, only events with initially strong MJO (RMM amplitude >1.0) during winter are examined unless otherwise stated. The total number of selected strong MJO events in each reforecast are indicated in Figure 2. For example, for NCEP-GEFS, 270 events are selected based on the observed RMM amplitude and phase at day-0 of each forecast and compared with the observation as a function of forecast lead time. Figure 1b shows the mean amplitude bias relative to the observation for initially strong MJO cases. The mean amplitude bias is $\frac{1}{n} \sum (F_A' - O_A')$, where $n$ is the number of selected events and subscript $A$ denotes the RMM amplitude of ensemble mean forecast anomalies ($F'$) and observed anomalies ($O'$) as a function of forecast lead time. Consistent with S2S models (Vitart 2017; Lim et al. 2018; Wang et al. 2018), most of the models tend to lose the amplitude as forecast time increases (Fig. 1b). Some models (NASA-GEOS5 and RSMAS-CCSM4) overestimate the amplitude during the first two weeks. Consistent with Janiga et al. (2018), Navy-ESPC produces stronger amplitudes than
observation after 10-days. To better understand the source of the bias, amplitude bias is compared by initial MJO phases (Fig. 2). During the first 10-days (gray vertical line), most models underestimate the amplitude when MJO locates initially over the Indo-Pacific warm pool (phase 1-4). NCAR-CESM1 and ECMWF-CY43R3 have overall weaker amplitude than the others. The reason for underestimation of MJO amplitude will be examined in later sections.

Figure 1c shows the mean phase bias for initially strong MJO cases. The mean phase bias is defined as $\frac{1}{n} \sum (F_p - O_p)$, where subscript P denotes the phase defined as the phase angle (°) on the RMM phase-space diagram (details in Kim et al. 2018). Similar to S2S models (Vitart 2017; Lim et al. 2018), all SubX models show negative values throughout the whole forecast period, indicating a slower propagation relative to the observation. When divided into initial MJO phases, this slower propagation appears in most of phases (not shown). The phase bias could be sensitive to MJO metrics. For example, the Navy-ESPC MJO was shown to be too fast in OLR (Janiga et al. 2018) while it is slow in the RMM measure.

3.2 MJO propagation

To further understand the predicted characteristics and biases associated with MJO eastward propagation over the Indo-Pacific, strong MJO events initialized at phase 1 and 2 are selected. Phase 1 and 2 are the phases in which the MJO convective signal is located over the central Indian Ocean and in which most models predict weaker and slower propagation through the MC than observed (Fig. 1 and 2). Although the spatial patterns of the MJO are slightly different between phase 1 and 2, we combine these two phases to increase the sample size. To avoid the double counting in observation due to sequential days that meets the criteria in a single strong MJO event, we screen the data with a 5-day interval which approximately matches with the initialization frequency in the reforecasts. In observations, a total of 246 initially strong MJO
phase 1 and 2 events are selected over 38 winters (1979-2017), while it ranges from 66 to 277 in
models due to different initialization frequency and reforecast period. A 5-day moving average
and 9-grid-point smoothing are applied for better visualization for Figure 3-5 and 7-8 (indicated
in each figure). Day-01 composite, for example, is the average of forecast day 1 to 5, day-02 is
average of day 2 to 6, and so on. Observed MJO propagation characteristics will be discussed
first, followed by model comparison.

Figure 3 shows the composites of OLR and 850 hPa horizontal wind anomalies at forecast
day-01. In observations (Fig. 3a), enhanced convection is centered over the equatorial Indian
Ocean and suppressed convection extends over a broad region from the eastern Maritime
Continent to the western Pacific (approximately from 110°E to 180°E). Strong equatorial
easterly wind anomalies to the east of the convective envelope are associated with the Kelvin
wave response to the convective heating, and the Rossby wave response enhances the easterlies
and poleward flow to the west of the suppressed anomaly (e.g., Adames and Kim 2016). After 10
days (Fig. 4a), the MJO convective envelope has moved to the MC with particularly strong
convection between the MC and Australia. The MJO-associated circulation induces positive
zonal advection of the seasonal mean moisture to the southern MC, thus helping the MJO to
detour south of the MC (Kim et al. 2017). At day-20 (not shown), the observed MJO convective
envelope propagates into the Pacific warm pool and the suppressed anomaly appears in the
Indian Ocean which is roughly a mirror image of the day-01 pattern. To visualize the broad-scale
observed MJO propagation, OLR and zonal wind at 850 hPa (U850) anomalies are averaged
over 20°S-20°N and displayed as a function of longitude and forecast lead (Fig. 5a). The OLR
anomaly weakens as it approaches the MC (west of 120°E) at about day-10 and re-amplifies as it
slowly propagates through the Timor Sea (around 120°E), a narrow channel between the
southern part of Indonesia and Northern Australia. Then, the convective anomaly approaches the
dateline at around day-20. After day-10, the suppressed convective anomaly appears in the
western Indian Ocean and propagates eastward. The U850 shows continuous eastward
propagation associated with the convective anomalies (Fig. 5a).

Overall, the SubX and S2S reforecasts capture the broad MJO propagation structure in the
OLR and wind anomalies but with phase and amplitude biases (Fig. 3-5). At day-01 (Fig. 3),
which is an average of the first five days, models have a comparable pattern with the
observation. However, the predicted amplitude in OLR in some models (NCEP-CESM1 and
ECMWF-CY43R3) is already weaker than the observed, while others (RSMAS-CCSM4 and
NASA-GEOS5) overpredict the amplitude, loosely consistent with Figure 2. The initial biases
could be attributed to different convective parameterization because OLR bias develops quickly.
Initialization data and assimilation methods can also impact the model difference at day-01,
although reasons behind diverse initial biases are beyond the scope of this study. At day-10,
predicted OLR and U850 anomalies become more diverse between the models in both phase and
amplitude (Fig. 4).

3.3 MJO propagation skill

Several earlier studies have examined the development of biases in MJO phase and
amplitude in dynamical forecasting systems using the RMM index (such as Fig. 1b and c).
However, because the RMM index is global, it is an indirect measure of the ability of a model to
capture MJO propagation through the MC. Moreover, because the fractional contribution of
upper-level zonal wind (U200) to the RMM is considerably higher than the contribution of OLR
and U850 anomalies, the RMM skill and biases mainly reflects the predicted U200 (e.g., Straub
2013), leading to an overly optimistic estimate of the ability of a model to predict OLR and U850
anomalies. The RMM skill, amplitude and phase biases shown in Figure 1 and 2, therefore, do not directly translate the biases emanating from the propagation of MJO convection and low-level circulation which are more closely linked to cumulus parameterizations.

To directly measure the MJO propagation ability in models, mean pattern correlation coefficient (PCC) and mean absolute amplitude (AMP) are calculated. Strong MJO events initialized at phase 1 and 2 are selected. PCC is a pattern correlation coefficient between individual observed and ensemble mean forecast anomalies calculated over the tropical Indo-Pacific (40°E-200°E and 20°S-20°N) and then averaged over all selected events in each model. AMP is the absolute amplitude of individual ensemble mean forecasted (or observed) anomalies averaged over the tropical Indo-Pacific (40°E-200°E and 20°S-20°N) and then averaged over all selected events in each model. To compare the change in skill as a function of exact lead time, the five-day moving average is not applied when calculating the PCC and AMP shown in Figure 6.

At day-01, models generally capture the observed dipole pattern of the MJO envelope over the Indo-Pacific (Fig. 3), resulting in a high PCC in all variables (Fig. 6a-c). Then the PCC decreases in all three variables reaching 0.5 before 10 days in all models with a faster drop in PCC_{OLR} than PCC_{UB50} or PCC_{U200}. As mentioned earlier, the RMM skill (> 3 weeks, Fig. 1a) leads to overly optimistic estimate of model MJO prediction ability. The relationship between RMM skill score for phase 1 and 2 (similar to Fig. 1a) only loosely matches the PCC of OLR and 850 and 200 hPa zonal winds. However, similar, to Fig. 1b, a continuous loss of amplitude (AMP) can be found in all models (Fig. 6d-f), except Navy-ESPC. In addition, we find little relationship between the magnitude of the AMP biases and the PCC. A particularly fast decrease in AMP is seen in NCAR-CESM1, ECMWF-CY43R3, NCEP-GEFS which generally produce
high RMM or PCC skill. The range of biases of $AMP_{OLR}$ are larger compared to the biases in circulations fields at the beginning of forecasts (< 1 week), indicating a quick development of convection biases. NCAR-CESM1, ECMWF-CY43R3, and NCEP-GEFS underestimate the $AMP_{OLR}$, while the rest of models overestimate it at the beginning of forecasts (Fig. 6d).

4. MJO Processes and Mean State

4.1 Moisture advection processes

To understand the sources of the biases in the MJO forecasts it is crucial to understand the representation of the processes associated with the MJO propagation. As mentioned in the introduction, studies have shown that the advection of the mean moisture by the MJO anomalous circulation is the key process that increases moisture to the east and decreases moisture to the west of the envelope of enhanced moisture associated with the MJO, thus controlling the propagation (e.g., D. Kim et al. 2014; Adames and Kim 2016; Jiang et al. 2018). Therefore, GCMs with dry mean biases tend to have weaker moisture advection and worse MJO propagation (Gonzalez and Jiang 2017) and lower MJO prediction skill (Lim et al. 2018). Here, we test whether the relationship between MJO propagation process and mean moisture applies to the models from SubX and S2S. The seasonal mean moisture advection by the anomalous wind, $-V' \cdot \nabla \bar{Q}_{850}$, is calculated where $V'$ is the horizontal wind anomaly at 850 hPa and $\bar{Q}_{850}$ is the winter climatology of specific humidity at 850 hPa, where the overbar denotes the climatology defined as the average of the first 4-weeks (1 to 28 forecast lead days). In order to smooth the horizontal moisture gradient ($\nabla \bar{Q}_{850}$), especially the strong gradients in the vicinity of the MC
region due to land-ocean contrasts, we interpolate both $\overline{Q}_{850}$ and $V'$ from 1.0° to 2.5° in both observations and reforecasts.

Studies have typically focused on moisture at 700 hPa (e.g., Gonzalez and Jiang 2017) but $Q_{850}$ is the only common moisture variable available in the SubX reforecasts. One may argue that 850 hPa level in the tropics is only marginally above the planetary boundary layer and $Q_{850}$ may not represent the column moisture field. To justify the use of $Q_{850}$ in representing the column moisture, we compare the column-integrated $Q$ ($\langle Q \rangle$) and $MSE$ ($\langle MSE \rangle$) calculated by ERAI over eight vertical levels from 1000 hPa to 100 hPa. Then, variables averaged over the Indo-Pacific (Fig. 7, green box, 80°E-130°E, 20°S-15°N) are compared as a function of forecast lead days when MJO starts from phase 1 and 2. The tendency and horizontal advection of $Q_{850}$, $\langle Q \rangle$, and $\langle MSE \rangle$ are similar to each other (not shown), although the overall pattern of $Q_{850}$ is noisier than the column-integrated variables.

Figure 7 shows the anomalous horizontal advection of winter mean specific humidity ($-V' \cdot \nabla \overline{Q}_{850}$) and anomalous wind at 850 hPa ($V'$) at day-01 (day 1 to 5 average) in observations and the reforecasts. In observations (Fig. 7a), the easterly and poleward wind to the east of the convection anomaly and the winter mean moisture ($\overline{Q}_{850}$) distribution (Fig. 9a) together induce moisture advection on the poleward side of the MC consistent with previous studies (D. Kim et al. 2014; Wang et al. 2017). The moisture recharge process is clearer if free tropospheric (such as 700 hPa) moisture observations are used but the overall behavior is similar (not shown). The positive moisture advection over the MC at day-01 helps the MJO in the Indian Ocean to propagate to the MC region after ten days as shown in Fig. 4a. To summarize the evolution of the moisture advection, $-V' \cdot \nabla \overline{Q}_{850}$ is averaged over a broad MC area (green box in Fig. 7). In observations, $-V' \cdot \nabla \overline{Q}_{850}$ is positive for up to two-weeks indicating a continuous moisture
advection to the MC area (Fig. 8a). This sustained moisture advection helps the MJO convective anomaly to propagate through the MC. After about two weeks, negative moisture advection helps the suppressed MJO phase to propagate over the MC. The moisture advection from $-V' \cdot \nabla \bar{Q}_{850}$ is approximately in phase with the moisture tendency and dominant among other terms and 90º out of phase with the moisture and OLR anomaly (not shown) as discussed in previous studies (e.g., D. Kim et al. 2014; Jiang 2017).

Models generally capture the pattern of the $-V' \cdot \nabla \bar{Q}_{850}$ at day-01 (Fig. 7b-i). However, the amplitude of moisture advection declines rapidly in all models (Fig. 8a), which gives rise to a weaker amplitude of moisture advection than the observed during the entire forecast period. During the first 10-days (Fig. 8b), the multi-model mean (MMM) of moisture advection is only 61.7 % of the observed value, and individual model ranges from 41.1 % (ESRL-FIM) to 80.0 % (NASA-GEOS5). It can be concluded that all SubX and S2S models predict weaker moisture advection than is seen in observations. This is consistent with the underprediction of MJO amplitude in most of the models (Fig. 6). The underprediction of MJO amplitude, however, could be associated with other MJO intensification processes as well, such as surface-fluxes (e.g., Sobel et al. 2010) or radiative-convective feedbacks (e.g., Raymond 2001). In addition, these processes could explain why the Navy-ESPC system overpredicts the amplitude of the MJO. Navy-ESPC has a negative OLR mean state bias over the active MJO region (not shown) which suggests it may be overestimating the strength of radiative-convective feedbacks.

4.2 Mean moisture biases

Two variables can impact the horizontal moisture advection ($-V' \cdot \nabla \bar{Q}_{850}$): wind anomalies ($V'$) and the mean moisture distribution ($\nabla \bar{Q}_{850}$). In the previous version of the ECMWF system, the quickly developing mean moisture bias has larger contributions to the moisture advection
bias than the MJO wind anomalies (Kim 2017). Observed $\tilde{Q}_{850}$ and model biases of $\tilde{Q}_{850}$ are compared in Figure 9 with land areas masked out due to the influence of terrain. During boreal winter, observed $\tilde{Q}_{850}$ is maximized in the tropical eastern MC, inducing both strong zonal and meridional gradient of $\tilde{Q}_{850}$, while all reforecasts show dry biases over the tropical Indo-Pacific region (Fig. 9). RSMAS-CCSM4, NCEP-GEFS, KMA-GloSea5 have an approximately 5% reduction of $\tilde{Q}_{850}$ in the tropical Indo-Pacific (40ºE-200ºE, 15ºS-15ºN) relative to observations, while Navy-ESPC, ESRL-FIM, and NASA-GEOS5 have an approximately 10% reduction. Dry biases are relatively small (<1 %) in NCAR-CESM1 and ECMWF-CY43R3.

The meridional moisture advection has a larger contribution to MJO propagation than the zonal component in both observation and reforecasts (D. Kim et al. 2014; Kim 2017; Lim et al. 2018). The meridional gradient of $\tilde{Q}_{850}$ (hereafter $\tilde{Q}_y$) is defined by simply taking the difference between the equatorial (80ºE-130ºE, 10ºS-5ºN) and average of two subtropical regions (80ºE-130ºE, 5ºN-15ºN and 80ºE-130ºE, 20ºS-10ºS), similar to Lim et al. (2018). Because of the decrease of moisture in poleward direction, observation shows strong negative meridional gradient ($\tilde{Q}_y = -1.7$ g/kg) while all models predict weaker meridional gradient ranging from -1.3 (ECMWF-CY43R3) to positive values of 1.0 g/kg (NASA-GEOS5). Simulation of positive $\tilde{Q}_y$ (poleward increase of mean moisture) in NASA-GEOS5 and ESRL-FIM is due to weak moisture in the MC area where the observation has a maximum. Basin-wide dry bias and weaker meridional moisture gradient can partially contribute to weaker horizontal moisture advection shown in Fig. 8.

Kim (2017) and Lim et al. (2018) argued that models with larger mean moisture pattern biases have weaker moisture advection, fast damping of MJO propagation, and thus lower MJO prediction skill. However, among the models examined here, we do not find significant linear
relationships between the mean moisture pattern, moisture advection, propagation and prediction skills. For example, NASA-GEOS5 shows the largest biases in $\bar{Q}_{850}$ and $\bar{Q}_{y}$ (Fig. 9h) but has the strongest moisture advection ($-V' \cdot \nabla \bar{Q}_{850}$) among models (Fig. 8). On the other hand, ECMWF has the smallest mean state bias but shows fast damping of the moisture advection, but high prediction skill. This is because moisture advection ($-V' \cdot \nabla \bar{Q}_{850}$) could also be attributed to the MJO circulation anomalies ($V'$). Weaker moisture advection shown in ECMWF-CY43R3 could result from the weaker wind anomalies, although the mean state bias is simulated reasonably. Stronger moisture advection in NASA-GEOS5 could result from the stronger wind, although the mean state bias is worse. Both models, however, have high RMM skills (Fig. 1a). Therefore, from eight models, it is hard to link the mean state bias directly to moisture advection performance, propagation and prediction skill. Biases in MJO associated fields (e.g. convection) and mean state cannot be clearly separated due to their interaction and tight coupling between each other. Further research with sensitivity experiments is needed. Nonetheless, it is clear that all SubX and S2S models examined in this study have dry biases in low-tropospheric moisture and simulate weaker moisture advection than observation which could potentially impact MJO propagation and limit the prediction skill.

4.3 Precipitation-moisture coupling and precipitation biases

To further understand the sources of the low-tropospheric dry biases shown in the SubX and S2S models, the simulation of precipitation-moisture coupling is evaluated. Tropical moisture is strongly coupled to precipitation and convection processes (e.g., Holloway and Neelin 2009) and the representation of these processes are sensitive to convective parameterization in numerical models. In the GPCP observations (Fig. 10a), the maximum winter mean precipitation is along the Intertropical Convergence Zone (ITCZ), South Pacific Convergence Zone (SPCZ) and over...
the MC. Figure 10 shows that models over predict precipitation over the western Indian Ocean as well as western Pacific with a well-known double ITCZ biases (Lin 2007). In summary, models tend to predict excess surface precipitation and have a lower-tropospheric dry bias over the Indo-Pacific region. These mean biases develop quickly and saturate in the first week.

To further understand the source of precipitation biases, the occurrence frequency (%) of daily total precipitation rates are compared (Fig. 11a). To calculate the occurrence frequency, first the number of occurrences in each precipitation rate bin is divided by the total number of days for each forecast lead time. Then the average across the first 4 weeks over the Indo-Pacific Ocean (60°E-180°E, 15°S-15°N) is computed from this; land area is masked out since convection is typically forced by the diurnal cycle more than moisture-convection coupling (Ahmed and Schumacher 2017). Only control simulations are used from the reforecasts in these calculations. Precipitation rates are binned with logarithmic bin sizes to account for more frequent events at low precipitation rates (Fig. 11a). Precipitation rates less than 0.01 mm/day are cut off. GPCP observations have a broad maximum in occurrence between 10 to 20 mm/day consistent with TRMM precipitation estimates (Kim et al. 2015). NCEP-GEFS has uneven distribution for precipitation rates at < 10 mm/day, and such distributions is seen in other areas, while the reason is not clear at this stage. All models produce frequent precipitation in light regimes (< 10 mm/day) and most of them underestimate the heavy precipitation regime (> 20 mm/day), except ESRL-FIM, NCEP-GEFS, and Navy-ESPC which overestimate frequency for > 40 mm/day. The most frequent precipitation rate is 15 mm/day in observations, 10 mm/day in NCAR-CESM1, ECMWF-CY43R3, and KMA-GloSea5, 5 mm/day in RSMAS-CCSM4 and NASA-GEOS5 and 2 mm/day in ESRL-FIM and Navy-ESPC. NCAR-CESM1, ECMWF-CY43R3, and KMA-GloSea5 have almost twice as frequent precipitation as that observed around 10 mm/day.
Overall, SubX and S2S reforecasts produce light precipitation that is too frequent which may be associated with the dry biases in the low-tropospheric mean state (Fig. 9).

Previous studies have shown the importance of low-tropospheric moisture to the onset of deep convection in the tropics (e.g., Bretherton et al. 2004; Peters and Neelin 2006; Holloway and Neelin 2009; Rushley et al. 2018). Over tropical oceans, a sharp increase in precipitation rate occurs when column relative humidity reaches about 80% saturation and continues to increase exponentially at higher saturations. However, this precipitation pickup tends to occur earlier than observed in most GCMs which causes these models to overestimate precipitation in dry regimes and underestimate it in humid regimes (Rushley et al. 2018). This common moisture-precipitation coupling bias can directly influence the simulation of MJO propagation (Kim et al. 2012; Jiang et al. 2015; Ahn et al. 2017). To evaluate the predicted moisture-precipitation coupling in the SubX and S2S reforecasts, daily total precipitation and $Q_{850}$ are compared. Note that $Q_{850}$ is the daily mean specific humidity field from observation and control simulations, not the climatology ($\bar{Q}_{850}$). Figure 11b shows the mean distribution of precipitation rate in each $Q_{850}$ bin averaged over the first 4-weeks over the tropical Indo-Pacific (60°E-180°E, 15°S-15°N) with land areas masked out. These results are not sensitive to the choice of the forecast lead since the biases develop in the first week of the forecasts. In observations, precipitation pickup begins around 10.5 g/kg, increases slowly up to about 12 g/kg, and increases exponentially beyond that (Fig. 11b). The inset in Figure 11b provides a closeup of the precipitation rates during drier conditions and shows the earlier onset of deep convection in all models compared to the observation. Recall from Figure 11a that these light precipitation rates are not rare and in fact occur more frequently in models than in observations. In summary, models overall predict larger
mean precipitation rates and more frequent precipitation in the lower humidity regime than in observations.

5. Summary and Discussion

In this study, MJO prediction skill, propagation processes, mean state biases, and precipitation-moisture coupling are evaluated in reforecasts from models in the SubX and S2S databases. Overall, SubX models show comparable RMM skill ranges (3-4 weeks) to S2S models and capture the broad structure of the MJO convective envelope at the beginning of the forecast. However, SubX and S2S models fail to predict the propagation of the MJO convection beyond 10-days with most of them having a fast decay in the amplitude of the convection associated with the MJO. To better understand the biases associated with the MJO propagation, we evaluate each model’s capability to represent moisture advection processes and relate this to mean moisture biases through the framework of the moisture mode theories (e.g., Raymond and Fuchs 2009; Sobel and Maloney 2012, 2013). Compared to observations, the horizontal moisture advection east of the MJO convective anomaly is underestimated in all models, primarily due to the underestimation of the meridional component of the horizontal moisture advection. This weak horizontal moisture advection is partly associated with dry biases in the lower-tropospheric across the Indo-Pacific.

Lim et al. (2018) demonstrated that MJO prediction skill and the bias pattern of the mean column-integrated water vapor are significantly correlated in seven S2S models. In this study, we do not find a strong correlation between biases in the horizontal moisture advection representation or mean moisture pattern with MJO prediction skill. This discrepancy could be due to our focus on moisture at 850 hPa, whereas Lim et al. (2018) used column-integrated
moisture. On the other hand, it could indicate that we do not yet fully understand all the different pathways in which biases in MJO propagation can arise due to errors in the representation of physical processes. For example, while our focus is mainly on horizontal moisture advection process, recent studies have shown that the vertical moisture advection also plays an important role in MJO propagation (Janiga and Zhang 2016; Wang et al. 2017). MJO propagation biases can also emanate from other processes such as the longwave cloud-radiation feedback (Lim et al. 2018), air-sea coupling (e.g., Demott et al. 2016), and diurnal cycle in the MC region (Zhang and Ling 2017). A more complete understanding of the contribution of these processes to errors in MJO propagation would lead to further improvements in MJO prediction skill.

SubX and S2S reforecasts contain the following biases over the tropical Indo-Pacific: a dry lower troposphere, excess of surface precipitation, more frequent occurrence of light precipitation rates, and a transition to stronger precipitation rates that begins at lower humidity than is seen in observations. This indicates that convection occurs too frequently in models and is not sufficiently inhibited when tropospheric moisture is low, which is likely due to the representation of entrainment. Previous studies have shown that changes in the representation of entrainment through increasing the sensitivity of deep convection to the environmental humidity generally improves MJO simulation (Tokioka et al. 1988; Bechtold et al. 2008, Hannah and Maloney 2011; Kim et al. 2012; Hirons et al. 2013, Klingaman and Woolnough 2014; Jiang et al. 2015) while degrading the mean state (Kim et al. 2011). Through increased entrainment rate, more moisture is likely to build up and precondition the deep convection while suppressing premature development of convection, thus improving MJO simulation (e.g., review by Kim and Maloney 2017). In other words, with low entrainment rate in the deep convective scheme, convection occurs before the lower troposphere is sufficiently moistened (i.e., insufficient
conditional instability) (Kuo et al. 2017). This results in a moisture-depleted atmosphere because
atmospheric moisture can be removed efficiently through convection-induced precipitation thus
the environment favors a low humidity state (Kuo et al. 2017). Given that the SubX and S2S
models have precipitation occurring in the lower humidity regime and possess climatological dry
biases in the lower-troposphere, it is likely that those models have lower sensitivity of the
convection to environmental moisture which can be related to the entrainment rate. However,
with the limited output variables and lack of detailed information about each model’s convection
scheme, discussing the impact of model parameterization to MJO prediction is beyond the scope
of this study. Additional output of model variables would help to resolve these issues.

490

491
Acknowledgments

HK was supported by NSF grant AGS-1652289, NOAA grant NA16OAR4310070, and the KMA R&D Program grant KMI2018-03110. GloSea5 reforecast was provided by Dr. Myong-In Lee from UNIST. MJ and the generation of the Navy-ESPC forecasts were supported by the Office of Naval Research and N2N6E through the Navy Earth System Prediction Capability Effort, and the Chief of Naval Research through the NRL Base Program (PE 0601153N). Navy-ESPC forecasts were performed as part of the NMME-SubX experiment funded by the NOAA-Modeling, Analysis, Predictions, and Projections (MAPP) Program (NOAA-OAR-CPO-2016-2004413). Computational resources were supported in part by a grant of time from the High Performance Computing Modernization Program and the Navy-ESPC forecasts were integrated at the Navy Department of Defense Supercomputing Resource Center, Stennis Space Center, Mississippi. The sources of data used in this study are: http://cola.gmu.edu/kpegion/subx/ for SubX reforecasts, http://s2sprediction.net/ S2S reforecasts, http://apps.ecmwf.int/datasets/data/interim-full-daily for ERA-Interim.
References


Kim, H. M., P. J. Webster, V. E. Toma, and D. Kim (2014), Predictability and prediction skill of the MJO in two operational forecasting systems, J. Climate, 27 (14), 5364-5378.


Kirtman, B. P. and coauthors (2017), The Subseasonal Experiment (SubX) [Data set]. IRI Data Library. https://doi.org/10.7916/D8PG249H


27

Lim, Y., S. Son, and D. Kim (2018), MJO prediction skill of the subseasonal-to-seasonal prediction models. J. Climate, https://doi.org/10.1175/JCLI-D-17-0545.1


Table 1. Hindcasts data information.

<table>
<thead>
<tr>
<th>Model</th>
<th>Ensemble members</th>
<th>Initialization interval</th>
<th>Hindcast period</th>
<th>Source and References</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCEP-GEFS</td>
<td>11</td>
<td>1/week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NASA-GEOS5</td>
<td>4</td>
<td>every 5 days</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Navy-ESPC</td>
<td>4</td>
<td>1/week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSMAS-CCSM4</td>
<td>3</td>
<td>1/week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESRL-FIM</td>
<td>4</td>
<td>1/week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAR-CESM1</td>
<td>20</td>
<td>1/week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECMWF-CY43R3</td>
<td>11</td>
<td>2/week</td>
<td>1997-2016</td>
<td>S2S, Vitart et al. (2017)</td>
</tr>
<tr>
<td>KMA/UKMO-GloSea5</td>
<td>3</td>
<td>4/month</td>
<td>1991-2010</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: (a) RMM prediction skill (bivariate correlation coefficient) between the model ensemble means and observation. Horizontal line denotes correlation of 0.5. (b) Amplitude bias and (c) phase bias (º) relative to observation for initially strong MJOs. Negative value in (b) and (c) indicates weaker and slower MJO than in observations, respectively.
Figure 2: Same as Figure 1b, except for initial MJO phases. The vertical lines denote forecast lead day-10. Numbers in the parenthesis indicate the number of selected initially strong MJO events.
Figure 3: Day-01 composites of OLR (W/m², shading), 850-hPa wind vector, and wind speed (m/s, vector colors) anomalies for initially strong MJO events. Numbers in the parenthesis indicate the number of selected events. Gray (black) arrow indicates windspeed > 0.5 m/s (2 m/s). 5-day moving average and 9-grid-point smoothing are applied.
Figure 4: Same as Figure 3 except for day-10.
Figure 5: Longitude-time composites of OLR (W/m²; shading) and U850 (contour interval 0.4 m/s) anomalies averaged over 20°S-20°N for initially strong MJO events. The purple line indicates zero U850 anomalies. The zonal and vertical lines indicate day-10 and 120°E (approximately the center of the MC), respectively. 5-day moving average is applied.
Figure 6: Mean pattern correlation coefficient (PCC) and mean absolute amplitude (AMP) calculated over 40°E-200°E and 20°S-20°N as a function of forecast lead days. Horizontal lines indicate (a-c) correlation of 0.5 for PCC and (d-f) observed AMP. Only initially strong MJO events are selected.
Figure 7: Same as Fig. 3 except for $-V' \cdot \nabla \bar{Q}_{850}$ (units: $10^{-10}/s$, shading). Green box indicates the Maritime Continent area ($80^\circ E$-130$^\circ E$, 20$^\circ S$-15$^\circ N$). 5-day moving average and 9-grid-point smoothing are applied. Only initially strong MJO events are selected.
Figure 8: (a) Observed and predicted $-V' \cdot \nabla \bar{Q}_{850}$ averaged over the MC (green box in Fig. 7) as a function of forecast lead days (5-day moving average is applied) and (b) averaged for 1-10 days. MMM (red bar) denotes the multi-model mean of eight models. Only initially strong MJO events are selected.
Figure 9: $\overline{Q}_{850}$ (g/kg) in (a) observation and (b-h) biases in models. Land area is masked out. Numbers on the top-right indicates the meridional moisture gradient of $\overline{Q}_{850}$ ($Q_y$, g/kg) defined in the text.
Figure 10: Same as Figure 9, except for precipitation rate (mm/day).
Figure 11: (a) Occurrence frequency (%) of daily total precipitation rate and (b) distribution of daily total precipitation rate in each $Q_{850}$ bins in observation (GPCP) and reforecasts’ control simulations averaged over the first 4-weeks over the Indo-Pacific warm pool ($60^\circ$E-$180^\circ$E, $15^\circ$S-$15^\circ$N). Precipitation is on a log-scale in (a). The black vertical lines in (a) indicate precipitation at 10-mm/day.