JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2022MS003044

Special Section:
Data assimilation for Earth system models

Assessing the Impact of Ocean In Situ Observations on MJO Propagation Across the Maritime Continent in ECMWF Subseasonal Forecasts

Danni Du[1], Aneesh C. Subramanian[1], Weiqing Han[1], Ho-Hsuan Wei[1,2], Beena Balan Sarojini[3], Magdalena Balmaseda[3], and Frederic Vitart[3]

[1]Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, CO, USA, [2]Now at CIRES/NOAA Physical Sciences Laboratory, Boulder, CO, USA, [3]The European Centre for Medium-Range Weather Forecasts, Reading, UK

Key Points:

- Ocean in situ data assimilation has an impact on subseasonal sea surface temperature forecasts but has little impact on forecasting Madden–Julian Oscillation (MJO) propagation across Maritime Continent
- Moist static energy budget analysis reveals that the MJO forecast errors are likely due to the underestimated meridional moisture advection
- The intraseasonal meridional wind biases contribute more to the underestimated meridional moisture advection than the systematic dry biases

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. Du,
dadu9013@colorado.edu

Citation:

Du, D., Subramanian, A. C., Han, W., Wei, H.-H., Sarojini, B. B., Balmaseda, M., & Vitart, F. (2023). Assessing the impact of ocean in situ observations on MJO propagation across the Maritime Continent in ECMWF subseasonal forecasts. *Journal of Advances in Modeling Earth Systems*, *15*, e2022MS003044. https://doi.org/10.1029/2022MS003044

Received 11 FEB 2022
Accepted 26 JAN 2023

Abstract Despite the well-recognized initial value nature of the subseasonal forecasts, the role of subsurface ocean initialization in subseasonal forecasts remains underexplored. Using observing system experiments, this study investigates the impact of ocean in situ data assimilation on the propagation of Madden–Julian Oscillation (MJO) events across the Maritime Continent in the European Centre for Medium-Range Weather Forecasts (ECMWF) subseasonal forecast system. Two sets of twin experiments are analyzed, which only differ on the use or not of in situ ocean observations in the initial conditions. Besides using the Real-time Multivariate MJO Index (RMMI) to evaluate the forecast performance, we also develop a new MJO tracking method based on outgoing longwave radiation anomalies (OLRa) for forecast evaluation. We find that the ocean initialization with in situ data assimilation, though having an impact on the forecasted ocean mean state, does not improve the relatively low MJO forecast skill across the Maritime Continent. Moist static energy budget analysis further suggests that a significant underestimation in the meridional moisture advection in the model forecast may hinder the potential role played by the ocean state differences associated with data assimilation. Bias of the intraseasonal meridional winds in the model is a more important factor for such underestimation than the mean state moisture biases. This finding suggests that atmospheric model biases dominate the forecast error growth, and the atmospheric circulation bias is one of the major sources of the MJO prediction error and should be a target for improving the ECMWF subseasonal forecast model.

Plain Language Summary Assimilating ocean subsurface observations can provide a better ocean state estimate for initializing weather and climate forecast in models. This study assesses the impact of the ocean subsurface data assimilation on subseasonal forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF). For the subseasonal prediction, the Madden–Julian Oscillation (MJO)—the dominant tropical atmospheric variability at the subseasonal time scale—is commonly used to evaluate the model skill. This study specifically focuses on the MJO propagation across the Maritime Continent. By using two quantitative methods, we find ocean in situ observations have no impact on forecasting the MJO propagation. Furthermore, we find the coupled forecast model has atmospheric biases in its meridional wind field, which leads to an underestimation in the meridional moisture advection, thus potentially reducing the MJO prediction skill. Such an atmospheric origin of the model error may hinder the potential impact of ocean subsurface data assimilation. While previous research emphasized the importance of improving the mean state moisture representation in the model for MJO forecast, our findings emphasize the importance of improving the meridional wind representation.

1. Introduction

Ocean observations are vital for improving our understanding of oceanic variability and its impact on climate. While satellite observations have the advantage of covering global surface ocean with fine spatial and temporal resolution, in situ measurements can detect both near-surface and subsurface variability, even though the data are limited in space and time. To improve the ocean state estimate, ocean data assimilation (DA) is applied in numerical models, where the information gained from satellite and in situ observations is allowed to be propagated in time and space to uncovered regions using dynamical and physical constraints (Penny et al., 2019). The resultant ocean reanalysis products have been used to initialize the ocean component of coupled forecast models

(Balmaseda, 2017; Meehl et al., 2014). While the importance of subsurface ocean initialization with in situ DA in seasonal, interannual, and decadal climate predictions is widely recognized (Balmaseda et al., 2009; Meehl et al., 2014), how ocean in situ DA influences subseasonal predictions is yet to be explored.

Of particular interest is the Madden–Julian Oscillation (MJO; Madden & Julian, 1971, 1972), a large-scale phenomenon of deep convection and circulation, the dominant mode of atmospheric intraseasonal oscillations at 20–90 day periods in the tropical climate system. On subseasonal timescales, MJO is recognized as the leading source of predictability (e.g., H. Kim et al., 2021; Waliser et al., 2003). Therefore, the prediction skill of MJO is commonly used as an indicator of model subseasonal prediction skill.

The MJO can modulate weather and climate phenomena around the globe, and such modulation depends on the location of its convection center (e.g., Cassou, 2008; C. Zhang, 2013). The Maritime Continent (MC), an archipelago sitting in the Indo-Pacific Warm Pool connecting the tropical Indian and Pacific oceans, is one of such key locations in MJO propagation. As the MJO propagates eastward from the Indian Ocean to the Western Pacific, it undergoes the MC barrier effect—tending to decay and sometimes stall when crossing the MC region (Hendon & Salby, 1994; Matthews, 2008)—due to topography interference (Wu & Hsu, 2009), strong diurnal convection (Ajayamohan et al., 2021; Hagos et al., 2016), and many other factors. Using observational analysis, recent studies have investigated the sea surface temperature (SST) differences between the MJO events that pass the MC region and the MJO events that are stalled. Significant SST differences have been found both locally in the MC region (C. Zhang & Ling, 2017) and remotely in the southeast Indian Ocean and the western-central tropical Pacific (L. Zhang & Han, 2020). Surface fluxes cannot explain the SST difference between those two types of MJO events, thus it might be the subsurface ocean that contributes to such differences (C. Zhang & Ling, 2017; L. Zhang & Han, 2018), which makes the subsurface ocean initialization with in situ DA potentially crucial for predicting the MJO propagation across the MC.

To assess the impact of ocean in situ DA on subseasonal predictions, we focus on forecasting the MJO propagation across the MC region using Observing System Experiments (OSEs) with the European Centre for Medium-Range Weather Forecasts (ECMWF) coupled subseasonal forecast system. The paper is organized as follows. Section 2 describes the ECMWF subseasonal forecast system, the data used in this study and the metric selected to evaluate the forecast performance. Section 3 explores the forecasted ocean mean state biases with and without ocean in situ DA and their differences, examines impact of such differences on MJO propagation across the MC, and interprets the results. Section 4 provides a summary and discussion.

2. Data and Methodology

2.1. Data

2.1.1. OSE Experiments and Model Forecast

Two sets of OSEs have been carried out to understand the role of in situ observations (including Argo, Moorings, XBT, CDT and marine mammals) in the Ocean Reanalysis System 5 (ORAS5) of ECMWF (similar to Zuo et al., 2019). These OSEs have been conducted with the low resolution configuration of ORAS5 (O5-LR) at approximately 1° horizontal resolution and 42 vertical levels. The only difference with respect to the O5-LR experiments in Zuo et al. (2019) is that the two OSEs in this study do not assimilate altimeter data and the bias correction has been switched off. One experiment assimilates all in situ observations (*all_obs*), while the other experiment (*no_insitu*) does not. However, *no_insitu* still contains a substantial amount of information from the real world by being constrained by SST and atmospheric reanalyzes of surface fluxes, as *all_obs* does.

Coupled subseasonal forecasts are initialized from these two OSEs respectively. The two sets of 32-day forecasts are done with the ECMWF model cycle 47R1. They are initialized on the first day of each month from 1993 to 2015 with each set having five ensemble members. The atmosphere initial conditions are perturbed by singular vectors and ensemble data assimilation (Vitart et al., 2019). The ocean initial conditions are generated by perturbing both the observations and surface forcings (Zuo et al., 2017). In this way, the atmospheric initial conditions are identical for these two sets of forecasts correspondingly, and since SST nudging is on, the only major difference between these two sets of forecasts is the subsurface initialization. We name the two sets of forecasts as the *all_obs* experiment (initialized from the OSE with all in situ observations assimilated) and the *no_insitu* experiment (initialized from the OSE with no in situ observations assimilated).

From the model forecast, we examine SST, outgoing longwave radiation (OLR) and variables needed for the Moist static energy (MSE) budget analysis (introduced later in Section 3.3). All the variables above are mapped onto a 1° × 1° horizontal grid. We also analyze the daily Real-time Multivariate MJO Index (RMMI, Wheeler & Hendon, 2004) from the forecast output, which monitors the location and the amplitude of the MJO.

For SST, the 12 hourly instantaneous values from the forecast output are used to calculate the daily mean by averaging the values at 0 and 12hr on each forecast lead day. For OLR, the forecast output is in the form of an accumulated sum (e.g., the output at 72hr would be the integrated sum of OLR from 0 to 72hr); thus we use the 24 hourly instantaneous values to calculate the daily increment, and then get the daily mean OLR flux. For the MSE budget analysis, we choose the 24 hourly instantaneous output to do daily tendency calculation. Additionally, for a flux term, we treat it in a similar way as we treat OLR. For an advection-related variable (e.g., wind velocity, temperature, geopotential, and specific humidity), we average the instantaneous values of 0 and 24hr at each day.

2.1.2. Observation and Reanalysis

We examine the model SST forecasts with the daily mean SST (0.25° × 0.25°) from ECMWF ORAS5. To compare with the forecast results, the ORAS5 SST is remapped to the 1° × 1° grid. For the OLR, we choose the daily interpolated OLR (2.5° × 2.5°) from the National Oceanic and Atmospheric Administration (NOAA; Liebmann & Smith, 1996). Since there is no direct OLR comparison on each grid point, in this study we do not remap the OLR from the forecast output with 1° × 1° resolution to the 2.5° × 2.5° grid in observation. For the MSE budget terms, we analyze the variables from ECMWF reanalysis version 5 (ERA5). Although we acknowledge that reanalysis is constrained by observations but is different from real observations from the field, for simplicity, in the following sections, we use the word "observation" referring to both the observation and the reanalysis data.

2.2. Methods of MJO_P Identification

The MC barrier effect is exaggerated in climate models, where fewer MJO events cross the MC than in reality (H. Kim et al., 2018). Therefore, in this study, we focus on the MJO events that continuously propagate from the Indian Ocean across the MC to the Western Pacific, denoted by MJO_P events, and use the number of successfully forecasted MJO_P events as a metric to evaluate the performance of *all_obs* and *no_insitu*. The MJO_P events are identified by the RMMI and a new MJO tracking method based on OLR anomaly (OLRa) respectively.

2.2.1. RMMI-Based MJO_P Identification

The observed RMMI is tailored into the same 32-day segments as those in the forecasts. With the 32-day RMMI, an MJO_P event is defined as one that experiences the propagation from Phase 2 or 3 (over the Indian Ocean), through Phase 4 or 5 (over the MC) and then into Phase 6 or 7 (over the Western Pacific) within the 32 days, and with an RMMI magnitude always greater than 1 during this process.

The RMMI intrinsically has its weight projected more on the circulation field, because there are 2 wind fields for circulation while only 1 for convection involved in the RMMI calculation, with all the fields equally weighted in the combined EOF analysis. However, the atmospheric heating generated by the MJO deep convection plays a key role in teleconnection (e.g., Henderson et al., 2017). Thus, in order to focus on the convective signals, we develop a new method of MJO tracking using the OLRa. Since the new OLRa-based method does not consider the large-scale circulation while the RMMI-based method weighs more on circulation than on OLRa, it is expected that this new method might pick up some events that are not the canonical MJO events as defined by RMMI.

2.2.2. OLRa-Based MJO_P Identification

As for the observed RMMI, the observed OLR is also tailored into the 32-day segments, with the anomalies calculated by removing the 23-year climatological daily mean from the daily OLR. The OLRa in forecasts is calculated by removing the 23-year climatology as a function of both the forecast initialization month and the forecast lead time. Similar to Kiladis et al. (2005) and Gottschalck et al. (2013), a 2D Fast Fourier Transform is then applied to band-pass filter the OLRa for its eastward propagation with zonal wavenumber 1–9. After the temporal and zonal filtering, a 3-day running mean is further applied to the 32-day segment to filter out synoptic variability, making the 32-day segment into a 30-day segment. In this way, the MJO-related convection signals are obtained.

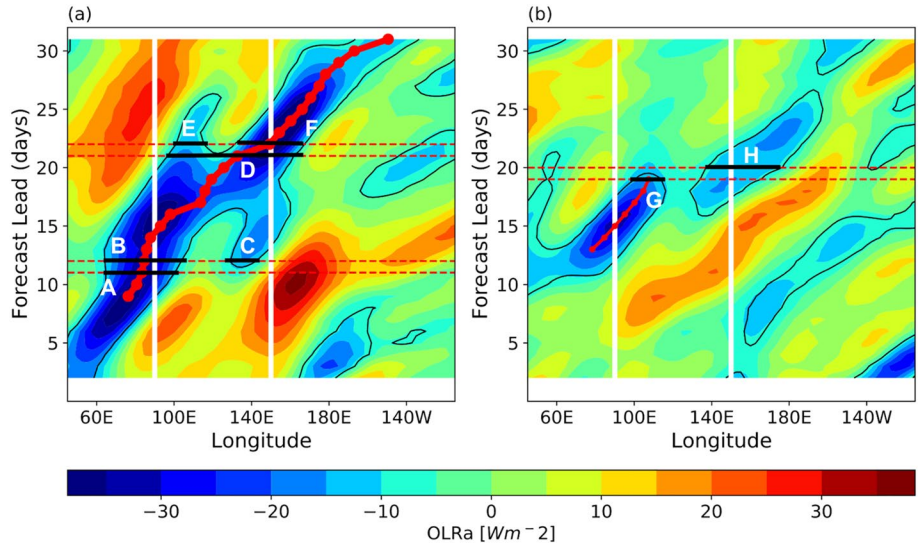

Figure 1. Schematic representation of the identification of the MJO_P events. The color shading shows filtered OLRa averaged over 15°N to 15°S. Negative anomalies exceeding 1 STD are defined as active MJO convection, and are denoted by the black contour. The two white vertical lines mark the western (90°E) and eastern (150°E) edges of the MC region. The red lines are the tracks for the MJO convection center. The horizontal black lines denote the longitudinal range of the active convection bands. (a) The tailored 32-day segment in observation (so that it shares the same time coordinate as in the forecast) starting from 05 January1997, as an example of an MJO_P event. (b) The same as (a) but starting from 01 January1998, which is not an MJO_P event. See text in Section 2.2.2 for details of the identification algorithm and meaning of bands A to H.

The filtered OLRa is then averaged from 15°N to 15°S with latitudinal weight adjustment, after which the standard deviation (STD) of the averaged OLRa can be calculated. We use −1 STD (the black contour in Figure 1) as the threshold of enhanced convection, and detect the MJO track based on those active MJO convection signals with negative anomalies exceeding 1 STD.

We describe the detailed steps in identifying the MJO tracks and selecting the MJO_P events as follows.

1. 90°E is chosen as the reference longitude, representing the western edge of the MC. It helps to capture the MJO convection from the Indian Ocean into the MC region. Based on this reference longitude, we can easily get the time index when the active MJO signal first approaches the western boundary of the MC (e.g., the starting day of the thick red line in Figure 1a and the starting day of the thin red line in Figure 1b), and also the corresponding location of the convection center (the mean longitude weighted by the OLRa magnitude at each longitude with active MJO convection, e.g., the starting longitude of the thick red line in Figure 1a and the starting longitude of the thin red line in Figure 1b).
2. From the starting day onward, the longitudinal convection band of the next day is identified. If there is an active convection band at day $i + 1$, and this band overlaps with the band at day i in longitude (e.g., band B overlaps with band A in Figure 1a), then the time index and its corresponding longitude of the convection center are recorded, and the tracking algorithm continues to propagate forward. If there is no active convection detected at day $i + 1$, or the detected band does not overlap with the band at day i (e.g., band H does not overlap with band G in Figure 1b), then the tracking algorithm for this MJO track would stop. If two or more bands are detected at day $i + 1$ for the same MJO track, we select the one overlapping with the band at day i (e.g., band C is not selected because it does not overlap with band A in Figure 1a). If still two or more bands are left, the distance between the convection center of day i and that of day $i + 1$ is calculated for each band at day $i + 1$, as well as the convection strength (the sum of the OLRa within each active band). Only the band that can minimize the value of "distance over strength" is selected into the current MJO track (e.g., band E and F both overlap with band D in Figure 1a, but only the band F is included in the propagating MJO track).
3. After identifying these MJO tracks in each 30-day segment, those whose convection center passes 150°E (the eastern edge of the MC) prior to the end of the 30 days are categorized as MJO_P events (e.g., the MJO track in Figure 1a).

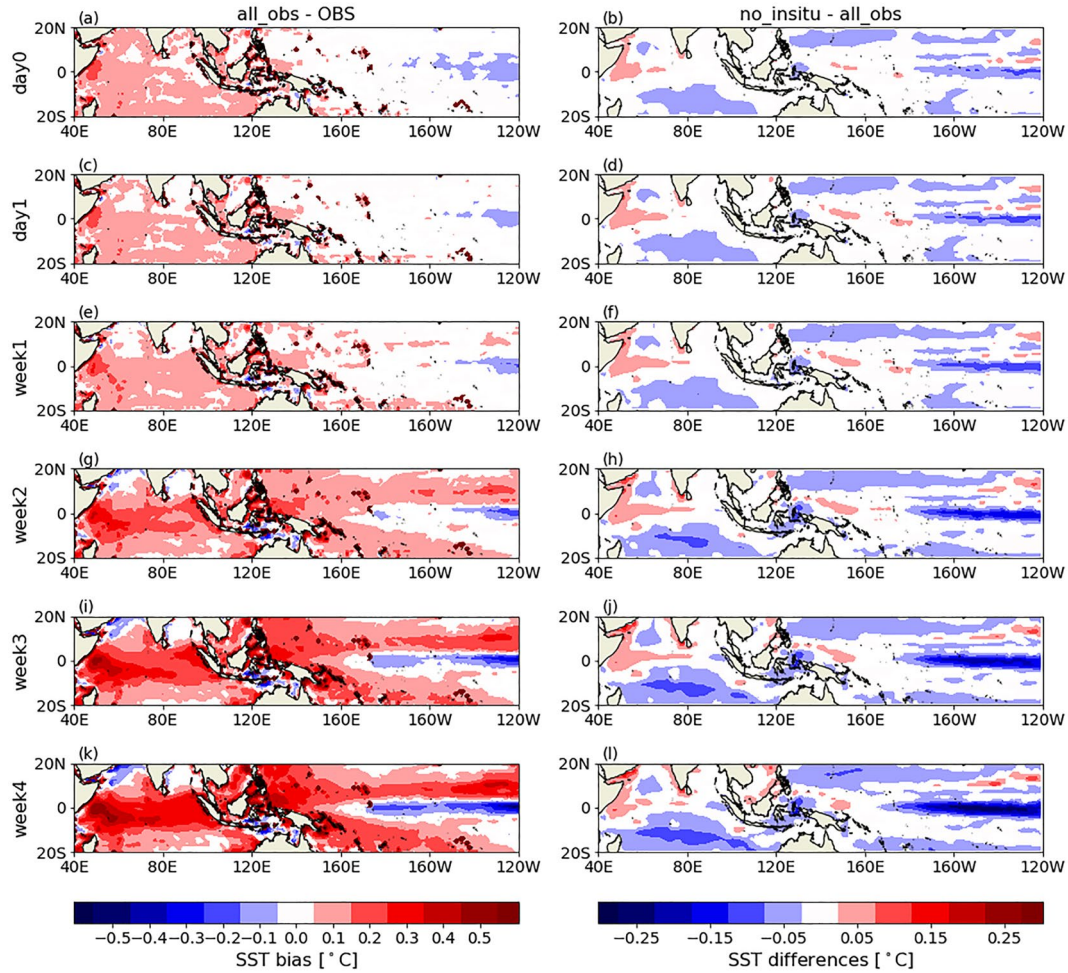

Figure 2. (Left) SST bias composite for *all_obs*. (a and c) show the daily mean SST biases in *all_obs* on the forecast initialization day and the forecast lead day 1. (e, g, i, and k) show the weekly mean SST biases in *all_obs* on the first, second, third, and fourth forecast lead week. (Right) Same as the left one but for SST differences between *all_obs* and *no_insitu*. The two-sided *t*-test has been applied and only values exceeding the 95% significance level are plotted. Notice that the colorbars for the left column and for the right column are different.

With the tracking method described above, we obtain the MJO_P events with their specific tracks characterized by the daily indices and the corresponding longitudinal locations of the convection center along with the MJO propagation.

3. Results

3.1. SST Mean State

Before starting the detailed MJO_P evaluation, we first examine the SST mean state. It would increase the chances for ocean subsurface DA to have an impact on the MJO forecast if some pronounced SST mean state differences between *all_obs* and *no_insitu* could be seen.

In general, *all_obs* and *no_insitu* have similar ocean mean state bias pattern and same bias growth sign (Figure S1 in Supporting Information S1). On the initialization day, the model has warm SST biases in the Indian Ocean and cold SST biases in the cold tongue region (Figure 2a). Since we average the SST at 0 and 12hr to represent the forecasted daily mean SST on day 0, it is not surprising to see non-zero biases in forecasts on the initialization day because the forecast error at lead time 12hr has already been introduced. Then the biases grow and expand gradually (Figures 2c, 2e and 2g, 2i and 2k). Significant warm SST biases develop in the Indian Ocean and the warm pool, and significant cold SST biases develop in the cold tongue region and the south MC seas.

Table 1
RMMI-Based Evaluation

	Number of MJO_P[a]	Number of MJO_P in OBS[b]	Number of MJO_IO[c]	Number of MJO_IO_P[c]	Passing rate
OBS	39	–	44	21	0.48
all_obs	16 [13,22,25,21,19][d]	12 [8,14,16,14,10]	45 [45,45,45,44,45]	11 [7,12,10,12,9]	0.24 [0.22]
no_insitu	17 [15,25,25,22,14]	13 [10,14,17,11,8]	45 [45,45,45,44,45]	12 [10,14,12,11,8]	0.27 [0.25]

[a]Please see definition of MJO_P events in Section 2.2. [b]MJO_P in OBS is defined as an event that happens in the observation and is also successfully predicted by the model forecasts. [c]Please see definition of MJO_IO events and MJO_IO_P events in Section 3.2.1. [d]Numbers in the bracket are for the five individual ensemble members.

Since SST nudging is on for both OSEs, most SST differences between *all_obs* and *no_insitu* are within 0.075°C on the initialization day (Figure 2b). The maximum difference happens in the cold tongue region, where the SST in *no_insitu* is colder than that in the *all_obs* (Figure 2b). The SST difference grows as the forecast goes on, becoming greater than 0.2°C in week 4 of the forecast (Figure 2l). Considering that the model has large cold biases in the cold tongue region, *all_obs* actually helps to reduce the cold biases in the cold tongue region.

Over other regions, though *all_obs* and *no_insitu* start from small SST differences, these differences also gradually develop with a growing amplitude and an expanding area (Figures 2d, 2f, 2h, 2j and 2l). Over the MC seas, the South Indian Ocean and the north part of the tropical Western Pacific, *all_obs* is warmer than *no_insitu* and becomes even warmer as the forecast goes on. Though most of those differences are still within 0.125°C, we speculate that there might be some cases where those SST differences in the individual MJO events could influence whether or not the MJO crosses the MC.

3.2. MJO_P Events

In this section, we examine whether the regional and large-scale SST differences between experiments *all_obs* and *no_insitu* affect the prediction skill in forecasting MJO across the MC.

3.2.1. Evaluation Using RMMI

Based on RMMI, we identify 39 MJO_P events in the observation, 16 MJO_P events in *all_obs* ensemble mean, and 17 in *no_insitu* ensemble mean (Table 1). Among the 39 observed MJO_P events, 12 (13) are captured by the ensemble mean forecast of *all_obs* (*no_insitu*). Both *all_obs* and *no_insitu* predict some spurious MJO_P events that are not in the observation. The two forecast experiments yield similar results in successfully forecasting the observed MJO_P events, except for the January 1995 event, when the *all_obs* predicts the MJO to continue staying at Phase 5 till the end, whereas the *no_insitu* predicts the MJO to transition to Phase 6 but shortly after return to Phase 5. However, when looking at the individual ensemble members for this event, we find that actually the MJO_P track is predicted by 3 ensemble members in *all_obs* and by 2 ensemble members in *no_insitu*. Hence, the two experiments may not have significantly different prediction skills for the 1995 event.

We also investigate the MJO events that are active over the Indian Ocean (Phase 2 or 3) on the forecast initialization day and the corresponding day in the observation (the MJO_IO events). During the 23-year observational period (1993–2015), 44 events are active over the Indian Ocean and 21 of them propagate into the Western Pacific within 32 days (the MJO_IO_P events), yielding a passing rate of 0.48 (Table 1). By contrast, in *all_obs* (*no_insitu*), only 11 (12) out of the 45 (45) events pass the MC region, with a passing rate of 0.24 (0.27). These results are consistent with existing studies in which the model exaggerates the MC barrier effect. And our results show no improvement in *all_obs* compared to *no_insitu*. Such deficiency is likely a prediction skill problem instead of a predictability problem, since it is common for all ensemble members (see bracket numbers in Table 1).

Overall, the MC prediction barrier exists in both experiments. *All_obs* and *no_insitu* have the same prediction skill for forecasting MJO propagation across the MC when evaluated with RMMI. Differences in ensemble mean analysis versus individual ensemble member analysis will not change our conclusion.

3.2.2. Evaluation Using OLRa

Based on the filtered OLRa, there are 79 MJO_P events detected in observation. Among them 35 (37) are detected in the ensemble mean OLRa field of *all_obs* (*no_insitu*). In addition to these, *all_obs* (*no_insitu*) forecasts 35 (31)

Table 2
OLRa-Based Evaluation

	Number of MJO_P	Number of MJO_P in OBS	Number of MJO_IO	Number of MJO_IO_P	Passing rate
OBS	79	–	44	22	0.50
all_obs	70 [70,65,65,69,68]	35 [30,25,33,27,29]	45 [45,45,45,44,45]	12 [15,7,11,10,14]	0.27 [0.25]
no_insitu	68 [76,69,58,75,62]	37 [35,26,21,32,32]	45 [45,45,45,44,45]	13 [13,10,8,16,13]	0.29 [0.27]

Note. The same as Table 1 but for OLRa-based evaluation.

extra MJO_P events which are not seen in the observation (Table 2). As expected, some MJO_P events are identified by both the RMMI-based method and the OLRa-based method, but each method also identifies independent events. Additionally, compared to the number of MJO_P events identified using RMMI, we see more MJO_P events using OLRa. In general, those extra MJO_P events also show evidence of propagation over the Maritime Continent in their RMMI. However, some of them can not maintain their RMMI amplitude to be greater than 1 in the process, some of them start from Phase 4 or 5 instead of 2 or 3, and some of them are occasionally interrupted by Phase 1 or Phase 8 during propagation. Overall, the OLRa-based method has looser restrictions on identifying MJO_P events than the RMMI-based method.

Among the 44 events that are initially active in Phase 2 or 3 (over the Indian Ocean) in the observation, 22 of them propagate into the Western Pacific when tracked using the OLRa, with a passing rate of 0.5. For model forecasts, 12 (13) out of 45 (45) reach the Western Pacific within the forecast time period in *all_obs* (*no_insitu*), with a passing rate of 0.27 (0.29). Similar to what we have found using RMMI, both experiments have a much lower passing rate compared to the observation, and so does each individual ensemble member (see bracket numbers in Table 2). Additionally, for the MJO_P events identified in the ensemble mean of one experiment but not in the other, we further analyze their individual ensemble members. There is no such a single event that all ensemble members in one experiment forecast the MJO_P but all ensemble members in the other do not.

Therefore, we conclude that when evaluated with the OLRa, the *all_obs* and *no_insitu* also have similar prediction skill, and the initialization with ocean in situ DA does not help improve the prediction skill of MJO propagation across the MC. We further perform MJO process diagnostics to identify additional sources of errors in the MJO evolution in these forecasts and present the results below.

3.3. Process Diagnostics

Existing theories suggest that MJO events are generated primarily from atmospheric internal variability (e.g., Adames & Kim, 2016; Sobel & Maloney, 2013); air-sea interaction can modify their amplitude, frequency, and propagation. Hence, if atmospheric model biases dominate the forecast biases, the potential improvement from ocean DA would be hindered. This is probably why even though there are growing regional and large-scale SST mean state differences between *all_obs* and *no_insitu*, the forecast skill of MJO across the MC still remains the same in the two experiments. To understand the causes for the forecast biases in both experiments, we carry out process diagnostics by doing MSE budget analysis in the MC region.

MSE budget analysis has been employed in MJO studies based on the "Moisture Mode" theory of MJO (e.g., Kiranmayi & Maloney, 2011; Seo et al., 2014; D. Kim et al., 2014, 2017; Y.-K. Lim et al., 2021), based on the hypothesis that the growth, decay and the eastward propagation of the MJO convection is governed by the intraseasonal moisture anomalies. The MSE is defined as $m = C_pT + gZ + L_vq$, where C_p is the specific heat of dry air at constant pressure, T is the air temperature, g is the gravitational acceleration, Z is the geopotential height, L_v is the latent heat of vapourization, and q is the specific humidity. The MSE budget equation can be written as

$$\left\{\frac{\partial m}{\partial t}\right\}' = -\left\{u\frac{\partial m}{\partial x}\right\}' - \left\{v\frac{\partial m}{\partial y}\right\}' - \left\{\omega\frac{\partial m}{\partial p}\right\}' + LH' + SH' + \langle LW\rangle' + \langle SW\rangle' \quad (1)$$

where u is the zonal wind velocity, v is the meridional wind velocity, ω is the pressure velocity, p is the pressure, LH is the surface latent heat flux, SH is the surface sensible heat flux, LW is the longwave radiation, and SW is the shortwave radiation. The angle bracket $\langle\rangle$ stands for the net radiation into the atmosphere (from both the ocean

surface and top of the atmosphere). The curly bracket { } denotes the mass-weighted vertical integral over the pressure levels from 100 hPa to 1,000 hPa. The prime symbol ′ denotes the intraseasonal anomaly. In this study, we use the 10-day running mean after removing the daily climatology to represent the intraseasonal anomalies. We apply a moving average instead of a band-pass filter due to the limitation caused by our 32-day forecast segments of a short period.

3.3.1. Pre-Convection MSE Accumulation

As discussed in Section 3.2.2, less than half of the OLRa-based MJO_P events in observation can be captured by the model forecast for both *no_insitu* and *all_obs*. Yet, the model does predict a similar amount of the total MJO_P events as what we get in observation (~70), though half of them are the "false" events that do not appear in the observation. This suggests that the model might be producing the MJO_P events because of wrong physical mechanisms, that is, the model might overestimate the contribution of some terms to the positive intraseasonal MSE tendency anomaly prior to the MJO convection onset, and underestimate the contribution of some others. We are especially interested in the terms whose positive/negative contribution to the MSE accumulation before the MJO convection onset is underestimated/overestimated by the model, because they can give us a hint of why the model is exaggerating the MC barrier effect.

To identify such terms, we apply the MSE budget analysis on MJO_P events detected in the observation, *all_obs* and *no_insitu* respectively. We choose the MJO_P events in the model forecasts instead of the missed MJO_P events in the model (the events that exist in the observation but the model fails to forecast) as, by choosing the MJO_P events produced by the model, we have the locations and the time indices for each MJO track, and thus can analyze the 10-day window (in order to extract the intraseasonal MJO signals by excluding synoptic variability) to best represent the pre-MJO MSE accumulation for a certain region. Thus, the MSE accumulation windows in the model and in the observation no longer have to be the same with respect to the forecast lead time, which helps us understand the model physics in MSE accumulation better. Additionally, we assume that the model has systematic biases in response to the MJO activity. Hence, the biases we find in the forecasted MJO_P events would also exist in the missed MJO_P events, and can potentially explain why the model fails to forecast those events. In order to understand the model mechanism better, we analyze the individual ensemble members instead of the ensemble mean.

The region over which we choose to apply the MSE budget analysis is the eastern part of the MC (130°E−150°E, 15°N–15°S). It connects the MC and the Western Pacific, thus is the key region for the MJO to propagate into the Western Pacific. All MSE budget terms in the following analyses are averaged over this box region.

To select the 10-day window for representing the MSE accumulation before the MJO convection onset in our region of interest, we search for the day when the MJO_P track for the first time arrives to the east of 130°E, and denote it as the day_130 for that MJO_P event. Note that in the OLRa-based MJO_P tracking method, we reduce the 32-day observation/forecast segment into a 30-day segment by performing a 3-day running mean, so the day i in the 30-day segment would be the day $i + 1$ in the original 32-day segment. After mapping the day_130 in the 30-day segment back into the 32-day segment, we transition the original 32-day observation/forecast to a new time axis with the day_130 now becoming the new day 0. Then, we compute the anomalous MSE tendency averaged within a 10-day window as a function of the days lagging day 0. If day i is beyond the 32-day segments, then the intraseasonal MSE tendency anomaly would be a null value; if only a part of the days from the 10-day window (day i-9 to day i) is within the 32-day segment, then we calculate the intraseasonal MSE tendency anomaly based only on that part.

In this way, we obtain the intraseasonal MSE tendency anomalies as a function of the lag for the day_130 of each MJO_P event. These MSE tendencies are then composited along the new time axis for the observation, *all_obs* and *no_insitu* respectively. From Figure 3a, we can clearly see that before the MJO convection arrives at 130°E, the eastern MC has a positive MSE accumulation, and after that, the eastern MC has a negative MSE tendency due to the moisture discharge caused by the precipitation. Such tendencies are seen in both the model and the observation, though the amplitude of both MSE accumulation and discharge is smaller in the model. Additionally, since *all_obs* and *no_insitu* closely follow each other in Figure 3a, we conclude that *all_obs* and *no_insitu* perform similarly regarding the MSE recharge and discharge. Lag −6 (the dashed vertical line in Figure 3a) is selected as the MSE accumulation window since it is the maximum point for the observation and also a local maximum for both *all_obs* and *no_insitu*, and therefore can be considered as a common MSE accumulation

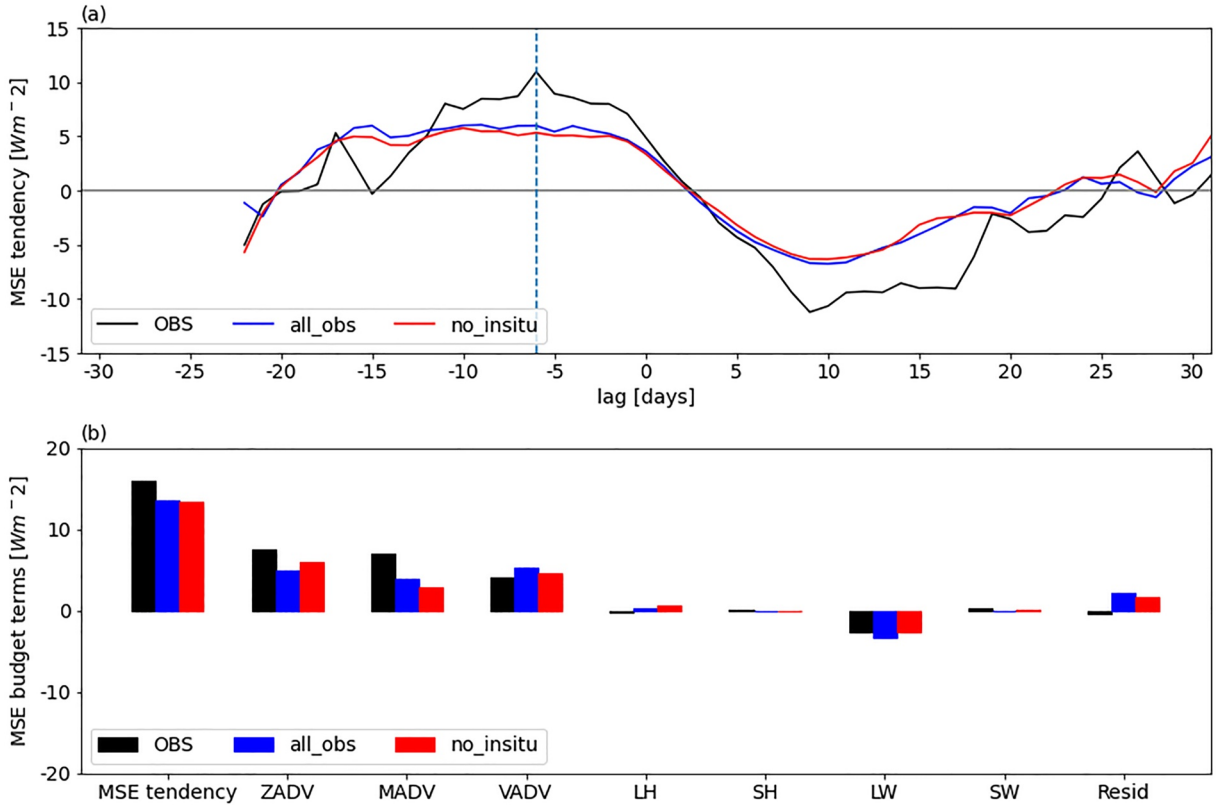

Figure 3. (a) The 10-day averaged MSE tendency for the MJO_P events in observation (black), *all_obs* (blue) and *no_insitu* (red) composited based on the lag to the day when the MJO_P track for the first time arrives to the east of 130°E (day_130). (b) MSE budget composite for the MJO_P events in observation (black), *all_obs* (blue) and *no_insitu* (red) that have a positive MSE accumulation in the lag −6 window. ZADV, MADV and VADV stand for zonal, meridional and vertical MSE advection respectively. All MSE budget terms are averaged over the eastern MC (130°E−150°E, 15°N–15°S).

period for both the observation and the model forecast. (Lag −5 and lag −2 have also been analyzed. The results we show next are not sensitive to the lag we choose.)

Then, the intraseasonal MSE tendency within the lag −6 window (from lag −15 to lag −6) is calculated for each MJO_P event. The MSE tendencies can be a null value (window out of the 32-day segment), a positive value or a negative one. This is expected because the MJO varies from event to event, and a single lag window selected based on the composite might not explain the pre-convection MSE accumulation for every individual event. Since we concentrate on the physical mechanisms related to the positive MSE accumulation, only those with a positive tendency are included in the composite analysis for the MSE budget. In the end, we obtain 50 observed MJO_P events, 202 *all_obs* MJO_P events and 193 *no_insitu* MJO_P events for composite MSE budget analysis. To avoid sampling bias, we examine the seasonality for these events. The selected MJO_P events in the observation, *all_obs* and *no_insitu* all approximately spread equally in all seasons (Figure S2 in Supporting Information S1). This further assures that the results in the later analyses reflect the process-induced biases or differences.

Overall, *all_obs* and *no_insitu* have similar amplitudes of composite MSE budget terms, which from a perspective of physical processes supports our conclusion in Section 3.2 that these two sets of experiments have similar prediction skills on MJO propagation across the MC. Noticeably, both *all_obs* and *no_insitu* underestimate the intraseasonal MSE tendency by 2–3 Wm^{-2}, and the most significant underestimation for the MSE budget terms happens in the meridional advection (Figure 3b). The positive meridional advection in the forecast composite is only ∼50% of that in the observational composite (and this is the same with lag −5 and lag −2 as well, see Figure S3 in Supporting Information S1). Besides the composite analysis, multiple linear regression is also applied to estimate the contribution to the MSE accumulation (the predictand) from each MSE budget term (the predictors) on the right hand side of Equation 1, with each MJO_P event corresponding to one sample. The regression coefficient for the meridional advection is 0.67 in observation, 0.35 in *all_obs* and 0.37 in *no_insitu*, which further

indicates that the model significantly underestimates the meridional advection (see Table S1 in Supporting Information S1 for the regression coefficients of other terms). Such underestimated meridional advection gives us a clue that the model might be exaggerating the MC barrier effect out of the insufficient meridional advection.

3.3.2. Possible Sources for the Underestimated Meridional Advection

Based on the moisture mode theory of the MJO, moisture is the core atmospheric variable determining the MJO variability. Therefore, we focus on the intraseasonal meridional moisture advection, and further decompose it into 9 terms by decomposing the meridional wind and the moisture content into the climatological component, the intraseasonal component and the high-frequency component respectively (see Text S1 in Supporting Information S1 for more details). Our results suggest that both *all_obs* and *no_insitu* mainly underestimate the following terms: the intraseasonal meridional wind advecting the climatological moisture, the intraseasonal meridional wind advecting the intraseasonal moisture, and the high-frequency meridional wind advecting the high-frequency moisture (Figure S4a in Supporting Information S1). These three terms can explain ∼70% of the total underestimation of the meridional MSE advection. Among these 3 terms, the intraseasonal meridional wind advecting the climatological moisture term can explain an underestimation of ∼1 Wm^{-2} in model forecasts, and is also the term that has the highest correlation coefficient (∼0.7) with the total intraseasonal meridional moisture advection (Figure S4b in Supporting Information S1). Our results are consistent with the existing studies, which have shown that the intraseasonal anomalous wind advecting the mean state moisture is a dominant process in the horizontal moisture advection (e.g., Adames, 2017; DeMott et al., 2018; Jiang et al., 2018; Kang et al., 2021; Maloney, 2009). Based on our results and previous studies, we single out the intraseasonal meridional wind advecting the climatological moisture term for further analysis.

When the intraseasonal meridional winds blow from the wet region to the dry region, they help moisten the tropical atmosphere. Previous studies have suggested that the forecast models tend to underestimate the MJO due to a systematic dry bias in the seasonal mean state moisture (Gonzalez & Jiang, 2017; H.-M. Kim, 2017; Y. Lim et al., 2018). Therefore, we are motivated to investigate the mean state moisture bias in *all_obs* and *no_insitu*.

Figure 4 shows that both *all_obs* and *no_insitu* have significant dry biases over the warm pool at 700 hPa. Such dry biases are also seen for the Integrated Water Vapor (Figure S5 in Supporting Information S1). To further understand how these dry biases in the low-frequency moisture field contribute to the underestimated meridional advection in the model forecast, we decompose the term of intraseasonal anomalous wind advecting the climatological moisture gradient into sub-terms step by step.

As aforementioned, we calculate the daily anomalies in each MSE budget term by removing the 23-year daily climatology, and then averaging them over a 10-day window to get the intraseasonal signals. Before averaging over the 10-day window, we have the anomalous daily meridional advection, which includes a term $-v'\frac{\partial Q}{\partial y}$ on each pressure level, where v' is the daily meridional wind velocity anomaly, and Q is the climatological specific humidity of that day.

By assuming the climatological specific humidity does not change over the 10-day window, we have

$$\overline{-v'\frac{\partial Q}{\partial y}} = -\overline{v'}\frac{\overline{\partial Q}}{\partial y} \tag{2}$$

where the bar stands for the time average over the 10-day window. Since the climatological Q does not change much within 10 days, we treat Q as a constant $\overline{Q}$. Hence, the intraseasonal meridional advection anomaly provided by the daily meridional wind anomaly advecting the climatological moisture gradient can be rewritten into the 10-day averaged intraseasonal meridional wind anomaly $\left(\overline{v'}\right)$ advecting the 10-day averaged climatological moisture gradient $\left(\frac{\overline{\partial Q}}{\partial y}\right)$. For simplicity, we drop the bar in the following derivation.

For each aforementioned MJO_P event with a positive pre-convection MSE tendency in observation, *all_obs* or *no_insitu*, we now have the corresponding v' and Q for its MSE accumulation window, and these two variables can both be written as the composite mean plus the deviation from the composite mean. Then, we are able to decompose the $-v'\frac{\partial Q}{\partial y}$ into

$$-v'\frac{\partial Q}{\partial y} = -\left(\widehat{v'} + v'_d\right)\frac{\partial\left(\widehat{Q} + Q_d\right)}{\partial y} = -\widehat{v'}\frac{\partial \widehat{Q}}{\partial y} - \widehat{v'}\frac{\partial Q_d}{\partial y} - v'_d\frac{\partial \widehat{Q}}{\partial y} - v'_d\frac{\partial Q_d}{\partial y} \tag{3}$$

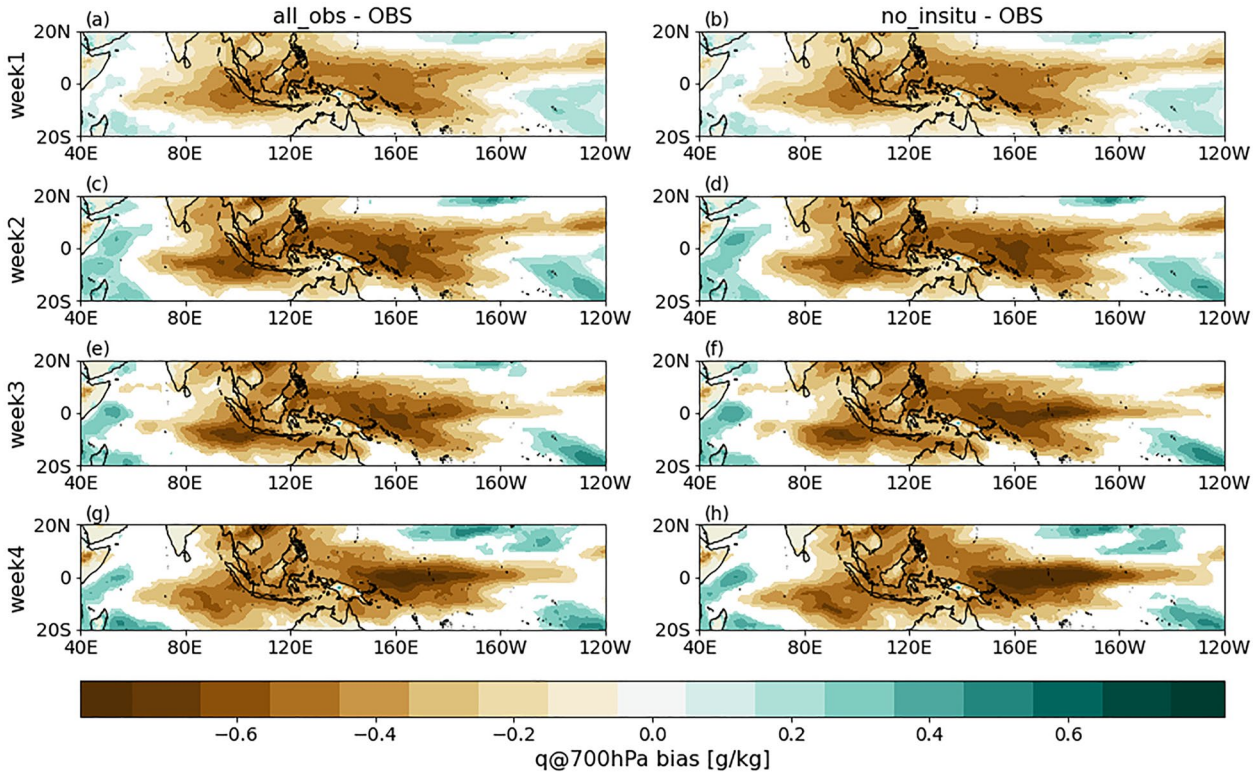

Figure 4. (a, c, e, and g) show the development of the weekly mean specific humidity bias at 700 hPa in *all_obs* from week1 to week 4. The right column is the same as the left but for *no_insitu*. The two-sided *t*-test has been applied and only values exceeding the 95% significance level are plotted.

where the hat denotes the composite mean, and the subscript d means the difference between the individual event and the composite mean. Then we write Equation 3 into its composite form

$$-v'\frac{\partial \hat{Q}}{\partial y} = -\hat{v}'\frac{\partial \hat{Q}}{\partial y} - \hat{v}'\frac{\partial \hat{Q}_d}{\partial y} - v'_d\frac{\partial \hat{Q}}{\partial y} - v'_d\frac{\partial \hat{Q}_d}{\partial y} = -\hat{v}'\frac{\partial \hat{Q}}{\partial y} - v'_d\frac{\partial \hat{Q}_d}{\partial y} \tag{4}$$

Only two terms are left on the right hand side. $-\hat{v}'\frac{\partial \hat{Q}}{\partial y}$ is the composite mean intraseasonal meridional wind anomaly $\left(\hat{v}'\right)$ advecting the composite mean climatological moisture gradient $\left(\frac{\partial \hat{Q}}{\partial y}\right)$. $-v'_d\frac{\partial \hat{Q}_d}{\partial y}$ is the composite of the wind deviation advecting the moisture gradient deviation. Since the composite variables $\hat{v}'$ and $\hat{Q}$ highlight the general moisture and circulation pattern during the MSE accumulation stage of the MJO_P events, we focus on the $-\hat{v}'\frac{\partial \hat{Q}}{\partial y}$ term only in the present discussion.

To compare the meridional advection in the model and the meridional advection in the observation, we drop the hat in the composite term $-\hat{v}'\frac{\partial \hat{Q}}{\partial y}$ for simplicity and rewrite it into

$$-v'_i\frac{\partial Q_i}{\partial y_i} = -\left(v'_{obs} + v'^{*}_i\right)\frac{\partial\left(Q_{obs} + Q^{*}_i\right)}{\partial y} = -v'_{obs}\frac{\partial Q_{obs}}{\partial y} - v'_{obs}\frac{\partial Q^{*}_i}{\partial y} - v'^{*}_i\frac{\partial Q_{obs}}{\partial y} - v'^{*}_i\frac{\partial Q^{*}_i}{\partial y} \tag{5}$$

where i represents either *all_obs* or *no_insitu*, v'^{*}_i is the model bias of the composite mean intraseasonal meridional wind anomaly based on the observation, and Q^{*}_i is the model bias of the composite mean climatological moisture.

Equation 5 shows that the difference between $-v'_i\frac{\partial Q_i}{\partial y_i}$ and $-v'_{obs}\frac{\partial Q_{obs}}{\partial y}$ are rooted in the following three terms:

$$T^{i}_{1} = -v'_{obs}\frac{\partial Q^{*}_i}{\partial y} \tag{6}$$

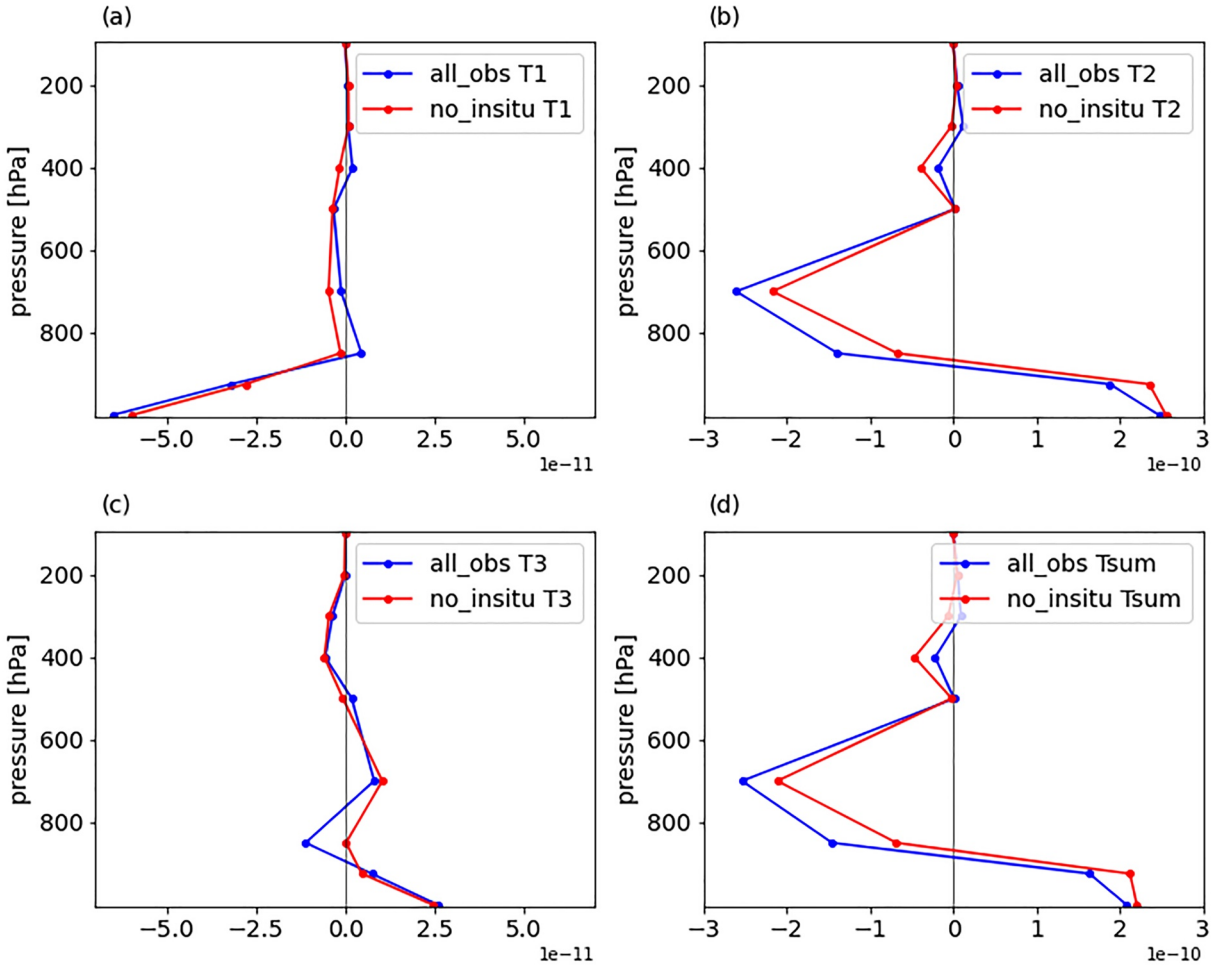

Figure 5. (a) Shows the vertical profile of T_1 in *all_obs* (blue) and in *no_insitu* (red). (b–d) are the same as (a) but for T_2, T_3, and T_{sum} respectively. See the definition of these terms in Section 3.3.2. The unit of the x axis is $kgkg^{-1}s^{-1}$. The terms are calculated using the lag −6 window. Note that the scale for (a) and (c) is one order of magnitude smaller than that of (b and d). These terms are calculated over the eastern MC (130°E−150°E, 15°N–15°S).

$$T_2^i = -v_i'^* \frac{\partial Q_{obs}}{\partial y} \tag{7}$$

$$T_3^i = -v_i'^* \frac{\partial Q_i^*}{\partial y} \tag{8}$$

T_1 is the observed meridional wind anomaly advecting the climatological moisture gradient bias, T_2 is the meridional wind anomaly bias advecting the observed climatological moisture gradient, and T_3 is the quadratic term where the wind bias is advecting the moisture gradient bias.

We calculate T_1, T_2, T_3 and their sum T_{sum} for both *all_obs* and *no_insitu* at each pressure level. The vertical profile of each term is shown in Figure 5. In both *all_obs* and *no_insitu*, T_2 is one order of magnitude greater than T_1 and T_3, and it dominates the vertical structure of T_{sum}. After applied the mass-weighted vertical integral, and multiplied by the constant L_v, T_2 leads to a meridional advection underestimation of 1 Wm^{-2} in *all_obs* and of 0.65 Wm^{-2} in *no_insitu*, and T_2 accounts for 91% and 85% of the integrated T_{sum} in *all_obs* and *no_insitu* respectively.

Therefore, it is the intraseasonal meridional wind bias instead of the dry bias in the model that indeed contributes the most to the difference between $v_i' \frac{\partial Q_i}{\partial y_i}$ and $-v_{obs}' \frac{\partial Q_{obs}}{\partial y}$, and is thus responsible for a significant part of the meridional advection underestimation in the model forecast. The same analysis on a lag −2 MSE accumulation window

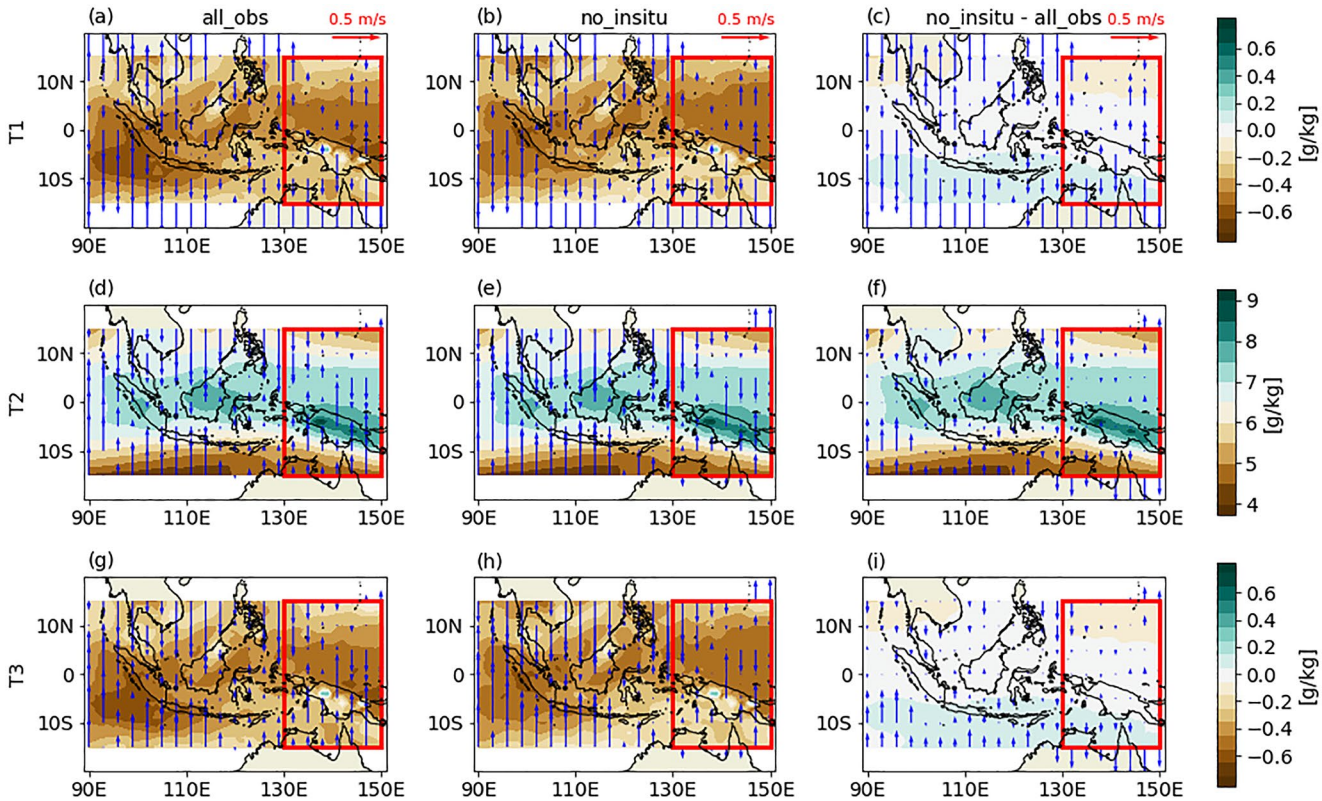

Figure 6. A visualization of the components of the three T terms at 700 hPa. The background colors represent the moisture component, and the overlapped arrows represent the meridional wind component. The first and the second columns visualize T_1, T_2 and T_3 for *all_obs* and *no_insitu* respectively. The third column shows the differences between *all_obs* and *no_insitu* in each wind and moisture component for each T term. The red box denotes our region of interest (130°E−150°E, 15°N–15°S).

(from lag −11 to lag −2) has also been done, and the conclusions remain the same (Figure S6 in Supporting Information S1).

To further illustrate why the wind bias contributes more than the dry bias, a visualization of each term on a Lon Lat map for the 700 hPa pressure level is provided in Figure 6. In general, over the eastern MC (the red box), the observed meridional winds diverge away from where the model maximum dry bias is and the model meridional wind biases converge to where the observed climatological maximum moisture is, which results in a negative T_1 and a negative T_2 respectively. For T_3, the wind biases converge to the maximum dry bias, thus causing a positive meridional advection. This is consistent with what we have seen in Figures 5a–5c. However, the gradient of the dry bias is much smaller than the gradient of the observed climatological moisture, while the wind bias is of the same magnitude as the observed meridional wind (e.g., Figures 6a and 6d). Hence, T_2 is much greater than T_1 and T_3. Though other studies consider the improvement in the moisture field as key to improving the MJO prediction furthermore (e.g., Gonzalez & Jiang, 2017; Jiang, 2017; H.-M. Kim, 2017; Y. Lim et al., 2018), this finding suggests that in the ECMWF subseasonal forecast system, the circulation field is still a large source of the MJO prediction error, and should be a target of future improvement.

4. Summary and Discussion

Two ocean OSEs have been carried out by ECMWF, one with all ocean in situ observations assimilated, and the other with all in situ observations removed from the data assimilation system. Two sets of subseasonal forecasts are initialized from those two ocean OSEs respectively (*all_obs* and *no_insitu*) and they are the same otherwise. Thus, we are able to assess the impact of the subsurface ocean initialization with ocean subsurface observation assimilation on predicting MJO propagation across the MC in these ECMWF subseasonal forecast experiments.

It should be noted that, due to the limited forecast sample size (12 initializations per year for 23 years) and the limited temporal output (12 hourly instantaneous values), the discussion of the seasonality and the diurnal cycle

are excluded in this study. We are aware that both seasonality and the diurnal cycle are important in understanding the MC barrier effect. However, they would not affect our conclusions in this study as discussed below.

First, we carry out composite analyses to compare the SST mean state in *all_obs* and *no_insitu*. During the forecast, the SST differences between two experiments grow. Compared to *no_insitu*, *all_obs* has warmer regional and large-scale SST. Therefore, we speculate if such SST differences can lead to an improved prediction for MJO propagation across the MC in *all_obs* compared to *no_insitu*.

However, based on the evaluation using RMMI and the newly developed OLRa tracking method, no improvement has been found in *all_obs*. Both *all_obs* and *no_insitu* exaggerate the MC barrier effect to a similarly large extent. And there is no single event where one experiment has better prediction skill compared to the other. The initial SST differences in the tropical oceans do not lead to a MJO forecast difference.

Why does the ocean subsurface DA have no impact on MJO propagation across the MC in ECMWF subseasonal forecast model as expected? We perform MJO process diagnostics to explore the possibility that the atmospheric biases of the model dominate the forecast error growth, which might hinder the potential role played by the ocean state differences. With MSE budget analysis, we confirm that *all_obs* and *no_insitu* perform similarly on predicting MJO propagation across the MC. We also find that compared to the observation, both *all_obs* and *no_insitu* significantly underestimate the meridional advection over the eastern MC during the pre-convection MSE accumulation period. Since the intraseasonal meridional wind advecting the climatological moisture is not only one of the major bias sources of the underestimated meridional advection in the model, but also highly correlated with the total meridional moisture advection in both the observation and the model forecasts, we further compare the term of the composite mean intraseasonal wind anomaly advecting the composite mean low-frequency moisture field between the model forecast and the observation. We find that both *all_obs* and *no_insitu* underestimate this term, which is overwhelmingly due to the biases in the intraseasonal meridional winds. Such an atmospheric origin of the forecast errors may help explain the forecast insensitivity to SST.

This study suggests that overcoming the MC prediction barrier in ECMWF subseasonal forecast would require reducing the MJO-related meridional wind bias. This is consistent with what H. Kim et al. (2019) has found when studying the SubX and S2S reforecasts. Their result shows that though with the best represented mean moisture state among all models, the ECMWF model still has fast damping of the horizontal advection, which is no better than many other models with larger dry biases. That said, improving the mean moisture field remains an important step to improving the MJO forecast in the ECMWF subseasonal forecast system. Bias development is a result of many processes entangled together, thus improving one field could also potentially lead to improvement in others.

Data Availability Statement

ORAS5 data are provided by ECMWF and downloaded from https://resources.marine.copernicus.eu. Interpolated OLR data are provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, from their web site at https://psl.noaa.gov/data/gridded/data.interp_OLR.html (Liebmann & Smith, 1996). ERA5 data are provided by ECMWF and downloaded from https://cds.climate.copernicus.eu/ (Hersbach et al., 2018a, 2018b). The subseasonal forecast data we used in this study has been published and is accessible at https://doi.org/10.25810/Z9EH-3940.

Acknowledgments
The authors sincerely appreciate the constructive comments and suggestions from three anonymous reviewers and the handling editor. This work is funded by the ONR MISOBOB research initiative (N00014-17-S-B001), NSF (NSF-OCE 1658132), NOAA (NA18OAR4310405), and NASA (21-OSST21-0026). The authors would also like to acknowledge high-performance computing support from Cheyenne (https://doi.org/10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. All analyses are done with the Casper data analysis and visualization cluster.

References

Adames, Á. F. (2017). Precipitation budget of the Madden–Julian Oscillation. *Journal of the Atmospheric Sciences*, *74*(6), 1799–1817. https://doi.org/10.1175/jas-d-16-0242.1

Adames, Á. F., & Kim, D. (2016). The MJO as a dispersive, convectively coupled moisture wave: Theory and observations. *Journal of the Atmospheric Sciences*, *73*(3), 913–941. https://doi.org/10.1175/jas-d-15-0170.1

Ajayamohan, R., Khouider, B., Praveen, V., & Majda, A. J. (2021). Role of diurnal cycle in the maritime continent barrier effect on MJO propagation in an AGCM. *Journal of the Atmospheric Sciences*, *78*(5), 1545–1565. https://doi.org/10.1175/jas-d-20-0112.1

Balmaseda, M. A. (2017). Data assimilation for initialization of seasonal forecasts. *Journal of Marine Research*, *75*(3), 331–359. https://doi.org/10.1357/002224017821836806

Balmaseda, M. A., AlVES, O. J., Arribas, A., Awaji, T., Behringer, D. W., Ferry, N., et al. (2009). Ocean initialization for seasonal forecasts. *Oceanography*, *22*(3), 154–159. https://doi.org/10.5670/oceanog.2009.73

Cassou, C. (2008). Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Nature*, *455*(7212), 523–527. https://doi.org/10.1038/nature07286

DeMott, C. A., Wolding, B. O., Maloney, E. D., & Randall, D. A. (2018). Atmospheric mechanisms for MJO decay over the maritime continent. *Journal of Geophysical Research: Atmospheres*, *123*(10), 5188–5204. https://doi.org/10.1029/2017jd026979

Gonzalez, A. O., & Jiang, X. (2017). Winter mean lower tropospheric moisture over the maritime continent as a climate model diagnostic metric for the propagation of the Madden–Julian Oscillation. *Geophysical Research Letters*, *44*(5), 2588–2596. https://doi.org/10.1002/2016gl072430
Gottschalck, J., Roundy, P. E., Schreck, C. J., III, Vintzileos, A., & Zhang, C. (2013). Large-scale atmospheric and oceanic conditions during the 2011–12 dynamo field campaign. *Monthly Weather Review*, *141*(12), 4173–4196. https://doi.org/10.1175/mwr-d-13-00022.1
Hagos, S. M., Zhang, C., Feng, Z., Burleyson, C. D., DeMott, C., Kerns, B., et al. (2016). The impact of the diurnal cycle on the propagation of Madden–Julian Oscillation convection across the maritime continent. *Journal of Advances in Modeling Earth Systems*, *8*(4), 1552–1564. https://doi.org/10.1002/2016ms000725
Henderson, S. A., Maloney, E. D., & Son, S.-W. (2017). Madden–Julian Oscillation Pacific teleconnections: The impact of the basic state and MJO representation in general circulation models. *Journal of Climate*, *30*(12), 4567–4587. https://doi.org/10.1175/jcli-d-16-0789.1
Hendon, H. H., & Salby, M. L. (1994). The life cycle of the Madden–Julian Oscillation. *Journal of the Atmospheric Sciences*, *51*(15), 2225–2237. https://doi.org/10.1175/1520-0469(1994)051<2225:tlcotm>2.0.co;2
Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2018a). ERA5 hourly data on pressure levels from 1979 to present [Dataset]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.bd0915c6
Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2018b). ERA5 hourly data on single levels from 1979 to present [Dataset]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.7283/633E-1497
Jiang, X. (2017). Key processes for the eastward propagation of the Madden–Julian Oscillation based on multimodel simulations. *Journal of Geophysical Research: Atmospheres*, *122*(2), 755–770. https://doi.org/10.1002/2016jd025955
Jiang, X., Adames, Á. F., Zhao, M., Waliser, D., & Maloney, E. (2018). A unified moisture mode framework for seasonality of the Madden–Julian Oscillation. *Journal of Climate*, *31*(11), 4215–4224. https://doi.org/10.1175/jcli-d-17-0671.1
Kang, D., Kim, D., Ahn, M.-S., & An, S.-I. (2021). The role of the background meridional moisture gradient on the propagation of the MJO over the maritime continent. *Journal of Climate*, *34*(16), 6565–6581.
Kiladis, G. N., Straub, K. H., & Haertel, P. T. (2005). Zonal and vertical structure of the Madden–Julian Oscillation. *Journal of the Atmospheric Sciences*, *62*(8), 2790–2809. https://doi.org/10.1175/jas3520.1
Kim, D., Kim, H., & Lee, M.-I. (2017). Why does the MJO detour the maritime continent during austral summer? *Geophysical Research Letters*, *44*(5), 2579–2587. https://doi.org/10.1002/2017gl072643
Kim, D., Kug, J.-S., & Sobel, A. H. (2014). Propagating versus nonpropagating Madden–Julian Oscillation events. *Journal of Climate*, *27*(1), 111–125. https://doi.org/10.1175/jcli-d-13-00084.1
Kim, H., Ham, Y., Joo, Y., & Son, S. (2021). Deep learning for bias correction of MJO prediction. *Nature Communications*, *12*(1), 1–7. https://doi.org/10.1038/s41467-021-23406-3
Kim, H., Janiga, M. A., & Pegion, K. (2019). MJO propagation processes and mean biases in the SubX and S2S reforecasts. *Journal of Geophysical Research: Atmospheres*, *124*(16), 9314–9331. https://doi.org/10.1029/2019jd031139
Kim, H., Vitart, F., & Waliser, D. E. (2018). Prediction of the Madden–Julian Oscillation: A review. *Journal of Climate*, *31*(23), 9425–9443. https://doi.org/10.1175/jcli-d-18-0210.1
Kim, H.-M. (2017). The impact of the mean moisture bias on the key physics of MJO propagation in the ECMWF reforecast. *Journal of Geophysical Research: Atmospheres*, *122*(15), 7772–7784. https://doi.org/10.1002/2017jd027005
Kiranmayi, L., & Maloney, E. D. (2011). Intraseasonal moist static energy budget in reanalysis data. *Journal of Geophysical Research*, *116*(D21), D21117. https://doi.org/10.1029/2011jd016031
Liebmann, B., & Smith, C. A. (1996). Description of a complete (interpolated) outgoing longwave radiation dataset. *Bulletin of the American Meteorological Society*, *77*(6), 1275–1277.
Lim, Y., Son, S.-W., & Kim, D. (2018). MJO prediction skill of the subseasonal-to-seasonal prediction models. *Journal of Climate*, *31*(10), 4075–4094. https://doi.org/10.1175/jcli-d-17-0545.1
Lim, Y.-K., Arnold, N. P., Molod, A. M., & Pawson, S. (2021). Seasonality in prediction skill of the Madden–Julian Oscillation and associated dynamics in version 2 of NASA's GEOS-S2S forecast system. *Journal of Geophysical Research: Atmospheres*, *126*(18), e2021JD034961. https://doi.org/10.1029/2021jd034961
Madden, R. A., & Julian, P. R. (1971). Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *Journal of the Atmospheric Sciences*, *28*(5), 702–708. https://doi.org/10.1175/1520-0469(1971)028<0702:doadoi>2.0.co;2
Madden, R. A., & Julian, P. R. (1972). Description of global-scale circulation cells in the tropics with a 40–50 day period. *Journal of the Atmospheric Sciences*, *29*(6), 1109–1123. https://doi.org/10.1175/1520-0469(1972)029<1109:dogscc>2.0.co;2
Maloney, E. D. (2009). The moist static energy budget of a composite tropical intraseasonal oscillation in a climate model. *Journal of Climate*, *22*(3), 711–729. https://doi.org/10.1175/2008jcli2542.1
Matthews, A. J. (2008). Primary and successive events in the Madden–Julian Oscillation. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *134*(631), 439–453. https://doi.org/10.1002/qj.224
Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., et al. (2014). Decadal climate prediction: An update from the trenches. *Bulletin of the American Meteorological Society*, *95*(2), 243–267. https://doi.org/10.1175/bams-d-12-00241.1
Penny, S. G., Akella, S., Balmaseda, M. A., Browne, P., Carton, J. A., Chevallier, M., et al. (2019). Observational needs for improving ocean and coupled reanalysis, S2S prediction, and decadal prediction. *Frontiers in Marine Science*, *6*, 391. https://doi.org/10.3389/fmars.2019.00391
Seo, H., Subramanian, A. C., Miller, A. J., & Cavanaugh, N. R. (2014). Coupled impacts of the diurnal cycle of sea surface temperature on the Madden–Julian Oscillation. *Journal of Climate*, *27*(22), 8422–8443. https://doi.org/10.1175/jcli-d-14-00141.1
Sobel, A., & Maloney, E. (2013). Moisture modes and the eastward propagation of the MJO. *Journal of the Atmospheric Sciences*, *70*(1), 187–192. https://doi.org/10.1175/jas-d-12-0189.1
Vitart, F., Balsamo, G., Bidlot, J., Lang, S., Tsonevsky, I., Richardson, D., & Balmaseda, M. (2019). *Use of ERA5 to initialize ensemble reforecasts*. European Centre for Medium Range Weather Forecasts.
Waliser, D., Lau, K., Stern, W., & Jones, C. (2003). Potential predictability of the Madden–Julian Oscillation. *Bulletin of the American Meteorological Society*, *84*(1), 33–50. https://doi.org/10.1175/bams-84-1-33
Wheeler, M. C., & Hendon, H. H. (2004). An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Monthly Weather Review*, *132*(8), 1917–1932. https://doi.org/10.1175/1520-0493(2004)132<1917:aarmmi>2.0.co;2
Wu, C.-H., & Hsu, H.-H. (2009). Topographic influence on the MJO in the maritime continent. *Journal of Climate*, *22*(20), 5433–5448. https://doi.org/10.1175/2009jcli2825.1
Zhang, C. (2013). Madden–Julian Oscillation: Bridging weather and climate. *Bulletin of the American Meteorological Society*, *94*(12), 1849–1870. https://doi.org/10.1175/bams-d-12-00026.1
Zhang, C., & Ling, J. (2017). Barrier effect of the Indo-Pacific maritime continent on the MJO: Perspectives from tracking MJO precipitation. *Journal of Climate*, *30*(9), 3439–3459. https://doi.org/10.1175/jcli-d-16-0614.1

Zhang, L., & Han, W. (2018). Impact of Ningaloo Niño on tropical Pacific and an interbasin coupling mechanism. *Geophysical Research Letters*, *45*(20), 11–300. https://doi.org/10.1029/2018gl078579

Zhang, L., & Han, W. (2020). Barrier for the eastward propagation of Madden–Julian Oscillation over the maritime continent: A possible new mechanism. *Geophysical Research Letters*, *47*(21), e2020GL090211. https://doi.org/10.1029/2020gl090211

Zuo, H., Balmaseda, M., De Boisseson, E., Hirahara, S., Chrust, M., & De Rosnay, P. (2017). A generic ensemble generation scheme for data assimilation and ocean analysis.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., & Mayer, M. (2019). The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: A description of the system and assessment. *Ocean Science*, *15*(3), 779–808. https://doi.org/10.5194/os-15-779-2019