We propose a set of MJO teleconnection diagnostics that enables an objective evaluation of model simulations, a fair model-to-model comparison, and a consistent tracking of model improvement. Various skill metrics are derived from teleconnection diagnostics including five performance-based metrics that characterize the pattern, amplitude, east–west position, persistence, and consistency of MJO teleconnections and additional two process-oriented metrics that are designed to characterize the location and intensity of the anomalous Rossby wave source (RWS). The proposed teleconnection skill metrics are used to compare the characteristics of boreal winter MJO teleconnections (500-hPa geopotential height anomaly) over the Pacific–North America (PNA) region in 29 global climate models (GCMs). The results show that current GCMs generally produce MJO teleconnections that are stronger, more persistent, and extend too far to the east when compared to those observed in reanalysis. In general, models simulate more realistic teleconnection patterns when the MJO is in phases 2–3 or phases 7–8, which are characterized by a dipole convection pattern over the Indian Ocean and western to central Pacific. The higher model skill for phases 2, 7, and 8 may be due to these phases producing more consistent teleconnection patterns between individual MJO events than other phases, although the consistency is lower in most models than observed. Models that simulate realistic RWS patterns better reproduce MJO teleconnection patterns.

1. Introduction

The Madden–Julian oscillation (MJO) (Madden and Julian 1971, 1972) is the dominant mode of tropical intraseasonal variability. It is characterized by a convection–circulation coupled system propagating eastward from the Indian Ocean to the Pacific with periods ranging from approximately 30 to 60 days. The MJO modulates atmospheric (e.g., tropical cyclones), oceanic (e.g., chlorophyll), and ocean–atmosphere coupled [e.g., El Niño–Southern Oscillation (ENSO)] disturbances/phenomena in the tropics (e.g., Neale et al. 2008; Lau and Waliser 2011; Jin et al. 2013; Zhang 2013; Kim et al. 2014b) and the impact extends outside of the tropical region. Anomalous MJO-induced upper-level
divergence can generate anomalous Rossby wave source (RWS) by producing divergent flow anomalies in the region of the strong absolute vorticity and its gradient associated with the midlatitude North Pacific westerly jet (Sardeshmukh and Hoskins 1988). Rossby waves excited by the tropical heating propagate poleward and eastward into the extratropics, modulating circulations therein (e.g., Horel and Wallace 1981; Hoskins and Karoly 1981).

The influence of tropical heating variations on the extratropics frequently manifests as changes in the teleconnection patterns that link variability over remote regions (Wallace and Gutzler 1981). For example, in the North Pacific, a negative Pacific–North America (PNA)-like pattern can be triggered about 5–10 days after MJO phase 3 (e.g., Hsu 1996; Mori and Watanabe 2008; Seo and Lee 2017; Tseng et al. 2019), which is characterized by enhanced MJO convection located over the eastern Indian Ocean and suppressed convection over the western Pacific. A positive PNA-like pattern can be triggered after MJO phase 7, which corresponds to reduced convection over the eastern Indian Ocean and enhanced convection over the western Pacific. In the North Atlantic, a positive North Atlantic Oscillation (NAO) (Walker and Bliss 1932) is more likely to occur about 5–15 days after MJO phase 3, and a negative NAO is likely to occur after phase 7 (Cassou 2008; Lin et al. 2009). The MJO can thus modulate midlatitude weather systems via modulating these teleconnection patterns. Henderson et al. (2016) found a significant decrease in eastern Pacific and Atlantic blocking events following MJO phase 3, and an increase following phase 7. Deng and Jiang (2011) found a poleward shift of North Pacific storm track about two weeks after MJO phase 3, and a southward shift after phase 7. A detailed review of various MJO impacts on the extratropics is offered by Stan et al. (2017).

Considering the substantial global impacts of the MJO and its teleconnections, it is crucial to accurately simulate and predict characteristics and variability of the MJO and its teleconnections in global climate models (GCMs). Numerous multimodel comparison studies have been conducted on MJO simulations in recent decades (e.g., Lin et al. 2006; Hung et al. 2013; Ahn et al. 2017). Performance-based MJO simulation diagnostics and metrics have been developed for a consistent evaluation of the MJO fidelity in models (Waliser et al. 2009). Although MJO simulation has generally improved in models from phase 5 of the Coupled Model Intercomparison Project (CMIP5) over CMIP3 models in terms of MJO variance and eastward propagation (Hung et al. 2013), models still tend to produce a weaker MJO with faster eastward propagation (e.g., Kim et al. 2014a; Jiang et al. 2015; Ahn et al. 2017). To understand the causes of such biases in models, various process-based MJO metrics have been developed recently (e.g., Kim et al. 2014a, 2015; Jiang et al. 2015; Ahn et al. 2017). Development of these metrics has elucidated fundamental MJO processes which, in turn, help guide model developments toward better MJO simulations (Jiang et al. 2019; Maloney et al. 2019).

Compared to the tropical MJO variability, evaluating the simulation ability of extratropical MJO teleconnections in GCMs has received less attention. Only a few recent studies discuss model performance in simulating MJO teleconnections. Yoo et al. (2015) showed that MJO teleconnections are substantially improved when a unified convection scheme (Park 2014) is used instead of the default shallow and deep convection schemes employed in the Community Atmosphere Model, version 5 (CAM5). Stan and Straus (2019) also showed that simulation and prediction of MJO teleconnections are improved by the quasi-explicit representation of moist convection using superparameterization (Grabowski 2001; Khairoutdinov and Randall 2001). Wu et al. (2016) found that the Beijing Climate Center atmospheric GCM (AGCM) is able to reproduce the MJO–NAO relationship for MJO phase 3 but not for phase 6. Henderson et al. (2017), which was the first multimodel comparison study on MJO teleconnections, evaluated MJO teleconnection patterns in 10 CMIP5 models and found that biases in simulating the position of the Pacific westerly jet cause errors in simulating MJO teleconnections, in addition to the effects of poor MJO performance. By comparing multimodel hindcasts, Vitart (2017) found that the amplitude of predicted MJO teleconnections is often too strong over the western North Pacific compared to reanalysis.

Our understanding of the current model performance in terms of MJO teleconnection simulation is still limited. One reason that makes rigorous model evaluation and comparison difficult is the lack of standardized MJO teleconnection diagnostics. Studies have used different diagnostics, making it hard to consistently assess and compare MJO teleconnection performance among different models and across model generations. Even for the diagnostic that most studies have used (i.e., the lag composite of geopotential height anomalies), the pressure level and the time lag used in the composites differ among studies. For example, Henderson et al. (2017) used 250 hPa and a 0–4-day lag whereas Lin and Brunet (2018) used 500 hPa and a 5–9-day lag. Tseng et al. (2018) and many other studies noticed remarkable differences of teleconnection patterns over different time lags for some MJO phases (e.g., phases 1 and 5). The teleconnection patterns also shift slightly poleward from...
lower to higher pressure levels (Kosaka and Nakamura 2006). Therefore, a standardized set of diagnostics/ metrics of MJO teleconnections is needed to objectively evaluate model success at simulating MJO teleconnections in a consistent framework, as will be derived in this study. Development of MJO teleconnection metrics would also enable the science community to quantitatively investigate the characteristics and mechanisms of MJO teleconnections as they are modulated by large-scale variability such as ENSO (Moon et al. 2011) or the quasi-biennial oscillation (QBO; Son et al. 2017; Wang et al. 2018).

The main goal of this study is to develop a set of standardized MJO teleconnection diagnostics and metrics. Specifically, those that can characterize pattern, amplitude, east–west position, persistence, and consistency of MJO teleconnections will be proposed. Process-based teleconnection metrics are also developed based on the linear Rossby wave theory to highlight the key dynamical processes associated with MJO teleconnections. Using the proposed diagnostics, our study extends the evaluation of MJO teleconnections performed in Henderson et al. (2017) to 29 GCMs with a focus on the PNA region in boreal winter. This expanded analysis will provide an assessment of model strengths and shortcomings in simulating MJO teleconnections in the current generation of climate models.

The paper is organized as follows. The GCMs and reference data are introduced in section 2. Section 3 describes five performance-based teleconnection skill metrics and evaluates MJO teleconnection performance in current GCMs. Two metrics derived from process-oriented diagnostics are described in section 4. Section 5 gives a summary and discussion. In a future companion paper, the role of the MJO and basic state in MJO teleconnection performance will be examined by relating teleconnection skill metrics developed in this study with various MJO and basic state metrics and by running a linear baroclinic model.

2. GCMs and reference dataset

A total of 29 GCMs are analyzed in this study (Table 1): 22 are from the CMIP5 archive (Taylor et al. 2012), 6 from the GEWEX (Global Energy and Water Cycle Experiment) Atmospheric System Study (GASS) and Year of Tropical Convection (YoTC) project (Jiang et al. 2015; Jiang 2017), and 1 Atmospheric Model Intercomparison Project (AMIP) run from the ECMWF (Davini et al. 2017) is used. Daily output such as horizontal winds, geopotential height, and outgoing longwave radiation (OLR) from a historical run ensemble member of each model is used. The period analyzed in this study is 1975–2005 for CMIP5 models, corresponding approximately to the period of the reference dataset described below. We analyzed the first ensemble member (r1i1p1) of each CMIP5 model to make a consistent comparison as some CMIP5 models only have one ensemble member. The GASS/YoTC project is a global model evaluation project with a specific focus on the processes associated with the MJO. Models derived from this project include both AGCMs and atmosphere–ocean coupled GCMs. For AGCMs, weekly sea surface temperature (SST) and sea ice concentration (SIC) derived from the NOAA Optimum Interpolation V2 dataset (Reynolds et al. 2002) were used as the boundary conditions. All models were integrated for 20 years and archived from 1991 to 2010, with the exception of SPCAM3, which is only archived from 1986 to 2003 for a total of 18 years. The ECMWF AMIP historical run was run with the Integrated Forecast System (IFS; cycle 36r4) atmospheric circulation model. The forcing and boundary conditions are set according to the CMIP5 historical forcing with SST and SIC derived from the Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST2.1.1). The experiments extend from 1979 to 2008, but output was only available at a period of 1980–2000 at the time of the analysis. Among 10 ensemble members, we analyzed the ensemble member that shows relatively realistic eastward MJO propagation, which is determined by its highest pattern correlation calculated against observed Hovmöller diagram of 25–90-day-filtered OLR anomalies with respect to convection over the eastern Indian Ocean following Jiang et al. (2015).

The reference OLR data are derived from the NOAA Advanced Very High Resolution Radiometer (AVHRR; Liebmann and Smith 1996) product dataset, while other variables are from the ECMWF interim reanalysis dataset (ERA-Interim, herein ERAI; Dee et al. 2011). The reference data period is from 1979 to 2017. Model output are interpolated to a 2.5° × 2.5° horizontal grid to be consistent with the reference data. We focus our analysis from October to March when the MJO is most active and when Northern Hemisphere teleconnections are most prominent.

The Real-time Multivariate MJO (RMM) indices are used to characterize the phase and amplitude of the MJO (Wheeler and Hendon 2004). The RMM indices of the reference data are obtained by projecting 15°S–15°N averaged OLR and 850- and 200-hPa zonal wind anomalies onto the observed combined empirical orthogonal function (CEOF) eigenvectors (Wheeler and Hendon 2004). To allow consistent comparison, the model RMM indices are also obtained by projecting their corresponding OLR and zonal wind anomalies onto the observed CEOF eigenvectors following...
Henderson et al. (2017). This projection method, however, can possibly lead to an artificially higher chance of model success in MJO simulations since modeled CEOFs may look different from the observed (Ahn et al. 2017). Anomalies are derived by subtracting the first three harmonics of the climatological seasonal cycle, and the most recent 120-day mean to reduce the influence of interannual variability following Wheeler and Henderson (2004). Normalization by the observed global variance is further applied prior to calculating the RMM indices. This study focuses on composites of strong MJO events defined as events with the RMM amplitude greater than one standard deviation.

3. Performance-based MJO teleconnection skill metrics

a. MJO teleconnection simulation (teleconnection skill metrics T1–T3)

MJO teleconnections are defined as the composites of 500-hPa geopotential height anomalies (Z500a) averaged over 5–9 days after each MJO phase (results are not sensitive to different pentad-means within the 3–11-day-lag time window). This average length (i.e., 5–9-day lag) is chosen to obtain a relatively stable extratropical response to the MJO given that it takes about a week or two for a tropically forced Rossby wave packet to reach high latitudes (Hoskins and Karoly 1981). This lag is applied to all Z500a results unless stated otherwise. Our analysis is mostly focused on MJO phases 3 and 7, which are the most effective MJO phases in exciting extratropical circulation anomalies (Stan et al. 2017; Tseng et al. 2019) and modulating extratropical events, such as eastern Pacific and Atlantic blockings, North Pacific storm tracks, atmospheric rivers, and extreme precipitation (e.g., Deng and Jiang 2011; Henderson et al. 2016; Mundhenk et al. 2016, 2018; Baggett et al. 2017; Wang et al. 2018). Figure 1 shows the Z500a and OLR anomaly composites for MJO phase 3 for the reference dataset and the 29 models. To test the significance of the results, a two-tailed Student’s t test is used, which has been advocated for application by modeling centers (Decremer et al. 2014) and does not require expensive computational cost compared to other advanced
techniques like bootstrapping. Because of the continuity of MJO activity between the phases, the degrees of freedom are determined as the number of days in each MJO phase bin divided by 5, which is the average duration in days of an individual MJO phase (Alaka and Maloney 2012; Henderson et al. 2016). The phase 3 teleconnection pattern over the PNA region (20°–80°N, 120°E–60°W) resembles a negative PNA pattern (Fig. 1a). Models can generally capture the phase 3 teleconnection pattern, especially the positive Z500a in the North Pacific, but with biases in location and amplitude. To quantitatively measure the capability of model simulations, three teleconnection skill metrics (T1–T3) are defined based on Fig. 1 and described below.

For T1, the pattern correlation coefficient (pattern CC) is calculated between the ERAI and model Z500a composites over the PNA region (20°–80°N, 120°E–60°W; black box in Fig. 1a) [similar to Yoo et al. (2015) and Henderson et al. (2017)]. This region is chosen for calculation because the MJO-associated Z500a vary within this approximate boundary (e.g., Fig. 4) and Z500a in

---

Fig. 1. 5–9-day averaged lagged response of Z500a (shading; m) to MJO phase 3 [OLR contours; green (brown) represents enhanced (suppressed) convection; interval: 10 W m⁻²] in (a) ERAI/NOAA and (b)–(dd) models. OLR anomalies are shown for day 0 (number and percentage of days in strong MJO phase are shown in the bottom-right corner). The dotted areas represent significant Z500a exceeding the 95% confidence level. The values in the top-right corner represent metrics pattern CC (T1, left value) and relative amplitude (T2, right value) for each model. Black box (20°–80°N, 120°E–60°W) in (a) denotes the PNA region while red box (30°–50°N, 140°E–120°W) represents the region for calculation of east–west position (T3) for phase 3 (middle red line indicates the longitude of maximum center of Z500a).
some models (e.g., CanESM2, MIROC-ESM, and NCAR-CAM5) extend westward toward 120°E in MJO phase 3. But results of pattern CC are not sensitive to regions changing from 110° to 130°E. The pattern CC quantifies the general pattern reproduction of MJO teleconnections in a model. For example, HadGEM2-CC (Fig. 1m, pattern CC of 0.88) simulates more realistic Z500a pattern than MIROC-ESM (Fig. 1p, pattern CC of 0.10) for MJO phase 3. However, models with higher pattern CC do not necessarily have Z500a values that are more significant. For example, ACCESS1.0 and GFDL-ESM2G (Figs. 1b,j) have high pattern CC but Z500a values are not significant over North America. Among the eight MJO phases, the multimodel mean of pattern CC is the largest for MJO phase 8 (0.66), followed by phase 3 (0.65), phase 7 (0.64), and phase 2 (0.61). In general, models simulate more realistic Z500a pattern for these MJO phases.

For T2, relative amplitude is defined as the model Z500a composite standard deviation over the PNA region divided by that of ERAI (Taylor 2001). This metric represents the relative amplitude of MJO teleconnections in a model compared to that in ERAI, where a value greater (less) than 1 indicates an overestimate (underestimate) of the teleconnection amplitude.

For T3, the east–west position is defined. This metric is developed given that current GCMs have strong zonal biases in the Pacific jet position that can lead to east–west shifts in MJO teleconnections (Henderson et al. 2017). East–west positions of MJO teleconnections are captured as follows. First, we find the location of the

Fig. 1. (Continued)
maximum absolute Z500a in the ERAI 5–9-day composite within the PNA region (e.g., 40°N, 170°W for phase 3). The region is then expanded by 50° westward and eastward, and 10° northward and southward from the maximum center (e.g., 30°–50°N, 140°E–120°W for phase 3). This region is used to ensure the calculation is focused on the target Z500a selected from ERAI because the maximum Z500a in a model may be over different regions. The east–west position is then calculated as the “Z500a-weighted” average longitude given from 0° to 360°, which is computed by the sum of longitude multiplied by Z500a and divided by the sum of Z500a across all longitudes within the selected region. Only anomalies with the same sign of the maximum center are included in calculation. The result for ERAI is 191° (~170°W) for phase 3, consistent with the longitude of the maximum center. A larger (smaller) metric value than ERAI indicates an eastward (westward) shift of MJO teleconnections. (Selected regions for each MJO phase are shown in red boxes in Fig. 4.)

To visualize the model ability in simulating MJO teleconnections, a Taylor diagram is constructed. It is designed to display the pattern CC (T1) (curve axis), relative amplitude (T2) (x and y axes), and the centered root-mean-square error (RMSE) (distance from the model point to the “REF” point) in a single diagram as they satisfy a geometric relationship [see details in Taylor (2001)]. The Taylor diagram of 5–9-day Z500a composites over the PNA region for MJO phases 3 and 7 is shown in Fig. 2. Generally, most models capture the overall Z500a pattern with a pattern CC exceeding 0.5. Specifically, CNRM-CM5 (Fig. 1h), HadGEM2-CC (Fig. 1m), and ECMWF (Fig. 1dd) have outstanding performance in simulating a realistic Z500a pattern related to phase 3 (pattern CC > 0.8) while CMCC-CESM, GFDL-ESM2M, and NorESM1-M produce the most realistic Z500a pattern for phase 7. However, HadGEM2-AO and MRI-ESM1 have a pattern CC < 0.5 for both MJO phases 3 and 7. There are models (MIROC-ESM, MIROC-ESM-CHEM, MPI-ESM-MR, MPI-ESM-P, and MRI-CGCM3) that produce a realistic Z500a pattern for one phase but not the other. This inconsistency of model simulations between the mirror MJO phases is also discussed in Henderson et al. (2017), who point out that a model may be able to reproduce the extratropical response, for example, to MJO phase 3, but not be able to reproduce the opposite response to phase 7.

The relative amplitude (T2) in Fig. 2 indicates that most models tend to overestimate (relative amplitude > 1.0) the amplitude of Z500a for both phase 3 (90% of models) and phase 7 (69% of models). The magnitude of the bias is generally larger for phase 3. The general overestimate of the Z500a amplitude is also found in other MJO phases except for phases 2 and 8 (not shown).

The east–west position (T3) of MJO teleconnections is summarized in Fig. 3a. About 60% of models have a westward-shifted MJO teleconnection for phase 3, while most (~86%) models show an eastward-shifted teleconnection for phase 7. MIROC-ESM, MIROC-ESM-CHEM, MPI-ESM-LR, TAMU-CAM4, and NCAR-CAM5 (ACCESS1.0, CMCC-CMS, GFDL CM3, IPSL-CM5A-LR, the MRI models, SPCAM3, and SPCCSM3) show a noticeable westward (eastward) shift in MJO teleconnections relative to ERAI for both MJO phases 3 and 7. The eastward shift in MJO teleconnections is also found in most of the remaining MJO phases (i.e., phases 1, 2, 6, and 8; not shown), indicating that models generally produce an eastward shift in MJO teleconnections compared to ERAI.

To identify regions over which the models have the highest ability to reproduce the observed Z500a pattern, sign agreement of 5–9-day Z500a composites is calculated between the models and ERAI. Results for individual MJO phases are shown in Fig. 4. The shadings show the ERAI Z500a composites. The dotted regions are where over 24 out of 29 models (>82%) produce the same sign as the ERAI Z500a composites. Over the North Pacific, the dots mostly overlap with the ERAI anomalies, indicating that the models generally reproduce a realistic Z500a pattern over the North Pacific. However, over North America, dots are sparse, especially for MJO phases 2 and 6, suggesting that MJO teleconnections over North America are more difficult...
to reproduce in models than that over the North Pacific for these MJO phases.

b. Pattern consistency of MJO teleconnections (teleconnection skill metric T4)

MJO teleconnection patterns depend strongly on the characteristics of the MJO, such as its amplitude (Wang et al. 2018), propagation speed (Bladé and Hartmann 1995; Yadav and Straus 2017), and continuity of propagation (Adames and Wallace 2014; Bao and Hartmann 2014). Teleconnection patterns are also sensitive to interannual variations in the basic state, such as ENSO (Moon et al. 2011) and the QBO (Son et al. 2017; Wang et al. 2018). The structure of an MJO can vary between MJO events under different basic states, thus producing variations in the resulting teleconnection patterns. Such event-to-event variations in the teleconnection patterns can lead to uncertainties when forecasting the MJO-related extratropical circulation anomalies. Tseng et al. (2018) suggested that some MJO phases tend to produce more consistent teleconnection patterns than other phases, and thus have the potential for more reliable prediction. To quantify how this feature is simulated in models, T4 is defined below based on Tseng et al. (2018) with some modification.

For T4, intraphase pattern consistency (IPC) is defined. For a given MJO phase, pattern CC is calculated between the day-5–9 Z500a composite (e.g., Fig. 1 for phase 3) and the Z500a of each MJO event over the PNA region. This is done for ERAI and each model. When the pattern CC is greater than 0.5, the corresponding Z500a pattern of that MJO event is considered to be consistent with the day-5–9 composite of that phase. The occurrence frequency (%) of events with a pattern CC > 0.5 is defined as IPC. An MJO phase with a larger IPC suggests that the teleconnection pattern related to that phase is more consistent between individual MJO events and hence more robust. The results for ERAI and all models are shown in Figs. 3b and 5. The values of IPC in ERAI have a similar range as Tseng et al. (2018). In ERAI, teleconnection patterns related to phases 2, 7, and 8 are more consistent than the other phases (IPC ~20%), indicating that extratropical variability related to these phases tend to be more predictable. Pattern consistency for phases 3–6 is relatively low (IPC ~5%–10%), which poses a great challenge for accurate prediction. The finding that MJO phase 3 has low IPC differs from the results of Tseng et al. (2018), who suggested a highly consistent teleconnection pattern related to phase 3 at a lag time of 0–10 days. This discrepancy is noticed in pentad means within the 2–10-day-lag time window and is likely due to different methodologies employed. We calculate the pattern CC of each event against the composite pattern whereas Tseng et al. (2018) calculated paired pattern CC for every two MJO events. Among the 29 models, only CMCC-CESM, GFDL-CM3, NorESM1-M, and SPCAM3 are able to capture the feature that phases 2, 7, and 8 have the highest IPC (not shown). Most (~65%) models have a common problem that they produce less consistent teleconnection patterns for phases 2, 7, and 8 but overestimate the consistency of teleconnection patterns for the remaining phases. It can be noticed in Fig. 5 that the multimodel mean of IPC is ~3% smaller for phase 7 and ~10% larger for phase 3 than IPC in ERAI. Biases in IPC are especially large in CanESM2 and NCAR-CAM5 (CMCC-CESM, MPI-ESM-MR, and MPI-ESM-P) for MJO phase 3 (7) (Fig. 3b).

c. Persistence of MJO teleconnections (teleconnection skill metric T5)

MJO teleconnections may persist for a number of days depending on whether and how fast the parent MJO transitions to another phase or decays. Long MJO teleconnection persistence can lead to an accumulation of related weather events and hence results in severe
impacts. To quantify the persistence of MJO teleconnections, daily-mean composites of Z500a from day 5 to day 40 of a given MJO phase are first calculated. This set of maps indicates the evolution of the MJO-related Z500a. Pattern CC is then calculated for each MJO phase between this set of evolution maps and the lagged 5–9-day pentad-mean Z500a composite (e.g., Fig. 1a) over the PNA region (similar to the concept of autocorrelation). By doing so, we obtain a time series of pattern CC that changes with lag from 5 to 40 days. The result for ERAI is shown in Fig. 6 (only lags from 5 to 25 days are shown). T5 is developed based on this figure.

For T5, persistence is defined as the number of days with pattern CC being larger than 0.5. In ERAI, Z500a related to phases 3, 6, and 7 have longer persistence (12 days for phase 3 and 13 days for phases 6 and 7) than the other phases. Among the 29 models, only CNRM-CM5 and NCAR-CAM5 produce similar results to ERAI with phases 3, 6, and 7 having the longest persistence for about two weeks (not shown). Longer teleconnection persistence for phases 3, 6, and 7 is possibly related to the dipole pattern of the RWS (defined in section 4) (e.g., Fig. 7 for phase 3). Opposite-signed RWS anomalies on the east and west ends of the Asian jet generate similar teleconnection patterns that interfere constructively and are associated with a slower decay of the teleconnection response (Tseng et al. 2019). An RWS that is symmetric about the jet produces interfering teleconnection responses that make the mid-latitude geopotential anomaly pattern highly sensitive to modest variations in the RWS. To support this contention, for phases 4 and 5 an RWS pattern symmetric about the jet occurs that causes destructive interference and a faster decay of the response (Tseng et al. 2019). Similar results are found in Goss and Feldstein (2017, 2018). Models tend to simulate a longer teleconnection persistence than ERAI for MJO phase 3 (~63% of models) (Fig. 3c). Specifically, MJO teleconnections

---

**FIG. 4.** Sign agreement of Z500a between ERAI and the models. Shadings are the 5–9-day averaged lagged response of Z500a to MJO phases for ERAI (m). Dotted regions indicate where over 24 out of 29 (>82%) models produce the same-sign anomaly with ERAI. The bottom color bar indicates the percentage of models that have the same sign as ERAI (a minus sign is added in front of the percentage to represent results for the negative anomalies). Red boxes denote regions for the calculation of east–west position (T3) for each phase, which shift along with the North Pacific Z500a (middle red line indicates the longitude of maximum center of Z500a).
4. Process-oriented diagnosis and teleconnection skill metrics

In this section, process-oriented diagnosis is applied to the models. Process-based metrics derived from this diagnosis are relevant to MJO teleconnection dynamics and hence provide information on critical processes that affect MJO teleconnection simulation skill.

As discussed in the introduction, the formation of the MJO-associated extratropical circulation anomalies is primarily explained by the Rossby wave propagation. Based on the linear, stationary Rossby wave theory, the RWS is often examined to diagnose the origin of the propagating Rossby wave (e.g., Sardeshmukh and Hoskins 1988; Straus et al. 2015; Seo and Lee 2017). Following Sardeshmukh and Hoskins (1988), the linearized RWS is expressed as follows:

$$\text{RWS} = \left( -\xi \nabla \cdot \mathbf{V}_x - \xi \nabla \cdot \mathbf{V}_y \right) + \left( -\mathbf{V}_x \cdot \nabla \xi - \mathbf{V}_y \cdot \nabla \xi \right).$$ (1)

where \( \mathbf{V}_x \) is the divergent component of horizontal winds, \( \xi \) is the absolute vorticity, the prime represents the anomalies, and the overbar denotes the seasonal mean. As demonstrated by Eq. (1), the RWS consists of a divergence and advection component. The first two terms indicate anomalous vorticity generation due to divergence of anomalous (first term) and climatological (second term) divergent winds. The last two terms represent anomalous vorticity generation due to advection of mean absolute vorticity by the anomalous divergent winds (third term) and advection of anomalous absolute vorticity by the climatological divergent winds (fourth term).

The 0–4-day-lagged averaged 250-hPa RWS anomalies associated with MJO phase 3 for ERAI and all models are shown in Fig. 7. This shorter lag (0–4-day instead of 5–9-day lag used in Figs. 1–6) is chosen to obtain a RWS response that is more directly related to the MJO heating with fewer impacts from midlatitude forcing. Climatology of 250-hPa zonal wind in ERAI and all models are also shown. In response to the enhanced convection over the Indian Ocean and suppressed convection over the western Pacific, a maximum negative RWS over southern Asia and a positive RWS over the eastern North Pacific are seen (Fig. 7a). An anomalous RWS of opposite sign is found for MJO phase 7 (not shown). Compared to ERAI, some longitudinal displacements of the RWS are simulated in models (Figs. 7b–dd). Henderson et al. (2017) suggests that a biased east–west position of the RWS may lead to a corresponding shift of MJO teleconnections. Such shifts are seen in most models such as CNRM-CM5, GFDL models, MRI models, and TAMU-CAM4.
To further investigate the dominant terms that contribute to the RWS and teleconnection performance, RWS anomalies are partitioned into individual terms in Eq. (1) and compared to the total RWS. The results show that the first term \(-\xi \nabla \cdot \mathbf{V}_y\) in ERAI has nearly identical spatial distribution and equivalent magnitude to the total RWS (not shown), indicating its dominant contribution to the RWS, consistent with some previous studies (Hsu 1996; Lin 2009; Seo and Son 2012; Takahashi and Shirooka 2014; Henderson et al. 2017).

The seasonal-mean background absolute vorticity \(\xi\) and the MJO-induced divergence \(-\nabla \cdot \mathbf{V}_y\) therefore play important roles in the RWS and its associated teleconnection anomalies.

The RWS pattern related to MJO phase 3 in Fig. 7 can be explained as follows. The MJO enhanced convection over the Indian Ocean is coupled with an ascending motion and upper-level divergence. The MJO-associated upper-level divergence field acts on the region of strong cyclonic vorticity on the north side of the subtropical westerly jet, giving rise to an anomalous negative RWS formed to the north of the MJO convection over southern Asia. The positive RWS at the jet exit region over the northeastern Pacific is an effect of the MJO-associated anomalous convergence related to the suppressed convection over the western Pacific. Given that the RWS is positively contributing to the vorticity generation (Sardeshmukh and Hoskins 1988), a negative (positive) RWS corresponds to an anomalous anticyclonic (cyclonic) vorticity generation in the Northern Hemisphere, and thus a positive Z500a formed locally over the North Pacific and negative Z500a found over the northwestern North America (Fig. 1a). Anomalous vorticity generated by the opposite-signed RWS anomalies triggers Rossby waves that propagate toward the PNA region and interfere constructively, leading to the teleconnection patterns observed in

![Fig. 7](image-url)

**Fig. 7.** As in Fig. 1, but for the 0–4-day-lagged averaged 250-hPa RWS anomalies (shading: \(10^{-10}\) s\(^{-2}\)). Contours are the climatology (contour interval: 10 m s\(^{-1}\); contours start from 30 m s\(^{-1}\)) of 250-hPa zonal wind. Red box (10°–45°N, 60°E–120°W) in (a) denotes the region for calculation of pattern CC of RWS (T6) and relative amplitude of RWS (T7).
Fig. 1. It is pointed out in Lukens et al. (2017) and Seo and Lee (2017) that the advection of mean absolute vorticity by anomalous divergent wind $-\nabla_x \cdot \nabla_x \xi$ may also be important although it is not the dominant process in our study. The mechanism of PNA-like pattern formation in response to the MJO is discussed in Seo and Lee (2017) based on Rossby wave propagation associated with the third term $-\nabla_x \cdot \nabla_x \xi$ of the RWS. They found that the wave trains emanating from the dipole-pattern RWS regions contain long waves (zonal wavenumbers 1 and 2) that propagate directly from the RWS regions to the PNA regions and shorter waves that first travel along the westerly jet and then emanate at the jet exit region. Process-oriented teleconnection skill metrics $T_6$ and $T_7$ based on the linear Rossby wave theory are defined below:

- **Pattern CC of RWS ($T_6$):** This is calculated similarly to pattern CC ($T_1$) of Z500a but over the red box in Fig. 7a for RWS anomalies.

- **Relative amplitude of RWS ($T_7$):** This is calculated similarly to relative amplitude ($T_2$) of Z500a but over the red box in Fig. 7a for RWS anomalies.

Because of the large contribution of the first ($-\nabla_x \cdot \nabla_x \xi$) and third ($-\nabla_x \cdot \nabla_x \xi$) terms in the RWS, results of $T_6$ and $T_7$ are effectively results of the sum of these two terms that represent the extratropical and tropical component of RWS, respectively (Sardeshmukh and Hoskins 1988). The Taylor diagram constructed with metrics $T_6$ and $T_7$ is shown in Fig. 8 for MJO phases 3 and 7. Most models can generally capture the overall pattern of RWS with a pattern CC greater than 0.8, although none shows a pattern CC greater than 0.8, possibly because the RWS is a noisy quantity consisting of multiple derivatives. For CNRM-CM5 and HadGEM2-CC (CMCC-CESM, GFDL-ESM2M, and NCAR-CAM5), which simulate a reasonable phase 3 (7) Z500a pattern (Fig. 2), the pattern of RWS is also reasonably reproduced (pattern CC of RWS > 0.66). In terms of RWS...
amplitude, a similar feature with MJO teleconnection simulation is found that most models overestimate the amplitude of the RWS (72% of models for phase 3 and 93% for phase 7).

Scatterplots of $T_6$ versus $T_1$ and $T_7$ versus $T_2$ are shown in Fig. 9. The correlation coefficient is 0.52 between $T_1$ and $T_6$ (significant at the 99% confidence level), and 0.32 between $T_2$ and $T_7$ (marginally significant at the 90% confidence level) for an average of MJO phases 3 and 7. When models simulate the RWS pattern ($T_6$) with pattern CC $0.5$, most of them also simulate the teleconnection pattern ($T_1$) with pattern CC $0.5$ (stars in Fig. 9a). The correlation coefficient increases to 0.8 and 0.48 respectively when all MJO phases are considered (both are significant at the 99% confidence level). But if the model with large-amplitude biases ($T_2 > 1.6$) is excluded (dots in Fig. 9b), the correlation between $T_2$ and $T_7$ for all MJO phases decreases from 0.48 to 0.32. The above results indicate that if a model better simulates a realistic RWS pattern, the MJO teleconnection pattern tends to be better captured. But a stronger RWS in a model may not necessarily lead to a stronger MJO teleconnection. This suggests that these process-oriented metrics do not explain teleconnection skill perfectly, especially for the amplitude of MJO teleconnections, partially because they are developed based on the linear Rossby wave dynamics. Other influences like nonlinear wave–mean flow interaction also play roles in MJO teleconnections.

5. Summary and discussion

Because of the lack of a standardized set of diagnostics/metrics, it has been difficult to objectively evaluate the model performance of MJO teleconnections, which obscures necessary model improvements. This study develops a set of teleconnection diagnostics for the purpose of quantitatively examining MJO teleconnection characteristics, objectively evaluating and comparing their simulations in/between models, and tracking progress in simulating these teleconnections. A set of seven teleconnection skill metrics is formulated based on MJO teleconnection diagnostics. Pattern CC ($T_1$) and relative amplitude ($T_2$) reveal the
general model reproduction of MJO teleconnection pattern and amplitude. East–west positions (T3) of teleconnection pattern and their shifts relative to reanalysis are diagnosed. Intraphase pattern consistency (IPC) (T4) helps examine the consistency of teleconnection pattern between individual MJO events for a given phase, which implies the predictability of the MJO-related global impacts. Persistence (T5) helps determine how long the teleconnections persist before decaying, which provides useful information for the MJO-related weather prediction. The pattern CC and relative amplitude of RWS compared to reanalysis (T6 and T7) reveal model fidelity in capturing the overall feature of the RWS that plays an important role in teleconnection formation.

Simulation of MJO teleconnections is quantitatively evaluated in 29 GCMS by applying standardized teleconnection skill metrics developed in this study. The results show that current GCMS generally produce MJO teleconnections with an eastward shift, larger amplitude, and longer persistence than in reanalysis. The simulation success of MJO teleconnections in models varies with MJO phase. Models generally have higher ability in simulating the observed MJO teleconnection for phases 2, 3, 7, and 8 than other phases. Most models produce a westward-shifted teleconnection for phase 3, but an eastward-shifted teleconnection for phase 7. Most models produce less consistent teleconnection patterns than reanalysis between individual MJO events for phases 2, 7, and 8, but overestimate the pattern consistency for the remaining phases. The model performance is also dependent on region. Current GCMS better reproduce MJO teleconnection patterns over the North Pacific than over North America, especially for phases 2 and 6. It is found that the RWS generated by the MJO-induced divergence plays a dominant role in determining model ability at simulating MJO teleconnections. When a model produces a realistic RWS pattern, the MJO teleconnection pattern may also be properly simulated. The robustness of metrics developed in this study is tested by analyzing the reference data over a shorter period (1979–2009 to be comparable to CMIP5 models) and shorter season (December to February). Although slight changes are found in the teleconnection amplitude, it does not influence the main conclusion of this study.

Some teleconnection skill metrics have a mutual relationship with each other. When correlations between teleconnection metrics for an average of all phases are calculated, the relative amplitude of MJO teleconnections (T2) is significantly correlated with IPC (T4) (correlation of 0.82), suggesting that a larger teleconnection amplitude is usually coupled with a higher consistency (smaller variation) of the teleconnection patterns between MJO events. This relationship alludes to the possibility that an overestimate of MJO teleconnection amplitude is partially due to less teleconnection pattern diversity between model MJO events. MJO teleconnections with a larger amplitude are also found to persist longer (correlation between T2 and T5 is 0.53). As the persistence of MJO teleconnections is a key factor to be considered when predicting the MJO-related impacts, the above results suggest that lack of teleconnection pattern diversity can influence the prediction accuracy of the MJO-related impacts.

This study suggests the importance of the RWS in MJO teleconnection simulations. The RWS changes with variations of the MJO-associated flow (Yasui and Watanabe 2010; Seo and Lee 2017; Wang et al. 2018) and in changes in the mean state (Hoskins and Ambrizzi 1993; Ting and Sardeshmukh 1993). Therefore, simulation of MJO teleconnections can be influenced by different characteristics of both the MJO and the basic state. In a companion paper, we will show how simulation of MJO pattern, amplitude, and propagation influences the model performance of MJO teleconnections. The role of location and amplitude of the subtropical westerly jet in MJO teleconnection simulation will also be investigated. The combination of these two studies will provide better understanding of simulation performance of MJO teleconnections. For example, HadGEM2-CC and HadGEM2-AO have the same atmospheric physics and ocean physics and only differ in their inclusion of the terrestrial carbon cycle and ocean biochemistry (Martin et al. 2011). But they show markedly different teleconnection patterns and RWS. Results in the companion paper suggest that this difference is possibly because 1) the MJO has better and more coherent eastward propagation in HadGEM2-CC than in HadGEM2-AO, which leads to more realistic teleconnection pattern in HadGEM2-CC, and 2) the jet shifts southward in HadGEM2-CC but extends eastward in HadGEM2-AO which causes different teleconnection pattern simulations.

Acknowledgments. Constructive and valuable comments from three reviewers are greatly appreciated. We thank WGNE MJO Task Force members for stimulating discussions during the course of this study. We thank Dr. Xianan Jiang for providing the GASS/YoTC model data and Dr. Aneesh Subramanian for providing the ECMWF model data. JW was supported by NSF Grant AGS-1652289, HK was supported by NSF Grant AGS-1652289, NOAA MAPP Grant NA16OAR4310070, and KMA R&D Program Grant KM12018-03110. DK was supported by the U.S. DOE Regional and Global Model.
Analysis program Grant DE-SC0016223, NOAA CVP Grant NA18OAR4310300, and NASA MAP Grant 80NSSC17K0227. EDM was supported by NSF Grants AGS-1441916 and AGS-1841754, NOAA MAPP Grants NA16OAR431064 and NA18OAR4310268, and NOAA CVP Grant NA18OAR4310299.

REFERENCES


