Underestimated MJO variability in CMIP6 models

Phong Vu Viet Le[1], Clement Guilloteau[2], Antonios Mamalakis[3], and Efi Foufoula-Georgiou[1]

[1]University of California, Irvine
[2]UC Irvine
[3]University of California Irvine

November 28, 2022

Abstract

The Madden-Julian Oscillation (MJO) is the leading mode of intra-seasonal climate variability, having profound impacts on a range of weather and climate phenomena. Here, we use a wavelet-based spectral Principal Component Analysis (wsPCA) to evaluate the skill of 20 state-of-the-art CMIP6 models in capturing the magnitude and dynamics of the MJO. The advantages of wsPCA are its ability to focus on desired frequencies and capture each propagative physical mode with one principal component (PC). We show that the MJO contribution to the total intra-seasonal climate variability is substantially underestimated in most CMIP6 models. The joint distribution of the modulus and angular frequency of the complex wavelet PC series associated with MJO is used to rank models relatively to the observations through the Wasserstein distance. Using Hovmöller phase-longitude diagrams, we show that precipitation variability associated with MJO is underestimated in most CMIP6 models for the Amazonia, Southwest Africa, and Maritime Continent.

Underestimated MJO variability in CMIP6 models

Phong V.V. Le[1,2], Clement Guilloteau[1], Antonios Mamalakis[1,3], and Efi Foufoula-Georgiou[1,4]

[1]Department of Civil and Environmental Engineering, University of California Irvine, CA, USA

[2]Faculty of Hydrology Meteorology and Oceanography, University of Science, Vietnam National University, Hanoi, Vietnam

[3]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA

[4]Department of Earth Systems Science, University of California Irvine, CA, USA

Corresponding author: Phong V.V. Le (phongl3@uci.edu); Efi Foufoula-Georgiou (efi@uci.edu)

Key Points:

- A wavelet-based spectral principal component analysis is used to examine CMIP6 models in reproducing the Madden-Julian Oscillation (MJO)
- CMIP6 models capture the average MJO propagation speed but significantly underestimate the MJO contribution to the total intra-seasonal climate variability
- Precipitation variability related to MJO over the Amazonia, Southwest Africa, and Maritime Continent is underestimated in CMIP6 models

The manuscript is under review in Geophysical Research Letters.

Date: *January 9, 2021*

Abstract

The Madden-Julian Oscillation (MJO) is the leading mode of intra-seasonal climate variability, having profound impacts on a wide range of weather and climate phenomena. Here, we use a wavelet-based spectral Principal Component Analysis (wsPCA) to evaluate the skill of 20 state-of-the-art CMIP6 models in capturing the magnitude and dynamics of the MJO. The advantages of wsPCA are its ability to focus on desired frequencies and capture each propagative physical mode with one principal component (PC). We show that the MJO contribution to the total intra-seasonal climate variability is substantially underestimated in most CMIP6 models. The joint distribution of the modulus and angular frequency of the complex wavelet PC series associated with MJO is used to rank models relatively to the observations through the Wasserstein distance. Using Hovmöller phase-longitude diagrams, we also show that precipitation variability associated with MJO is underestimated in most CMIP6 models for the Amazonia, Southwest Africa, and Maritime Continent.

Plain Language Summary

Dominant modes (i.e. coherent spatio-temporal patterns of variability) of the climate system, such as the Madden-Julian Oscillation (MJO), influence a wide range of weather and climate phenomena worldwide. The ability of state-of-the-art climate models to accurately simulate these modes is crucial for advancing our understanding of the climate system and reliably predicting its future trends. The Coupled Model Intercomparison Project phase 6 (CMIP6) will be the foundation for the upcoming Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report. Here, we use a wavelet-based spectral principal component analysis (wsPCA) to quantitatively assess how well historical simulations from 20 CMIP6 models capture MJO as compared to observations. We first show that the MJO magnitude is not reproduced well in most of CMIP6 models. We then reveal that MJO-related precipitation variability in the Amazonia, Southwest Africa, and Maritime Continent is significantly underestimated in many CMIP6 models. Our results highlight the need to better simulate the coupled ocean-atmosphere dynamics in order to improve the representation of MJO in climate models. Moreover, studies using projected states of MJO for assessing future tropical and extratropical impacts should be examined with caution.

1. Introduction

The Madden-Julian Oscillation (MJO) is the dominant mode of intra-seasonal (1-3 months) variability in the tropical atmosphere, characterized by an eastward-moving band of rain clouds (Madden & Julian, 1971, 1972). The MJO interacts with a wide range of tropical weather and climate phenomena, including monsoonal systems (Lorenz & Hartmann, 2006; Taraphdar et al., 2018), tropical cyclone activity (Bessafi & Wheeler, 2006; Klotzbach, 2010; Maloney & Hartmann, 2000), and the El Niño-Southern Oscillation (ENSO) (Hendon et al., 2007; Lau & Waliser, 2012; Takayabu et al., 1999). As a strong tropical heating source, the MJO also exhibits teleconnections to the extratropics affecting regional hydroclimate (Jones et al., 2004; Roxy et al., 2019). Given the planetary-scale climatic impacts of the MJO, the ability of state-of-the-art coupled general circulation models (CGCMs) to accurately capture its magnitude, location and dynamics is of vital importance for subseasonal-to-seasonal prediction (Robertson et al., 2015; Woolnough, 2019) and assessment of future global climate (Meehl, Stocker, et al., 2007).

A number of efforts have focused on assessing CGCMs, primarily those participating in the Coupled Model Intercomparison Projects (CMIP) (Lambert & Boer, 2001; Meehl, Covey, et al., 2007; Taylor et al., 2012) in terms of their ability to properly capture organized spatio-temporal modes across scales. Despite much progress in climate modeling, considerable shortcomings in simulating major modes of climate variability remain, persisting from one model generation to the next (Eyring et al., 2019). For instance, at intra-seasonal timescales, previous generation CGCMs typically exhibit poor representation of MJO dynamics both in amplitude and the eastward propagating pattern (Ahn et al., 2017; Hung et al., 2013; Jiang et al., 2015; Lin et al., 2006; Zhang et al., 2006). The primary factors hypothesized to affect MJO simulations in CGCMs include model resolution and physics, especially the air–sea coupling across multiple spatial scales (Jiang et al., 2020; Zhang, 2005).

The CMIP6 set of models (Eyring et al., 2016) will be the foundation for the Intergovernmental Panel on Climate Change Sixth Assessment Report. Featuring substantial improvements in the physical parameterizations and inclusion of additional Earth system processes, the CMIP6 is expected to provide a rich opportunity to evaluate the aforementioned shortcomings in simulating MJO. Thus far, very few studies have investigated the performance of CMIP6 models in capturing the MJO. Recently, Orbe et al. (2020) analyzed six U.S. climate models participating in CMIP6 and reported improvements in the amplitudes of the MJO-related winds and precipitation compared to the CMIP5. By analyzing 34 models, Ahn et al. (2020) showed that the propagation of MJO over the Maritime Continent in CMIP6 models is more realistic than in the CMIP5. The connection between MJO and the quasi-biennial oscillation (QBO) in CMIP6 models has also been explored (Kim et al., 2020). Nevertheless, there is still a general lack of understanding of the MJO representation in the state-of-the-art climate models.

Two classical ways of identifying MJO dynamics is through a space-time spectral analysis (STSA) (Hendon & Wheeler, 2008; Kiladis et al., 2005; Wheeler & Kiladis, 1999) and an empirical orthogonal function (EOF) analysis (Lo & Hendon, 2000; Maloney & Hartmann, 1998; Waliser et al., 2003; Wheeler & Hendon, 2004). While the STSA requires the selection of windows in the wavenumber-frequency domain containing the signal of interest, EOF-based methods require bandpass filtering and seasonal partitioning to isolate the intra-seasonal components of the data. The frequency-domain (spectral) variants of EOF analysis (Hannachi et al., 2007; Schmidt et al., 2019) rely on the eigen-decomposition of the Fourier cross-spectral matrix (CSM), which offers the possibility to look for modes in specific frequency bands and handle propagating effects.

We propose to use the wavelet-based spectral principal component analysis (wsPCA), which is based on the eigen-decomposition of the CSM computed through a continuous complex analytic wavelet transform (Guilloteau et al., 2020). The wsPCA allows robust estimation of the CSM and seamlessly removes trends in the data without any pre-processing. The complex wavelet principal component (wPC) time series resulting from the wsPCA are characterized by their instantaneous magnitude and phase, which are useful quantities to describe the temporal evolution of dynamical climatic modes.

In this study, we analyze global precipitation (PPT) and outgoing longwave radiation (OLR) daily time series to assess MJO variability in observations, reanalysis, and as simulated by 20 CMIP6 models under historical forcing. Particularly, we first demonstrate the use of the wsPCA to robustly extract the spatio-temporal patterns of the MJO. We then evaluate the dynamics of MJO simulated in CMIP6 models by comparing them to those inferred by the observations. Finally, we evaluate MJO-related precipitation variability as simulated by CMIP6 models in the Amazonia, Southwest Africa, and Maritime Continent. The rest of this paper is organized as follows. Section 2 describes the data and methodology used. Section 3 presents the main results of this study followed by a summary in Section 4.

2. Materials and Methods

2.1. Data

CMIP6 Models: Daily-averaged outputs, including PPT and OLR, from historical simulations of 20 CMIP6 models (Table S1) during the period 1983-2014 are examined. Model output is taken only from the first ensemble member (r1i1p1f1) of each model, which uses the same observed evolution of forcing in the 20th century. All model outputs are bilinearly interpolated to a common equal-area scalable earth (EASE) grid of approximately 220 km resolution (Brodzik et al., 2014). Anomaly time series of each field are obtained by removing the climatic mean of each day of the year (DOY) from the raw data. The climatic mean is calculated as the average over the study period of the 15-day period centered on each DOY.

Observations and Reanalysis: For observations, we employ the daily global interpolated OLR obtained from the National Center for Atmospheric Research. Daily PPT is obtained from the PERSIANN-CDR database (Ashouri et al., 2015). For reanalysis, daily-averaged fields of the above variables are obtained from the ERA5 datasets (Hersbach et al., 2020). Observations and reanalysis datasets are obtained over the same period (1983-2014) and interpolated onto the same EASE grid as CMIP6 models for comparison. Moreover, daily observed precipitation obtained from other datasets (TRMM, IMERG, GPCP, and CMORPH) is used for comparison with the PERSIANN-CDR (Table S2).

2.2. Methodology

We use the wsPCA (Guilloteau et al., 2020) to identify organized spatio-temporal modes of variability within the MJO timescales. The wsPCA relies on the estimation of the CSM between time series at different locations using the Morlet continuous wavelet transform (CWT) and the extraction of its eigenvectors in various frequency bands. Consider a dataset consisting of L time-ordered snapshots of a variable at N gridded locations, $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_N)^T \in \mathbb{R}^{N \times L}$. The i^{th} row of $\boldsymbol{X}$, that is $\boldsymbol{x}_i^T = \left(x_{i,t_1}, \dots, x_{i,t_L}\right) \in \mathbb{R}^L$, represents the time series at the i^{th} location. Meanwhile, the j^{th} column $\left(x_{1,t_j}, \dots, x_{N,t_j}\right)^T \in \mathbb{R}^N$ represents the vectorized snapshot at time t_j. The CWT of $\boldsymbol{x}_i$ is defined as $w_i(\nu,t) = \frac{1}{\sqrt{\nu}}\int_{-\infty}^{+\infty} x_i(\tau)\Psi^*\left(\frac{\tau-t}{\nu}\right)d\tau$, where $\Psi^*(\mathrm{t})$ is the complex conjugate of the Morlet wavelet defined in its simplified form as $\Psi(t) \approx \pi^{-\frac{1}{4}}e^{i2\pi f_0 t}e^{\frac{-t^2}{2}}$, ν is the scale parameter, and f_0 is the central frequency of the Morlet wavelet (Addison, 2002). We choose $f_0 = \sqrt{\frac{1}{2\ln 2}}$, often used in practice when accurate time localization of the wavelet transform is sought. At the scale ν, corresponding to the Fourier frequency $f = \frac{f_0}{\nu}$, the CWT of all $\boldsymbol{x}_i$ time series can be arranged into a matrix of wavelet coefficients:

$$\mathbf{W}_f = \begin{bmatrix} w_1(f,t_1) & \cdots & w_1(f,t_L) \\ \vdots & \ddots & \vdots \\ w_N(f,t_1) & \cdots & w_N(f,t_L) \end{bmatrix} \in \mathbb{C}^{N\times L} \tag{1}$$

The empirical CSM at frequency f is then computed as $\boldsymbol{S}_f = \frac{1}{L-1}\boldsymbol{W}_f\boldsymbol{W}_f' \in \mathbb{C}^{N\times N}$ and its eigen-decomposition constitutes the wsPCA:

$$\boldsymbol{S}_f\boldsymbol{U}_f = \boldsymbol{U}_f\boldsymbol{\Lambda}_f \tag{2}$$

where $\boldsymbol{W}_f'$ denotes the conjugate transpose of $\boldsymbol{W}_f$; $\boldsymbol{\Lambda}_f \in \mathbb{R}^{N\times N}$ is the diagonal matrix of the eigenvalues ($\lambda_{i,f}^2 \in \mathbb{R}_+, i = 1,..,N$) and $\boldsymbol{U}_f \in \mathbb{C}^{N\times N}$ is the matrix of column eigenvectors ($\boldsymbol{u}_{i,f} \in \mathbb{C}^N,\ i = 1,..,N$) of $\boldsymbol{S}_f$, respectively. We note here that $tr(\boldsymbol{\Lambda}_f) = tr(\boldsymbol{S}_f)$. If the interest is in extracting modes which span a desired frequency band, $\boldsymbol{S}_f$ can be integrated over that frequency band before performing the eigen-decomposition. Here we define the MJO band-integrated ($4 \leq f \leq 12$ cpy) CSM as:

$$\boldsymbol{S}_{\mathrm{MJO}} = \int_4^{12}\frac{\boldsymbol{S}_f}{f}df \in \mathbb{C}^{N\times N} \tag{3}$$

The diagonal matrix of eigenvalues and the matrix of column eigenvector of $\boldsymbol{S}_{\mathrm{MJO}}$ are $\boldsymbol{\Lambda}_{\mathrm{MJO}} = \mathrm{diag}\left(\lambda_{i,\mathrm{MJO}}^2 \in \mathbb{R}_+, i = 1,..,N\right) \in \mathbb{R}^{N\times N}$ and $\boldsymbol{U}_{\mathrm{MJO}} = \left(\boldsymbol{u}_{i,\mathrm{MJO}} \in \mathbb{C}^N, i = 1,..,N\right) \in \mathbb{C}^{N\times N}$, respectively. For unique solution of the eigen-decomposition in $\mathbb{C}^{N\times N}$, we impose unit L2-norm for each eigenvector and a zero argument to the scalar element with the largest modulus in each eigenvector. The wPC series of wavelet coefficients associated with the eigenvector $\boldsymbol{u}_{i,\mathrm{MJO}}$ at frequency f is calculated as:

$$\boldsymbol{\kappa}_{i,f} = \boldsymbol{W}_f'\boldsymbol{u}_{i,\mathrm{MJO}} \in \mathbb{C}^L \tag{4}$$

and the MJO band-integrated complex wPC series is then defined as:

$$\boldsymbol{\kappa}_{i,\mathrm{MJO}} = \int_4^{12}\frac{\boldsymbol{\kappa}_{i,f}}{f}df \in \mathbb{C}^L \tag{5}$$

By design, the wsPCA separates modes of variability having distinct frequency supports. The eigenvectors are represented as maps of complex loading coefficients whose argument characterizes the relative phase shift (i.e. time delays) of the wPC time series between different geographical locations, allowing wsPCA to handle potential non-synchronicity between the time series $\boldsymbol{x}_i$ and propagation effects. The use of the Morlet wavelet in particular allows optimal time-frequency localization and insensitivity to linear trends (Guilloteau et al., 2020).

2.3. MJO Diagnostics

***Spectral energy within the MJO frequency band*:** The energy distribution of the analyzed signal across frequencies is described by the wavelet power spectral density (PSD):

$$PSD(f) = \frac{f_0}{f} \times \frac{1}{N(L-1)} \sum_{n=1}^{N} \sum_{l=1}^{L} |w_n(f, t_l)|^2 = \frac{f_0}{f} \times \frac{tr(\mathbf{S}_f)}{N} \tag{6}$$

and the energy contained within the MJO frequency band is given by:

$$\overline{PSD}_{\mathrm{MJO}} = \frac{1}{12-4} \int_4^{12} PSD(f)df \tag{7}$$

The fraction of spectral power (FSP) contributed by the first wPC at frequency f is:

$$FSP_1(f) = \frac{\lambda_{1,f}^2}{tr(\boldsymbol{\Lambda}_f)} \tag{8}$$

and over the MJO frequency band is:

$$\overline{FSP}_{1,\mathrm{MJO}} = \frac{\lambda_{1,\mathrm{MJO}}^2}{tr(\boldsymbol{\Lambda}_{\mathrm{MJO}})} \tag{9}$$

By comparing $PSD(f)$, $\overline{PSD}_{\mathrm{MJO}}$, and $FSP_1(f), \overline{FSP}_{1,\mathrm{MJO}}$ computed from observations and CMIP6 model outputs, an assessment can be made of the ability of CMIP6 models to reproduce the total variance within intra-seasonal time scales and to model the MJO mode with the right contribution to the total intra-seasonal variance.

Patterns and propagation speed of MJO: Unlike classical PCA for which two eigenvectors and corresponding PCs are needed to capture the MJO (Wheeler & Hendon, 2004), for wsPCA only the first complex eigenvector $\boldsymbol{u}_{1,\mathrm{MJO}}$ and the first complex wPC series $\boldsymbol{\kappa}_{1,\mathrm{MJO}}$ are needed. Specifically, the maps of $|\boldsymbol{u}_{1,\mathrm{MJO}}|$ and $arg(\boldsymbol{u}_{1,\mathrm{MJO}})$ capture the magnitude and phase, respectively, of the MJO pattern. To compare MJO patterns between observations and models, the complex correlation coefficient is calculated as:

$$\rho^{\boldsymbol{u}_{1,\mathrm{MJO}}} = \frac{\boldsymbol{u}_{1,\mathrm{MJO}}^{\mathrm{obs}} \cdot \boldsymbol{u}_{1,\mathrm{MJO}}^{*\mathrm{mod}}}{\left|\boldsymbol{u}_{1,\mathrm{MJO}}^{\mathrm{obs}}\right|_2 \cdot \left|\boldsymbol{u}_{1,\mathrm{MJO}}^{*\mathrm{mod}}\right|_2} \in \mathbb{C} \tag{10}$$

where $\boldsymbol{u}^*$ is the complex conjugate and $|\boldsymbol{u}|_2$ is the L2-norm of $\boldsymbol{u}$, respectively.

The wPC1 series $\boldsymbol{\kappa}_{1,\mathrm{MJO}}$ is used to quantitatively diagnose the magnitude and propagation dynamics of MJO. In the two-dimensional complex space defined by the real and imaginary parts of $\boldsymbol{\kappa}_{1,\mathrm{MJO}}$, we form a wsPCA MJO index akin to the previous indices (Kiladis et al., 2014; Wheeler & Hendon, 2004). Based on the variable used (i.e. OLR or PPT), we designate this index as the wsPCA-based OLR MJO index (wOMI) or the wsPCA-based PPT MJO index (wPMI), respectively. To allow comparison between models and observations, the wPC1 series ($\boldsymbol{\kappa}_{1,\mathrm{MJO}}$) of each model are normalized by the standard deviation of that obtained from observations. Specifically, we work with:

$$\widehat{\boldsymbol{\kappa}}_{1,\mathrm{MJO}} = \frac{\boldsymbol{\kappa}_{1,\mathrm{MJO}}}{\lambda_{1,\mathrm{MJO}}^{\mathrm{obs}}/\sqrt{2}} \in \mathbb{C}^L \tag{11}$$

Note here that $\lambda_{1,\mathrm{MJO}}^{\mathrm{obs}} = \sqrt{2}\sigma_{\mathfrak{R}(\boldsymbol{\kappa}_{1,\mathrm{MJO}}^{\mathrm{obs}})} = \sqrt{2}\sigma_{\mathfrak{I}(\boldsymbol{\kappa}_{1,\mathrm{MJO}}^{\mathrm{obs}})}$. At any time t, the modulus and argument of $\hat{\kappa}_{1,\mathrm{MJO}}(t)$ define the instantaneous intensity and phase of the MJO, respectively. Since $\arg(\hat{\kappa}_{1,\mathrm{MJO}}(t)) \in [0, 2\pi]$, the eight traditional phases of MJO correspond to angular sectors each spanning over $\frac{\pi}{4}$ radians in the complex plane. The angular frequency $\omega_{1,\mathrm{MJO}}(t) = \Delta\arg(\hat{\kappa}_{1,\mathrm{MJO}}(t))/\Delta t$ represents the instantaneous propagation speed of MJO. In the rest of the paper, we only discuss wOMI as the primary MJO index as we will show later that the MJO pattern is reproduced more accurately by models using OLR than PPT.

In order to compare models to observations in term of their ability to capture both the magnitude and instantaneous propagating speed of MJO, we form the bivariate probability density function (PDF) of $|\hat{\kappa}_{1,\mathrm{MJO}}(t)|$ and $\omega_{1,\mathrm{MJO}}(t)$ for models and observations and compare them using a distance metric. We choose the Wasserstein (or Earth Mover's) distance (WD) (Kantorovich, 2006; Rubner et al., 2000) which is a nonlinear metric defined as the minimal amount of work, or optimal mass transport (Villani, 2008), needed to transform a discrete probability distribution to another. This metric allows to rank CMIP6 models based on their skill to reproduce the magnitude and dynamics of MJO.

3. Results

Figure 1 (a1,b1 – top row) shows the power spectral density (PSD) of PPT and OLR for the observations, reanalysis, and 20 CMIP6 models. The PSD indicates that much of the energy of both variables is concentrated within the ENSO timescale (2-7 years), highlighting the dominant influence of this interannual variability mode on the climate system. At intra-seasonal timescales (1-3 months), the PSD obtained from the multi-model ensemble (MME) mean of PPT is comparable to that from the observations, whereas the PSD of OLR in CMIP6 models is generally higher than that of observations and reanalysis.

The fraction of power spectra contributed by wPC1 (FPS_1) is presented in Figure 1 (a2,b2 – bottom row). We note that FPS_1 is high at low-frequencies (interannual and lower frequencies) for both PPT and OLR (40-70% of the spectral power is contributed by wPC1 within the ENSO timescale). At intra-seasonal timescales, the FPS_1 of OLR ranges from 4-18% and that of PPT is slightly lower. Nevertheless, while the observations and reanalysis show a well-defined peak in FPS_1 within the MJO timescale (reaching up to 18%) indicating a coherent signal of MJO (inset plots), many models substantially underestimate FPS_1 within the MJO timescale and show no well-defined peak. This result implies that, although CMIP6 models do not lack total variance within intra-seasonal timescales (Figure 1 – top panels), they fail to properly model the MJO mode of variability. Comparisons of the PSD and FPS_1 among observed precipitation products are further shown in Figure S1.

The spatial pattern of $|\boldsymbol{u}_{1,\mathrm{MJO}}|$ computed from observed OLR shows a coherent spatio-temporal mode spanning from the tropical Indian Ocean to the Western Pacific (Figure 2a) and the pattern of $\arg(\boldsymbol{u}_{1,\mathrm{MJO}})$ shown in Figure 2b clearly indicates eastward propagation of MJO, demonstrating the robustness of the wsPCA to identify the MJO as the dominant mode in the 4-12 cpy frequency band. The spatial patterns of the magnitude and argument of $\boldsymbol{u}_{1,\mathrm{MJO}}$ of OLR and PPT for all CMIP6 models, reanalysis, and observations are presented in Figures S2-S5 for comparison. Furthermore, the lag-longitude diagrams of the reconstructed OLR and PPT anomalies within the MJO timescale are shown in Figure S6. It can be seen that the average eastward propagation speed of MJO as estimated from the observations, reanalysis, and a large number (13/20) of the models is about 5 m/s. Our results suggest that the majority of CMIP6 models are able to capture well the average propagation speed of MJO which is consistent with previous studies (Ahn et al., 2020; Orbe et al., 2020). Nevertheless, many models underestimate the MJO variability as reflected by the lower values of the normalized magnitudes $\frac{\lambda_{1,\mathrm{MJO}}\times\sqrt{N}}{\sqrt{tr(\boldsymbol{\Lambda}_{\mathrm{MJO}})}}\,|\boldsymbol{u}_{1,\mathrm{MJO}}|$ compared to those of the observations (Figures S2&S4). Moreover, the magnitude of the wPC1 time series $\boldsymbol{\kappa}_{1,f}$ across frequencies is presented in Figure 2c, showing considerable interannual variability in MJO activity. Figure 2d shows the trajectory in the complex plane of the daily wOMI obtained from observations during the study period. The trajectories of daily wOMI and wPMI obtained from all datasets are further presented in Figures S7-S8.

Comparison of the $\overline{PSD}_{\mathrm{MJO}}$ and $\overline{FSP}_{1,\mathrm{MJO}}$ for PPT and OLR is shown in Figure 3a-b, respectively. For PPT, while the CMIP6 models show a spread of the $\overline{PSD}_{\mathrm{MJO}}$ above and below the value of the observations indicating no systematic bias, the $\overline{FPS}_{1,\mathrm{MJO}}$ estimated from the models is consistently smaller than that from the observations, indicating that the models systematically underestimate the MJO variability. For OLR, most CMIP6 models exhibit higher $\overline{PSD}_{\mathrm{MJO}}$ than observations (except model IPSL-CM6A-LR(13); Figure 3b); however all models show lower $\overline{FPS}_{1,\mathrm{MJO}}$ than that of the observations, further confirming that CMIP6 models consistently underestimate the contribution of the MJO to intra-seasonal climate variability. For both variables, $\overline{PSD}_{\mathrm{MJO}}$ of the reanalysis is slightly higher than in the observations, but the $\overline{FPS}_{1,\mathrm{MJO}}$ is lower. The scatter plot of the modulus of the complex pattern correlation coefficients $|\rho_{OLR}^{\boldsymbol{u}_{1,\mathrm{MJO}}}|$ and $|\rho_{PPT}^{\boldsymbol{u}_{1,\mathrm{MJO}}}|$ as defined in Equation (10) is shown in Figure 3c. Most of the models show correlations in the range of 0.6-0.85 for both variables, confirming that, for all models, the first dynamical mode extracted by the wsPCA in the 4-12 cpy frequency band is actually the MJO, and indicating quite good agreement of the modeled MJO patterns to the observed ones. We note however that the complex pattern correlation only indicates agreement between the unit-norm first complex eigenvectors (Equation 10) and does not take into account the discrepancy between their corresponding eigenvalues (variance explained), a discrepancy that has been separately assessed in Figure 3a-b. Two models showing very low values of $|\rho_{OLR}^{\boldsymbol{u}_{1,\mathrm{MJO}}}|$ and $|\rho_{PPT}^{\boldsymbol{u}_{1,\mathrm{MJO}}}|$ are the IPSL-CM6A-LR(13) and CanEMS(5). Finally, most of the scatter points are below the $1{:}1$ line, implying that CMIP6 models generally reproduce more accurately the patterns of OLR than PPT.

Figure 4a compares the relationship of the magnitude and propagation speed of MJO for models, reanalysis, and observations for all days during 1983-2014 (these can be seen as joint PDFs). Note that the normalized wPC series $\hat{\boldsymbol{\kappa}}_{1,\mathrm{MJO}}$ (see Section 2.3) are plotted to allow comparison between models and observations. We find that while the average propagation speed (mean of the PDF of $\omega_{1,\mathrm{MJO}}(t)$) is quite similar among all models ($0.1-0.13$ rad/day, equivalent cycles of $60-48$ days), CMIP6 models underestimate the magnitude $|\hat{\boldsymbol{\kappa}}_{1,\mathrm{MJO}}|$ of the MJO mode. The marginal PDFs of $|\hat{\kappa}_{1,\mathrm{MJO}}(t)|$ and $\omega_{1,\mathrm{MJO}}(t)$ for all datasets are shown in Figure 4b further demonstrating that most of CMIP6 models capture the MJO propagation speed but underestimate the amplitude of MJO compared to the observations (as also shown in Figure 4a). Moreover, Figure 4c shows the ranked WD between the joint PDFs of $\left|\hat{\kappa}_{1,\mathrm{MJO}}(t)\right|$ and $\omega_{1,\mathrm{MJO}}(t)$ inferred by the observations (reference) and those obtained from the reanalysis (red bar) and CMIP6 models (grey bars). The smaller the values of the WD, the better the performance of a model to reproduce the observed MJO magnitude and speed. Relatively good models that show the smallest WD values include the NESM3(19) and SAM0-UNICON(20) that are consistent with recent reports on the improvements of MJO simulations in these models (Shin & Park, 2020; Yang et al., 2020).

In Figure 5, we evaluate the impact of the MJO on precipitation over the Amazonia, Southwest Africa, and Maritime Continent regions. The Hovmöller phase-longitude diagrams of PPT anomalies show that the MME mean produces smaller MJO-related precipitation variability compared to the observations during all eight MJO phases and in all regions (Figure 5a). Details of the Hovmöller diagrams for each model in each region are further shown in Figures S9-S11. These diagrams suggest that a large number of CMIP6 models underestimate the MJO signal to regional precipitation compared to the observations. Among the three regions, the models produce the most realistic precipitation variability in the Maritime Continent where MJO activity is the greatest. Furthermore, the scatter plots of the WD and correlation coefficients of the Hovmöller diagrams of PPT (ρ_{PPT}^{Hov}) between models and observations for each region are presented in Figure 5b-d. It can be seen that models showing good performance in reproducing the MJO magnitude (i.e. models with low WD value) also tend to exhibit higher correlation of ρ_{PPT}^{Hov} with observations and larger MJO-related precipitation variability in the Amazonia and Maritime Continent, while this tendency is not observed in Southwestern Africa. Our results suggest that CMIP6 models which underestimate MJO magnitude also reproduce weak MJO teleconnections to regional precipitation.

4. Conclusions

In this study, we have analyzed historical simulations of 20 CMIP6 models to assess their ability to capture the space-time dynamics of MJO. For the first time, we applied the wsPCA to extract the pattern, magnitude, and eastward propagation of MJO from daily PPT and OLR. The key advantage of wsPCA compared to other PCA methods is that the cross-spectral matrix (CSM) between time series across locations is estimated using a complex CWT enabling robust estimation of the CSM in any desired frequency band. Moreover, the wsPCA is non-parametric and simple to implement compared to nonlinear dimensionality reduction approaches, such as the nonlinear Laplacian spectral analysis (NLSA) (Giannakis & Majda, 2012), which significantly facilitates the extraction of dynamical modes from a large number of models. We defined the wsPCA MJO indices (wOMI and wMPI) based on the real and imaginary parts of the MJO band-integrated (4-12 cpy) complex wPC1 series to evaluate the magnitude and phase of the MJO mode at the daily scale and compare models with observations. We then investigated the influence of MJO to precipitation variability in CMIP6 models over three different regions.

The analysis herein showed that most CMIP6 models are able to realistically capture the eastward propagation of MJO as also reported in recent studies (Ahn et al., 2020; Orbe et al., 2020). However, the simulation of the MJO magnitude in CMIP6 remains a challenging problem. We demonstrated that although CMIP6 models exhibit enough spectral power or total variance within the intra-seasonal timescales as compared to observations, they tend to underestimate the variability contributed by the MJO mode. Furthermore, we showed that precipitation variability associated with the MJO is underestimated in the CMIP6 models in the Amazonia, Southwest Africa and Maritime Continent. Our results highlight the need to better simulate the coupled ocean-atmosphere dynamics in climate models to improve the MJO representation and MJO-driven tropical and extratropical rainfall.

Acknowledgments, Samples, and Data

The authors acknowledge support provided by the National Science Foundation (NSF) under the TRIPODS+X program (grant DMS-1839336) and the EAGER program (grant ECCS-1839441), as well as by NASA's Global Precipitation Measurement program (grant 80NSSC19K0684). Upon request, the code that supports the findings of this paper can be provided by the corresponding authors.

The authors acknowledge the FAIR data policy. The CMIP6 data set is available at https://esgf-node.llnl.gov/projects/cmip6. The ECMWF ERA5 reanalysis data set was downloaded from https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5. The PERSIANN-CDR precipitation data was downloaded from https://chrsdata.eng.uci.edu/. The interpolated OLR data of NCAR and the GPCP v2.3 precipitation data were provided by the NOAA/ESRL PSD, Boulder, CO, USA (https://psl.noaa.gov/data/gridded/index.html). The TRMM (34B2) and IMERG precipitation data were provided by the NASA's Precipitation Measurement Missions (https://gpm.nasa.gov/data/directory). The CMORPH precipitation data was provided by the NOAA/CPC (https://www.cpc.ncep.noaa.gov/products/janowiak/cmorph_description.html).

References

Addison, P. S. (2002). *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*: Taylor & Francis.

Ahn, M.-S., Kim, D., Kang, D., Lee, J., Sperber, K. R., Gleckler, P. J., et al. (2020). MJO Propagation Across the Maritime Continent: Are CMIP6 Models Better Than CMIP5 Models? *Geophysical Research Letters, 47*(11), e2020GL087250.

Ahn, M.-S., Kim, D., Sperber, K. R., Kang, I.-S., Maloney, E., Waliser, D., et al. (2017). MJO simulation in CMIP5 climate models: MJO skill metrics and process-oriented diagnosis. *Climate Dynamics, 49*(11), 4023-4045.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., et al. (2015). PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bulletin of the American Meteorological Society, 96*(1), 69-83.

Bessafi, M., & Wheeler, M. C. (2006). Modulation of South Indian Ocean Tropical Cyclones by the Madden–Julian Oscillation and Convectively Coupled Equatorial Waves. *Monthly Weather Review, 134*(2), 638-656. https://doi.org/10.1175/MWR3087.1

Brodzik, M. J., Billingsley, B., Haran, T., Raup, B., & Savoie, M. H. (2014). Correction: Brodzik, M.J., et al. EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets. ISPRS International Journal of Geo-Information 2012, 1, 32–45. *ISPRS International Journal of Geo-Information, 3*(3), 1154-1156.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development, 9*(5), 1937-1958.

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., et al. (2019). Taking climate model evaluation to the next level. *Nature Climate Change, 9*(2), 102-110.

Giannakis, D., & Majda, A. J. (2012). Nonlinear Laplacian spectral analysis for time series with intermittency and low-frequency variability. *Proceedings of the National Academy of Sciences, 109*(7), 2222-2227. https://www.pnas.org/content/pnas/109/7/2222.full.pdf

Guilloteau, C., Mamalakis, A., Vulis, L., Le, P. V. V., Georgiou, T. T., & Foufoula-Georgiou, E. (2020). Rotated spectral principal component analysis (rsPCA) for identifying dynamical modes of variability in climate systems. *Journal of Climate*, 1-59. https://doi.org/10.1175/JCLI-D-20-0266.1

Hannachi, A., Jolliffe, I. T., & Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology, 27*(9), 1119-1152.

Hendon, H. H., & Wheeler, M. C. (2008). Some Space–Time Spectral Analyses of Tropical Convection and Planetary-Scale Waves. *Journal of the Atmospheric Sciences, 65*(9), 2936-2948. https://doi.org/10.1175/2008JAS2675.1

Hendon, H. H., Wheeler, M. C., & Zhang, C. (2007). Seasonal Dependence of the MJO–ENSO Relationship. *Journal of Climate, 20*(3), 531-543. https://doi.org/10.1175/JCLI4003.1

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society, 146*(730), 1999-2049.

Hung, M.-P., Lin, J.-L., Wang, W., Kim, D., Shinoda, T., & Weaver, S. J. (2013). MJO and Convectively Coupled Equatorial Waves Simulated by CMIP5 Climate Models. *Journal of Climate, 26*(17), 6185-6214.

Jiang, X., Adames, Á. F., Kim, D., Maloney, E. D., Lin, H., Kim, H., et al. (2020). Fifty Years of Research on the Madden-Julian Oscillation: Recent Progress, Challenges, and Perspectives. *Journal of Geophysical Research: Atmospheres, 125*(17), e2019JD030911. https://doi.org/10.1029/2019JD030911. https://doi.org/10.1029/2019JD030911

Jiang, X., Waliser, D. E., Xavier, P. K., Petch, J., Klingaman, N. P., Woolnough, S. J., et al. (2015). Vertical structure and physical processes of the Madden-Julian oscillation: Exploring key model physics in climate simulations. *Journal of Geophysical Research: Atmospheres, 120*(10), 4718-4748. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD022375

Jones, C., Waliser, D. E., Lau, K. M., & Stern, W. (2004). Global Occurrences of Extreme Precipitation and the Madden–Julian Oscillation: Observations and Predictability. *Journal of Climate, 17*(23), 4575-4589. https://doi.org/10.1175/3238.1

Kantorovich, L. V. (2006). On the Translocation of Masses. *Journal of Mathematical Sciences, 133*(4), 1381-1382. https://doi.org/10.1007/s10958-006-0049-2

Kiladis, G. N., Dias, J., Straub, K. H., Wheeler, M. C., Tulich, S. N., Kikuchi, K., et al. (2014). A Comparison of OLR and Circulation-Based Indices for Tracking the MJO. *Monthly Weather Review, 142*(5), 1697-1715. https://doi.org/10.1175/MWR-D-13-00301.1

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.

Kim, H., Caron, J. M., Richter, J. H., & Simpson, I. R. (2020). The Lack of QBO-MJO Connection in CMIP6 Models. *Geophysical Research Letters, 47*(11), e2020GL087295. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL087295

Klotzbach, P. J. (2010). On the Madden–Julian Oscillation–Atlantic Hurricane Relationship. *Journal of Climate, 23*(2), 282-293. https://doi.org/10.1175/2009JCLI2978.1

Lambert, S. J., & Boer, G. J. (2001). CMIP1 evaluation and intercomparison of coupled climate models. *Climate Dynamics, 17*(2), 83-106.

Lau, W. K., & Waliser, D. E. (2012). El Nino southern oscillation connection. In *Intraseasonal Variability in the Atmosphere-Ocean Climate System* (pp. 297-334): Springer.

Lin, J.-L., Kiladis, G. N., Mapes, B. E., Weickmann, K. M., Sperber, K. R., Lin, W., et al. (2006). Tropical Intraseasonal Variability in 14 IPCC AR4 Climate Models. Part I: Convective Signals. *Journal of Climate, 19*(12), 2665-2690. https://doi.org/10.1175/JCLI3735.1

Lo, F., & Hendon, H. H. (2000). Empirical Extended-Range Prediction of the Madden–Julian Oscillation. *Monthly Weather Review, 128*(7), 2528-2543. https://doi.org/10.1175/1520-0493(2000)128<2528:EERPOT>2.0.CO;2

Lorenz, D. J., & Hartmann, D. L. (2006). The Effect of the MJO on the North American Monsoon*. *Journal of Climate, 19*(3), 333-343. https://doi.org/10.1175/JCLI3684.1

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.

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.

Maloney, E. D., & Hartmann, D. L. (1998). Frictional Moisture Convergence in a Composite Life Cycle of the Madden–Julian Oscillation. *Journal of Climate, 11*(9), 2387-2403. https://doi.org/10.1175/1520-0442(1998)011<2387:FMCIAC>2.0.CO;2

Maloney, E. D., & Hartmann, D. L. (2000). Modulation of Hurricane Activity in the Gulf of Mexico by the Madden-Julian Oscillation. *Science, 287*(5460), 2002-2004. https://science.sciencemag.org/content/sci/287/5460/2002.full.pdf

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., et al. (2007). THE WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research. *Bulletin of the American Meteorological Society, 88*(9), 1383-1394.

Meehl, G. A., Stocker, T. F., Collins, W. D., Friedlingstein, P., Gaye, A. T., Gregory, J. M., et al. (2007). Global climate projections. In *Climate Change 2007: The physical science basis* (pp. 747-846). United Kingdom: Cambridge University Press.

Orbe, C., Van Roekel, L., Adames, Á. F., Dezfuli, A., Fasullo, J., Gleckler, P. J., et al. (2020). Representation of Modes of Variability in Six U.S. Climate Models. *Journal of Climate, 33*(17), 7591-7617.

Robertson, A. W., Kumar, A., Peña, M., & Vitart, F. (2015). Improving and Promoting Subseasonal to Seasonal Prediction. *Bulletin of the American Meteorological Society, 96*(3), ES49-ES53.

Roxy, M. K., Dasgupta, P., McPhaden, M. J., Suematsu, T., Zhang, C., & Kim, D. (2019). Twofold expansion of the Indo-Pacific warm pool warps the MJO life cycle. *Nature, 575*(7784), 647-651. https://doi.org/10.1038/s41586-019-1764-4

Rubner, Y., Tomasi, C., & Guibas, L. J. (2000). The Earth Mover's Distance as a Metric for Image Retrieval. *International Journal of Computer Vision, 40*(2), 99-121. https://doi.org/10.1023/A:1026543900054

Schmidt, O. T., Mengaldo, G., Balsamo, G., & Wedi, N. P. (2019). Spectral Empirical Orthogonal Function Analysis of Weather and Climate Data. *Monthly Weather Review, 147*(8), 2979-2995. https://doi.org/10.1175/MWR-D-18-0337.1

Shin, J., & Park, S. (2020). Impacts of ENSO and Madden–Julian oscillation on the genesis of tropical cyclones simulated by general circulation models and compared to observations. *Environmental Research Letters, 15*(3), 034046. http://dx.doi.org/10.1088/1748-9326/ab7466

Takayabu, Y. N., Iguchi, T., Kachi, M., Shibata, A., & Kanzawa, H. (1999). Abrupt termination of the 1997–98 El Niño in response to a Madden–Julian oscillation. *Nature, 402*(6759), 279-282. https://doi.org/10.1038/46254

Taraphdar, S., Zhang, F., Leung, L. R., Chen, X., & Pauluis, O. M. (2018). MJO Affects the Monsoon Onset Timing Over the Indian Region. *Geophysical Research Letters, 45*(18), 10011-10018. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078804

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society, 93*(4), 485-498.

Villani, C. (2008). *Optimal transport: old and new* (Vol. 338): Springer Science & Business Media.

Waliser, D. E., Lau, K. M., 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.

Wheeler, M. C., & Kiladis, G. N. (1999). Convectively Coupled Equatorial Waves: Analysis of Clouds and Temperature in the Wavenumber–Frequency Domain. *Journal of the Atmospheric Sciences, 56*(3), 374-399.

Woolnough, S. J. (2019). Chapter 5 - The Madden-Julian Oscillation. In A. W. Robertson & F. Vitart (Eds.), *Sub-Seasonal to Seasonal Prediction* (pp. 93-117): Elsevier.

Yang, Y.-M., Wang, B., Cao, J., Ma, L., & Li, J. (2020). Improved historical simulation by enhancing moist physical parameterizations in the climate system model NESM3.0. *Climate Dynamics, 54*(7), 3819-3840. https://doi.org/10.1007/s00382-020-05209-2

Zhang, C. (2005). Madden-Julian Oscillation. *Reviews of Geophysics, 43*(2).

Zhang, C., Dong, M., Gualdi, S., Hendon, H. H., Maloney, E. D., Marshall, A., et al. (2006). Simulations of the Madden–Julian oscillation in four pairs of coupled and uncoupled global models. *Climate Dynamics, 27*(6), 573-592. https://doi.org/10.1007/s00382-006-0148-2

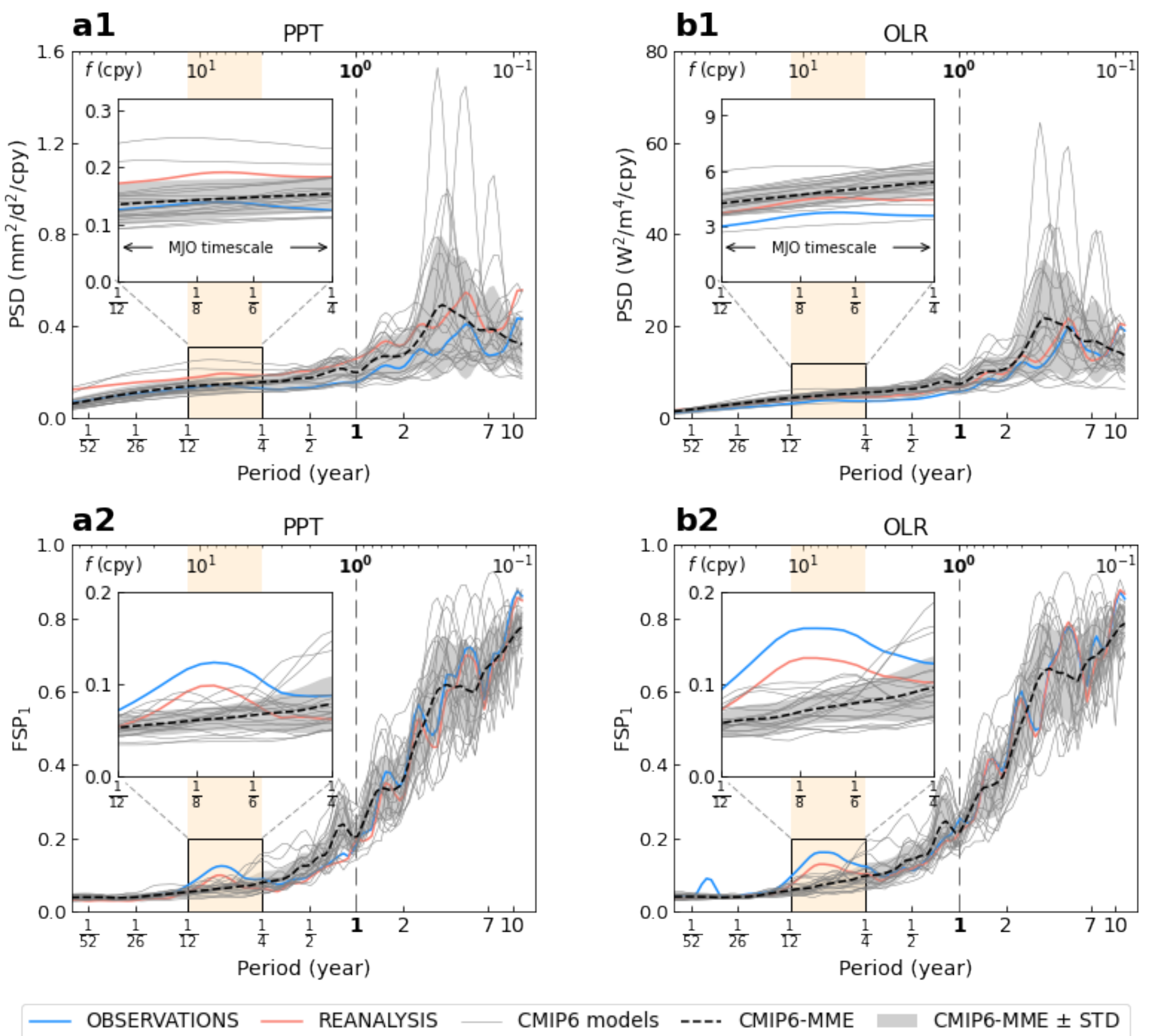

Figure 1. (Top) Power spectral density of (a1) daily precipitation rate (PPT) and (b1) daily outgoing longwave radiation (OLR). (Bottom) Fraction of spectral power explained by wPC1 for (a2) PPT and (b2) OLR. Blue lines correspond to observations, red lines to reanalysis data, dashed black lines correspond to the multi-model ensemble (MME) mean of 20 CMIP6 models, and the grey shaded regions represent MME $\pm$ standard deviation (here individual models are not distinguished from one another). The MJO timescale (yellow shaded vertical bands) ranges from 1-3 months. Frequency f in cycles per year (cpy) is shown in the top horizontal axes.

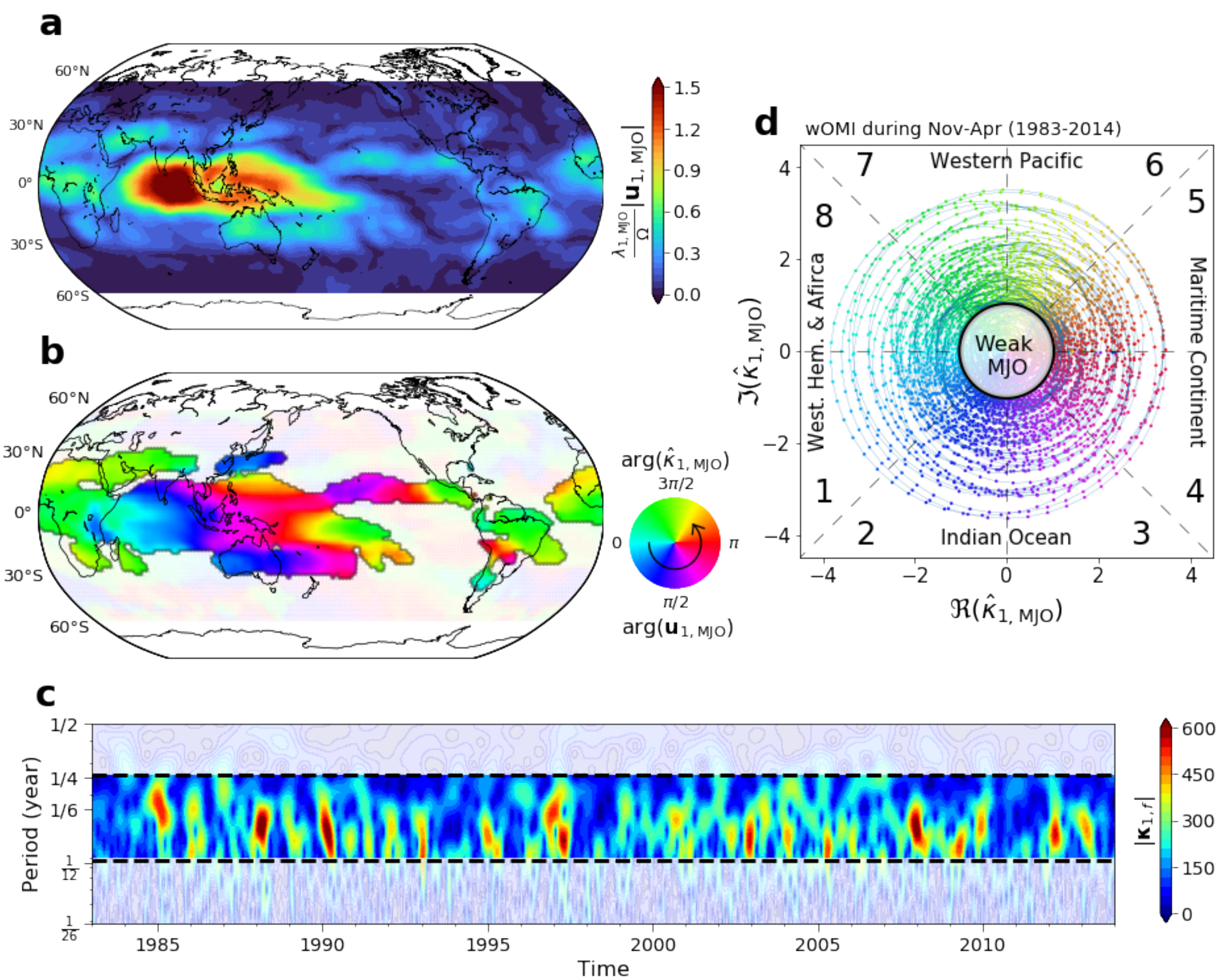

Figure 2. (a-b) Spatial patterns of the MJO band-integrated complex eigenvector $\boldsymbol{u}_{1,\mathrm{MJO}}$ for observed OLR for (a) modulus (magnitude) and (b) argument (phase). The modulus is shown for the unit-norm eigenvector with a scaling factor $\frac{\lambda_{1,\mathrm{MJO}}}{\Omega}$, with $\Omega = \sqrt{\frac{tr(\Lambda_{\mathrm{MJO}})}{N}}$. The counter-clockwise, circular arrow in the colorscale indicates the direction of the propagation of the extracted wave. **(c)** Magnitude of the complex wavelet PC1 time series $|\boldsymbol{\kappa}_{1,f}|$ associated with $\boldsymbol{u}_{1,\mathrm{MJO}}$ across frequencies for observed OLR. **(d)** Trajectory in the complex plane of the wsPCA-based OLR MJO index (wOMI) $\hat{\boldsymbol{\kappa}}_{1,\mathrm{MJO}}$ for observed OLR. The wOMI is displayed during boreal winter season (Nov-Apr) from 1983-2014 with one sample per day plotted. Points that lie inside the black unit circle correspond to days that are classified as weak/inactive MJO. The same colorscale as in panel (b) is used to represent the values of $\arg(\hat{\kappa}_{1,\mathrm{MJO}}(t))$, indicating the eastward propagation of MJO. See text for definition of variables.

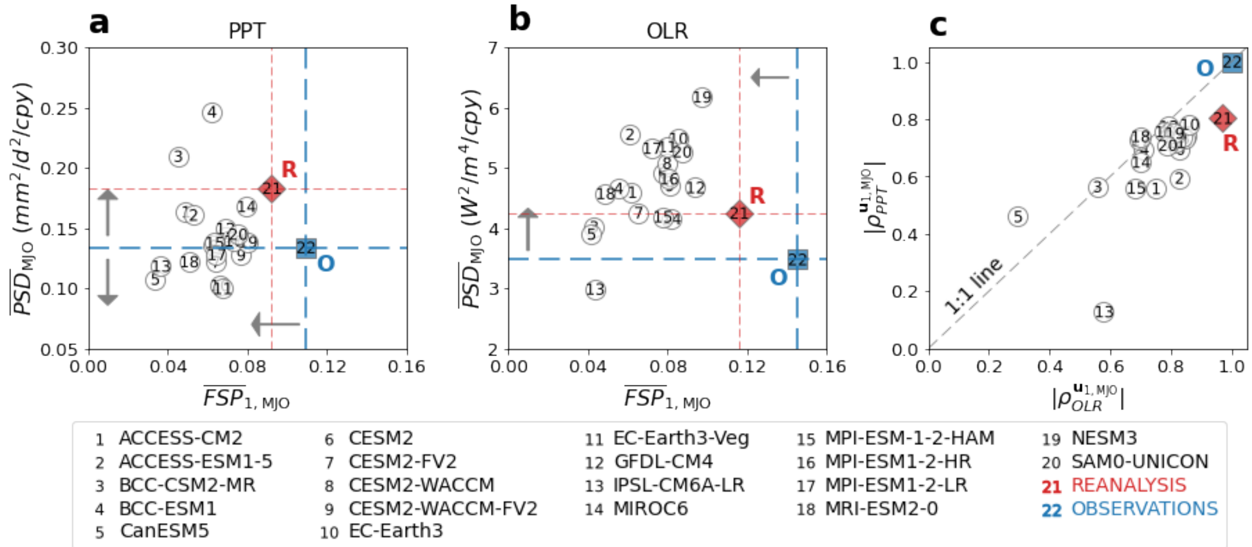

Figure 3. Comparison of spectral energy within the MJO frequency band ($\overline{PSD}_{\mathrm{MJO}}$) and the fraction of energy explained by wPC1 ($\overline{FSP}_{1,\mathrm{MJO}}$) for **(a)** Precipitation rate and **(b)** Outgoing longwave radiation for observations, reanalysis products and models. The systematic underestimation of $\overline{FSP}_{1,\mathrm{MJO}}$ in the models is apparent. **(c)** Scatter plot of the correlation coefficients of the patterns for the first complex eigenvectors of the modeled and observed OLR ($\left|\rho_{OLR}^{\boldsymbol{u}_{1,\mathrm{MJO}}}\right|$) and modeled and observed PPT ($\left|\rho_{PPT}^{\boldsymbol{u}_{1,\mathrm{MJO}}}\right|$) as defined in Equation (10). Numbers inside markers represent CMIP6 models (1-20), reanalysis (21), and observations (22).

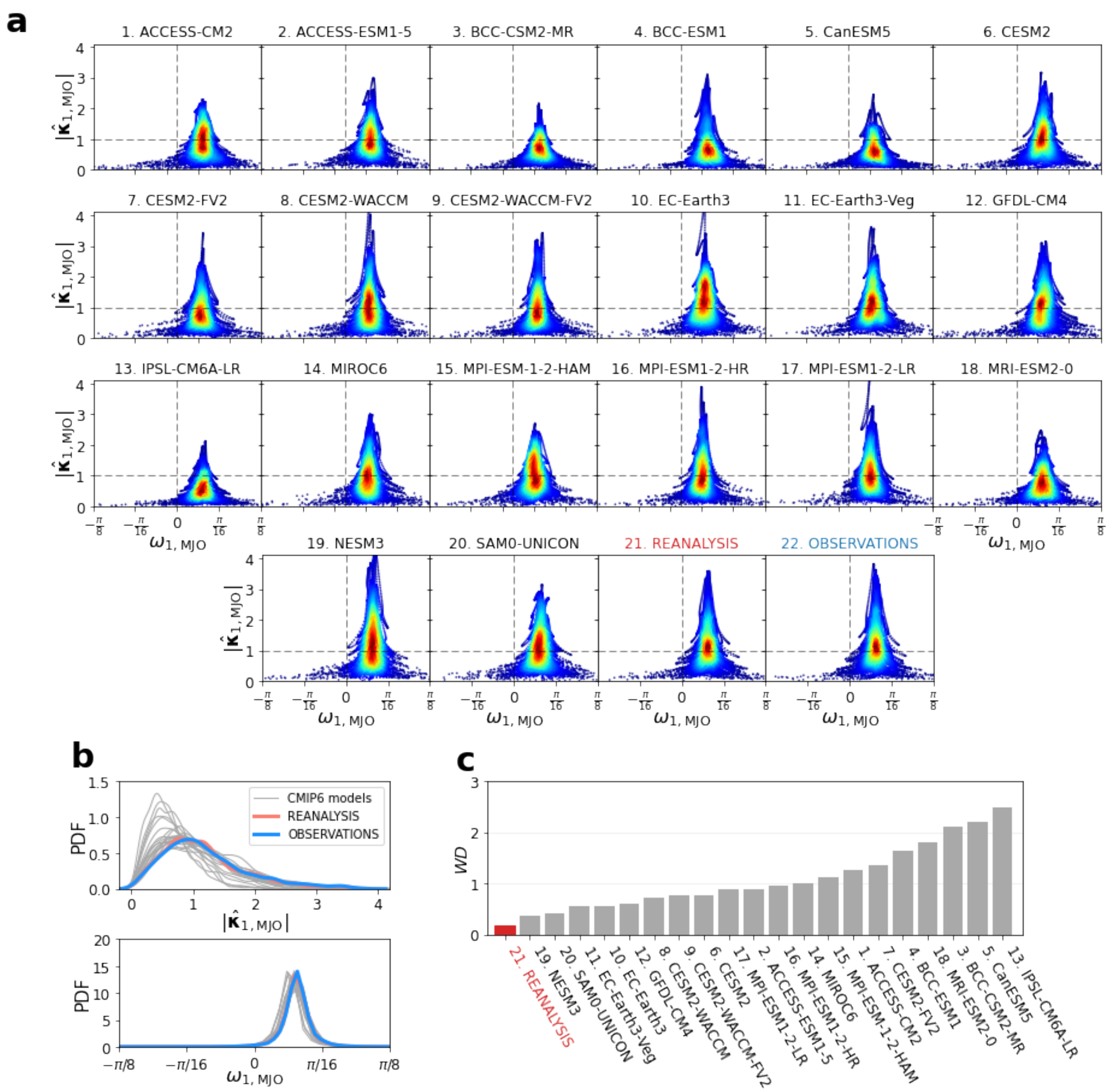

Figure 4. (**a**) Relationship between the normalized magnitude $\left|\hat{\kappa}_{1,\mathrm{MJO}}(t)\right|$ and angular frequency $\omega_{1,\mathrm{MJO}}(t)$ of CMIP6 models, reanalysis products, and observations computed at the daily scale. Points under the unit horizontal dashed lines are classified as weak MJO. The color scale represents the joint PDF of $\left|\hat{\kappa}_{1,\mathrm{MJO}}(t)\right|$ and $\omega_{1,\mathrm{MJO}}(t)$, with warmer color indicating higher probability. (**b**) Probability density functions of the MJO band-integrated wPC1 series for (top) magnitude $\left|\hat{\kappa}_{1,\mathrm{MJO}}(t)\right|$ and (bottom) angular frequency $\omega_{1,\mathrm{MJO}}(t)$. (**c**) Wasserstein distance (WD) between the joint probability distribution of $\left|\hat{\kappa}_{1,\mathrm{MJO}}(t)\right|$ and $\omega_{1,\mathrm{MJO}}(t)$ obtained from observations (reference) and those obtained from reanalysis (red bar) and CMIP6 models (grey bars). The WD values of models are sorted from low to high, indicating the ranking of CMIP6 models in reproducing the MJO magnitude and propagation dynamics.

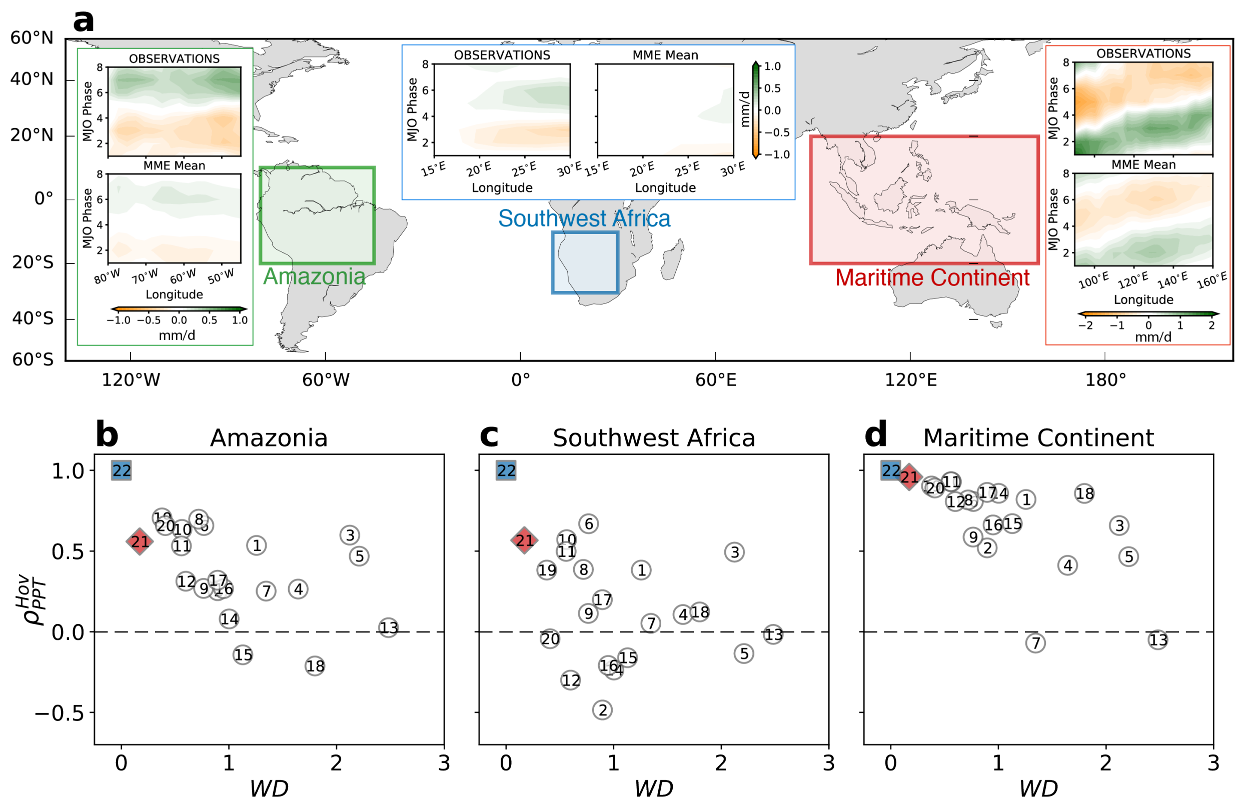

Figure 5. (**a, Top**) Comparison of modeled (CMIP6 MME Mean) and observed MJO-related precipitation anomalies around the climatic mean over the Amazonia (10ºN-20ºS, 45ºW-80ºW), Southwest Africa (10ºS-30ºS, 15ºE-30ºE), and Maritime Continent (20ºS-20ºN, 90ºE-160ºE). (Bottom) Scatter plots of the WD between the observed and modeled joint PDFs of $|\hat{\kappa}_{1,\mathrm{MJO}}(t)|$ and $\omega_{1,\mathrm{MJO}}(t)$ (see Figure 4) and the pattern correlation coefficients of the Hovmöller diagram (see Figures S7-S9) between models and observations for the (**b**) Amazonia, (**c**) Southwest Africa, and (**d**) Maritime Continent. Numbers inside markers represent CMIP6 models (1-20), reanalysis (21), and observations (22) as in Figure 3. It is seen that models that better reproduce MJO magnitude and propagation dynamics (low WD value) also tend to better reproduce the MJO-related precipitation variability over Amazonia and Maritime Continent, but not necessarily in Southwestern Africa.

Supporting Information for

Underestimated MJO in CMIP6 models

Phong V.V. Le[1,2], Clement Guilloteau[1], Antonios Mamalakis[1,3], and Efi Foufoula-Georgiou[1,4]

[1]Department of Civil and Environmental Engineering, University of California Irvine, CA, USA

[2]Faculty of Hydrology Meteorology and Oceanography, University of Science, Vietnam National University, Hanoi, Vietnam

[3]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA

[4]Department of Earth Systems Science, University of California Irvine, CA, USA

Contents of this file

Table S1: The selected 20 CMIP6 models used in our study with names**,** institutions and horizontal grid resolution of the atmospheric and ocean variables. The models were selected based on data availability at the time of writing the manuscript. The ID assigned to each model is used throughout this study.

ID	**Model**	**Institution Name**	**Average grid resolution** (longitude x latitude)	
			Atmosphere	**Ocean**
1	ACCESS-CM2	Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia	1.87° × 1.25°	1.0° × 1.0°
2	ACCESS-ESM1-5		1.87° × 1.25°	1.0° × 1.0°
3	BCC-CSM2-MR	Beijing Climate Center, Beijing, China	1.1° × 1.1°	1.0° × 0.78°
4	BCC-ESM1		2.8° × 2.8°	1.0° × 0.78°
5	CanESM5	Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, BC, Canada	2.8° × 2.8°	1.0° × 0.62°
6	CESM2	National Center for Atmospheric Research, Boulder, CO, USA	0.9° × 1.25°	0.9° × 1.25°
7	CESM2-FV2		1.9 ° × 2.5°	1.9° × 2.5°
8	CESM2-WACCM		0.9° × 1.25°	0.9° × 1.25°
9	CESM2-WACCM-FV2		1.9° × 2.5°	1.9° × 2.5°
10	EC-Earth3	Consortium of various institutions from Spain, Italy, Denmark, Finland, Germany, Ireland, Portugal, Netherlands, Norway, the United Kingdom, Belgium, and Sweden	0.7° × 0.7°	1.0° × 0.62°
11	EC-Earth3-Veg		0.7° × 0.7°	1.0° × 0.62°
12	GFDL-CM4	Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, NJ, USA	1.0° × 1.0°	0.25° × 0.16°
13	IPSL-CM6A-LR	Institut Pierre Simon Laplace, Paris, France	2.5° × 1.25°	1.0° × 0.54°
14	MIROC6	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, National Institute for Environmental Studies, and RIKEN Center for Computational Science, Japan	1.4° × 1.4°	1.0° × 0.70°
15	MPI-ESM-1-2-HAM	Max Planck Institute fur Meteorologie, Forschungszentrum Julich, University of Oxford, Finnish Meteorological Institute, Leibniz Institute for Tropospheric Research, ETH Zurich	1.87° × 1.87°	1.52° × 0.82°
16	MPI-ESM1-2-HR		0.94° × 0.94°	0.45° × 0.45°
17	MPI-ESM1-2-LR		1.87° × 1.87°	1.4° × 0.82°
18	MRI-ESM2-0	Meteorological Research Institute, Tsukuba, Japan	1.1° × 1.1°	1.0° × 0.5°
19	NESM3	Nanjing University of Information Science and Technology, Nanjing, China	1.87° × 1.87°	1.0° × 0.62°
20	SAM0-UNICON	Seoul National University, Seoul, Republic of Korea	1.25° × 0.94°	1.1° × 0.47°

Table S2: List of observed daily global precipitation products used for comparison.

ID	Name	Abbreviation	Period of record	Spatial resolution
1	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record	PERSIANN-CDR	1983-present	$0.25^{o} \times 0.25^{o}$
2	Integrated Multi-satellitE Retrievals for GPM	IMERG	2001-present	$0.1^{o} \times 0.1^{o}$
3	Tropical Rainfall Measuring Mission (34B2)	TRMM	1998-present	$0.25^{o} \times 0.25^{o}$
4	Global Precipitation Climatology Project	GPCP	1996- present	$1.0^{o} \times 1.0^{o}$
5	CPC MORPHing technique	CMORPH	2002- present	$0.25^{o} \times 0.25^{o}$

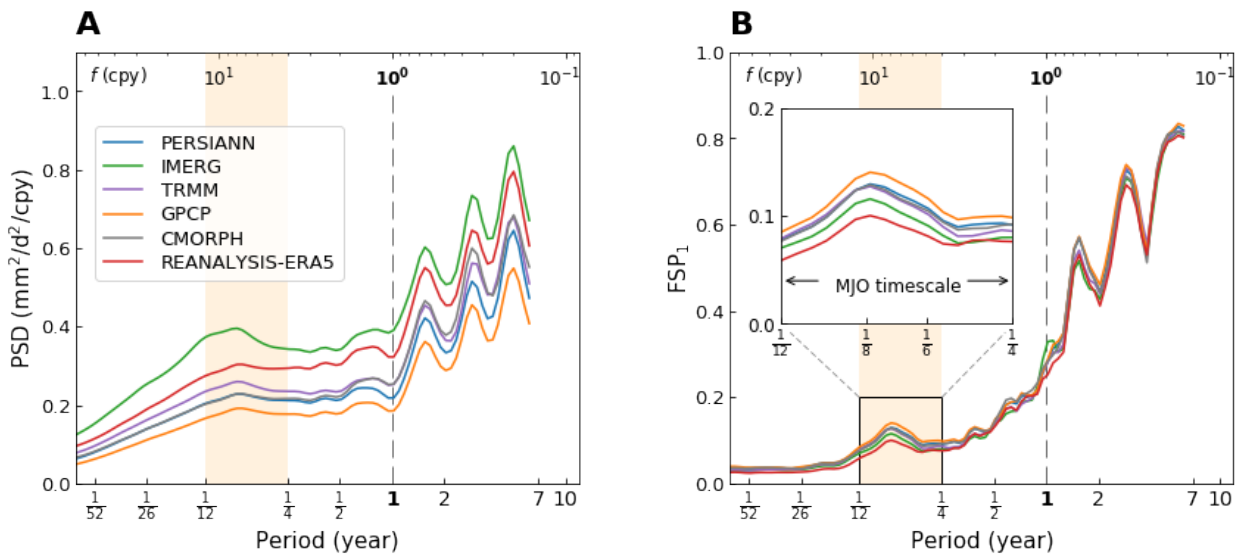

Figure S1. (a) Power spectral density and (b) Fraction of spectral power explained by the first wavelet principal component (wPC1) of daily precipitation rate obtained from 5 different observed datasets and the reanalysis products during the common period 2002-2019. The highest and lowest values of power spectral density is found in the IMERG and GPCP, respectively, but all observed datasets are in good agreement in terms of capturing the MJO mode. The MJO timescale (yellow shaded vertical bands) ranges from 1-3 months. Frequency f in cycles per year (cpy) is shown in the top horizontal axes.

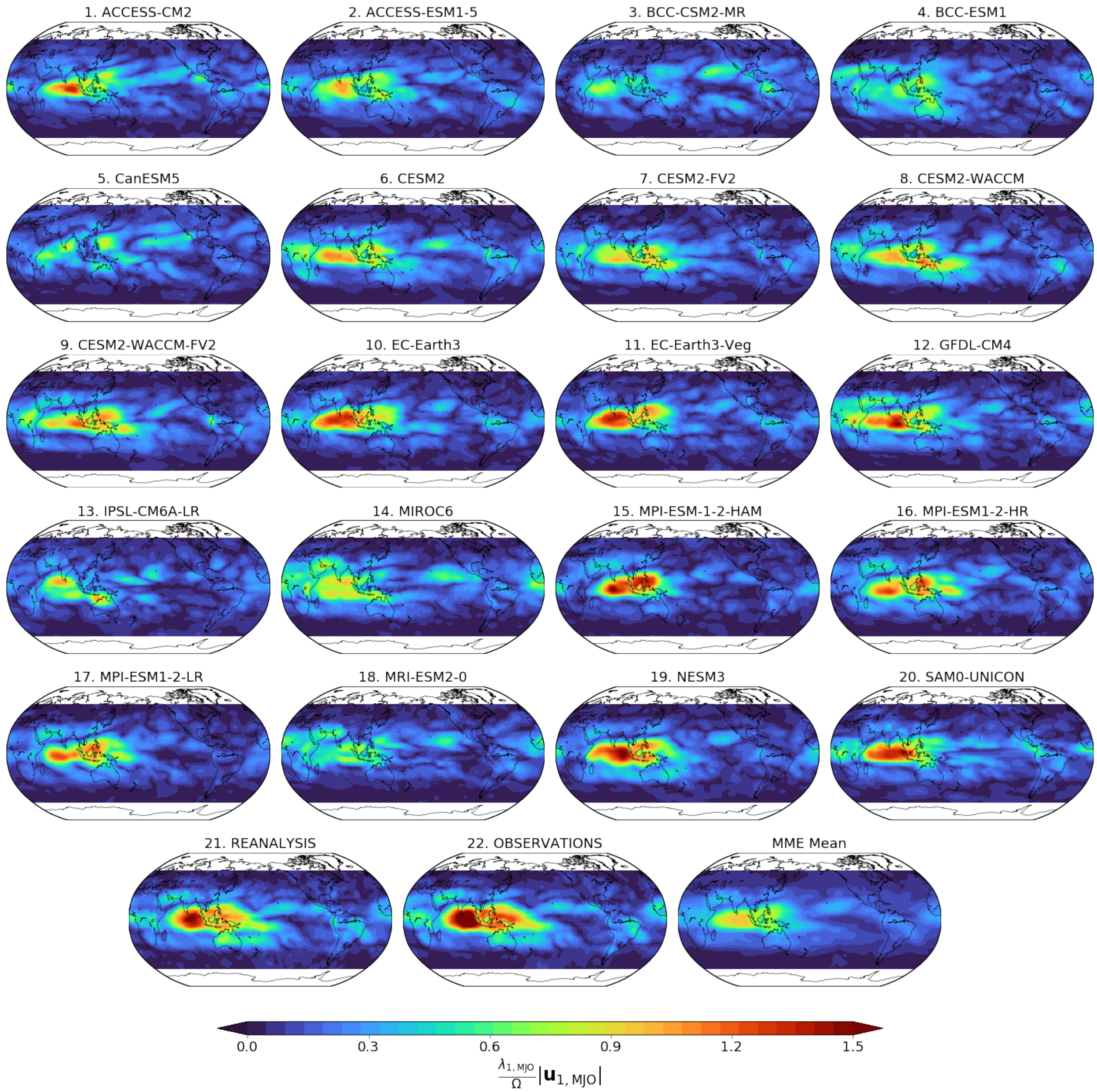

Figure S2. Spatial patterns of the modulus (magnitude) of the MJO band-integrated first complex eigenvector $\boldsymbol{u}_{1,\mathrm{MJO}}$ of OLR for the 20 CMIP6 models, reanalysis, and observations. MME Mean represents the mean of the 20 models. The map is shown for the unit-norm eigenvector with a scaling factor $\frac{\lambda_{1,\mathrm{MJO}}}{\Omega}$ representing the contribution of wPC1 to the total energy in the MJO frequency band, with $\Omega = \sqrt{\frac{tr(\boldsymbol{\Lambda}_{\mathrm{MJO}})}{N}}$. See text for definitions.

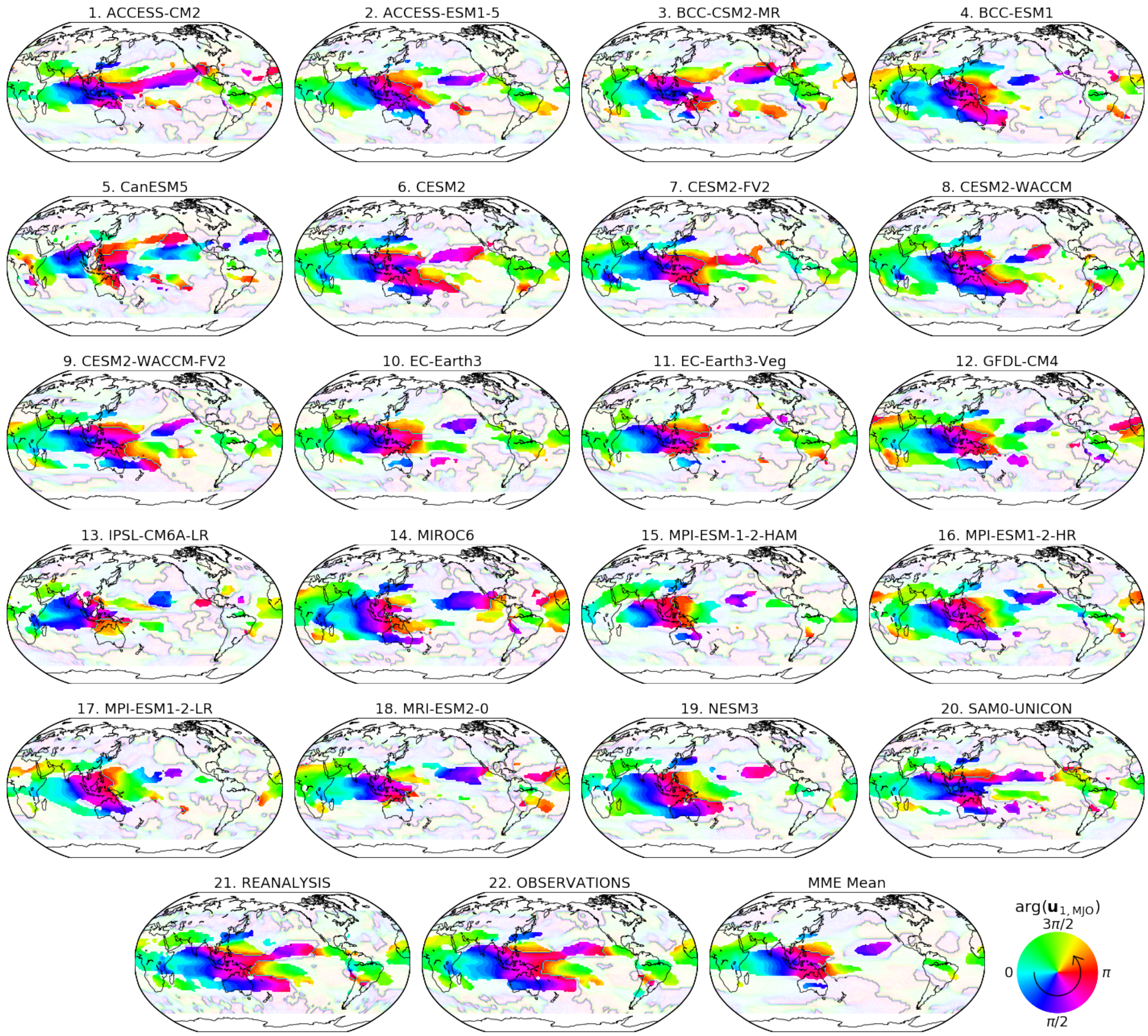

Figure S3. Spatial patterns of the argument (phase) of the MJO band-integrated first complex eigenvector $\boldsymbol{u}_{1,\mathrm{MJO}}$ of OLR for 20 CMIP6 models, reanalysis, and observations. MME Mean represents the mean of 20 models. The counter-clockwise, circular arrow in the colorscale indicates the direction of propagation of the extracted waves.

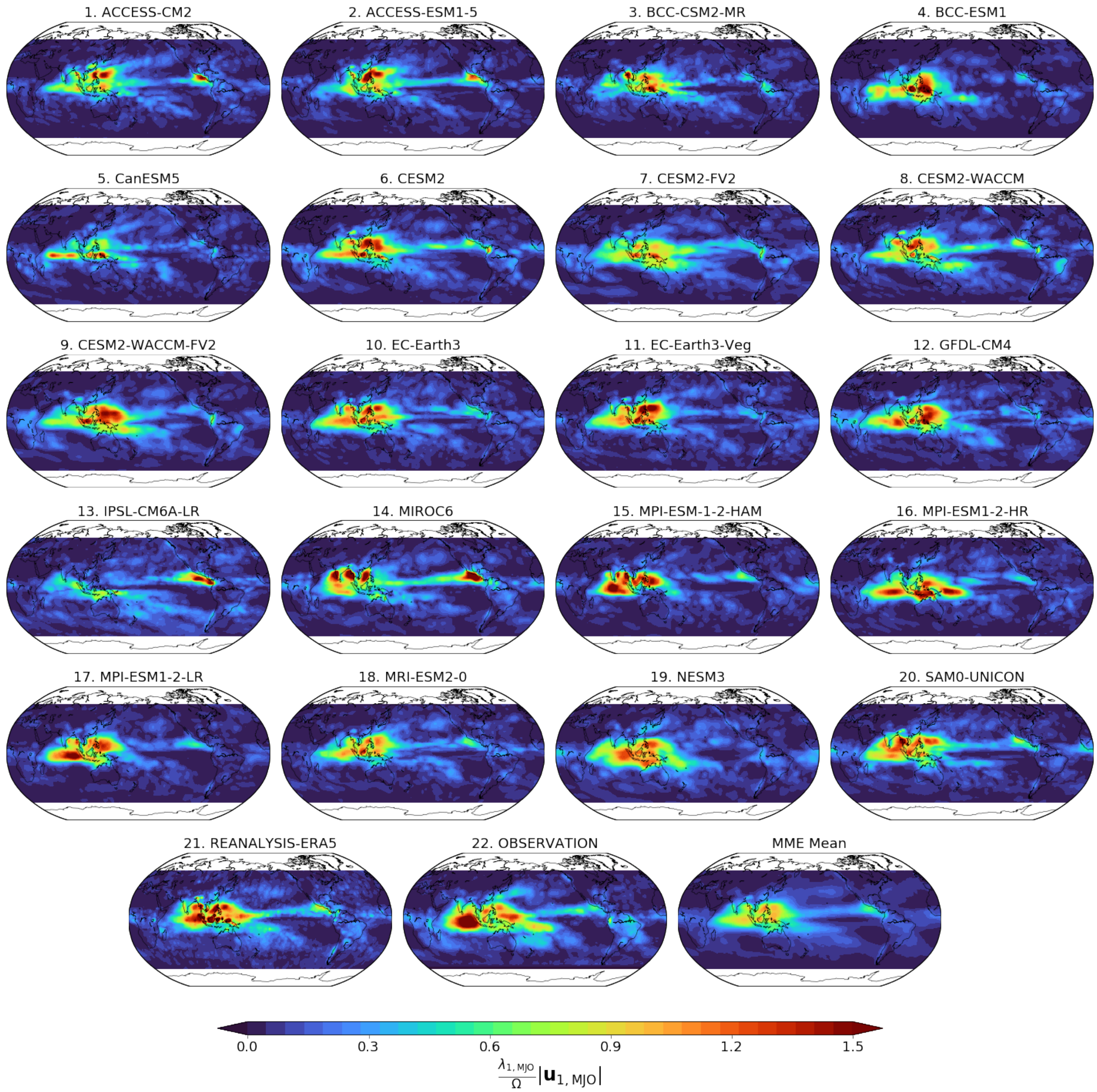

Figure S4. Spatial patterns of the modulus (magnitude) of the MJO band-integrated first complex eigenvector $\boldsymbol{u}_{1,\mathrm{MJO}}$ of PPT for the 20 CMIP6 models, reanalysis, and observations. MME Mean represents the mean of the 20 models. The map is shown for the unit-norm eigenvector with a scaling factor $\frac{\lambda_{1,\mathrm{MJO}}}{\Omega}$ representing the contribution of wPC1 to the total energy in MJO frequency band, with $\Omega = \sqrt{\frac{tr(\boldsymbol{\Lambda}_{\mathrm{MJO}})}{N}}$

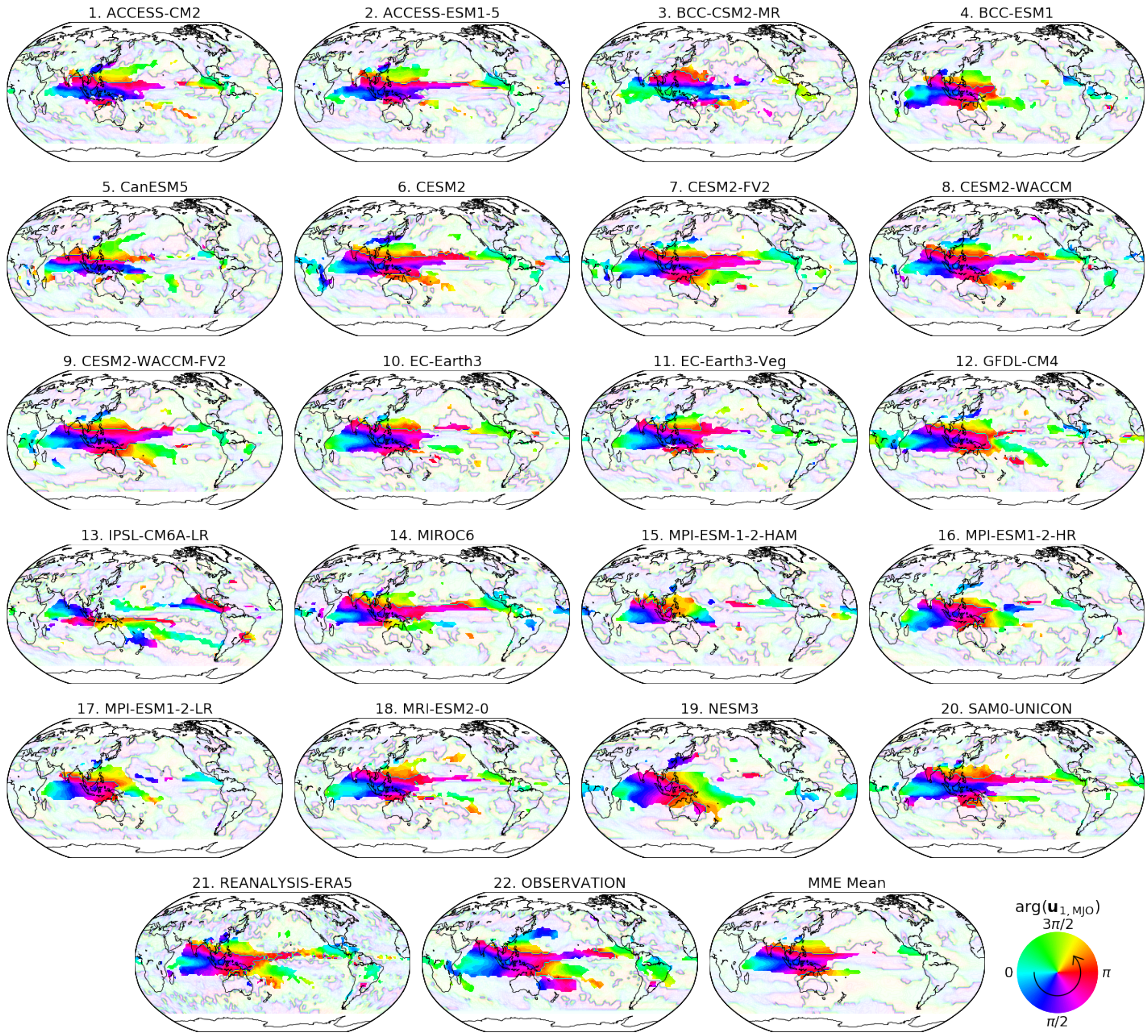

Figure S5. Spatial patterns of the argument (phase) of the MJO band-integrated first complex eigenvector $\boldsymbol{u}_{1,\mathrm{MJO}}$ of PPT for 20 CMIP6 models, reanalysis, and observations. MME Mean represents the mean of 20 models. The counter-clockwise, circular arrow in the colorscale indicates the direction of propagation of the extracted waves.

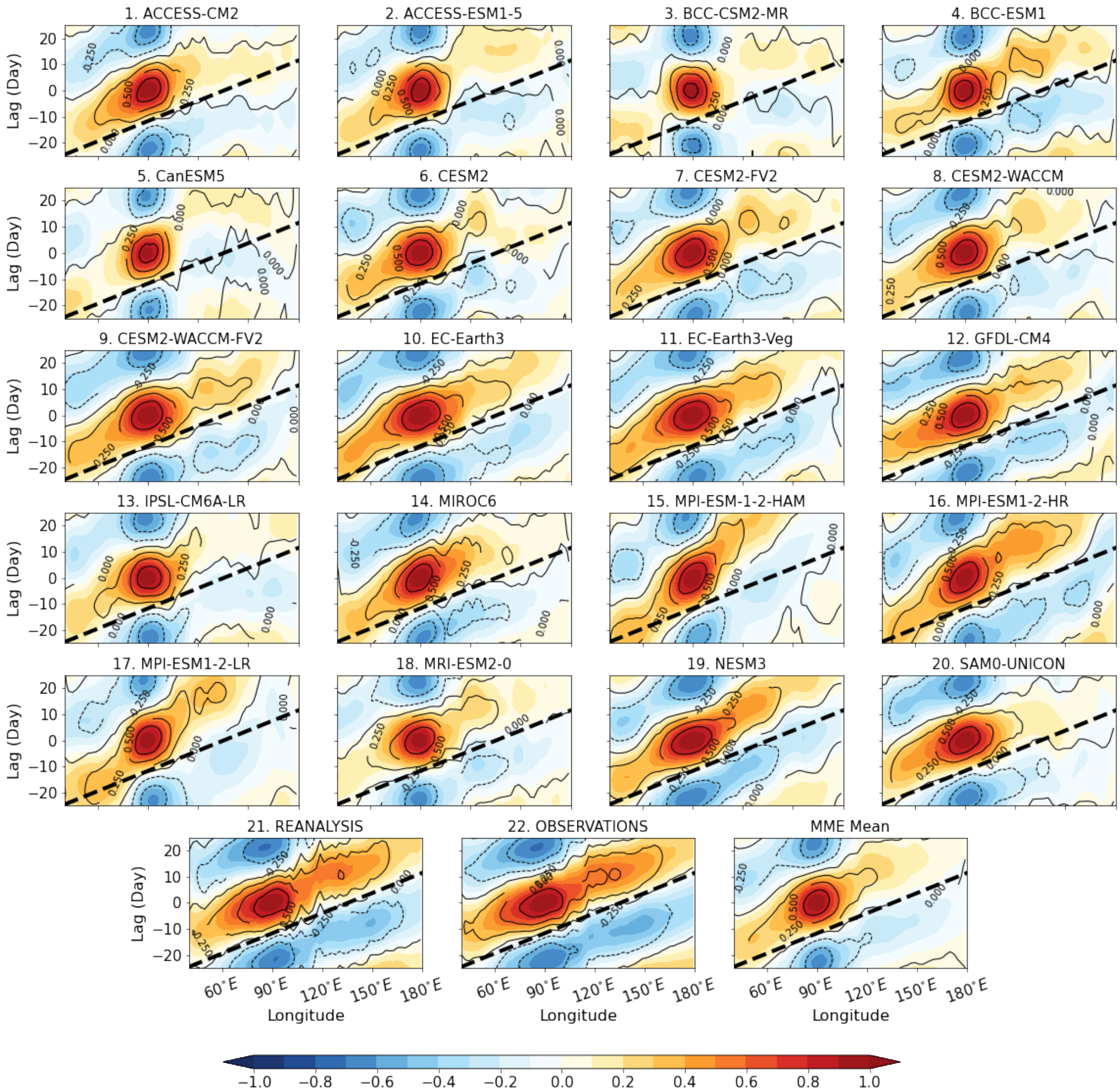

Figure S6. Lag-longitude diagram of 10ºS-10ºN-averaged OLR anomalies (colors) and PPT anomalies (contours) reconstructed within the MJO frequency band (30-90 days) against the corresponding OLR and PPT anomalies at the Indian Ocean reference region (10ºS-10ºN, 80º-100ºE) from 1983-2014. The reconstruction of OLR and PPT anomalies was performed through inverse wavelet transform of the wPC1 time series $\boldsymbol{\kappa}_{1,f}$ for frequencies f within the MJO frequency band. Black dashed lines indicate an eastward propagation speed of 5 m/s.

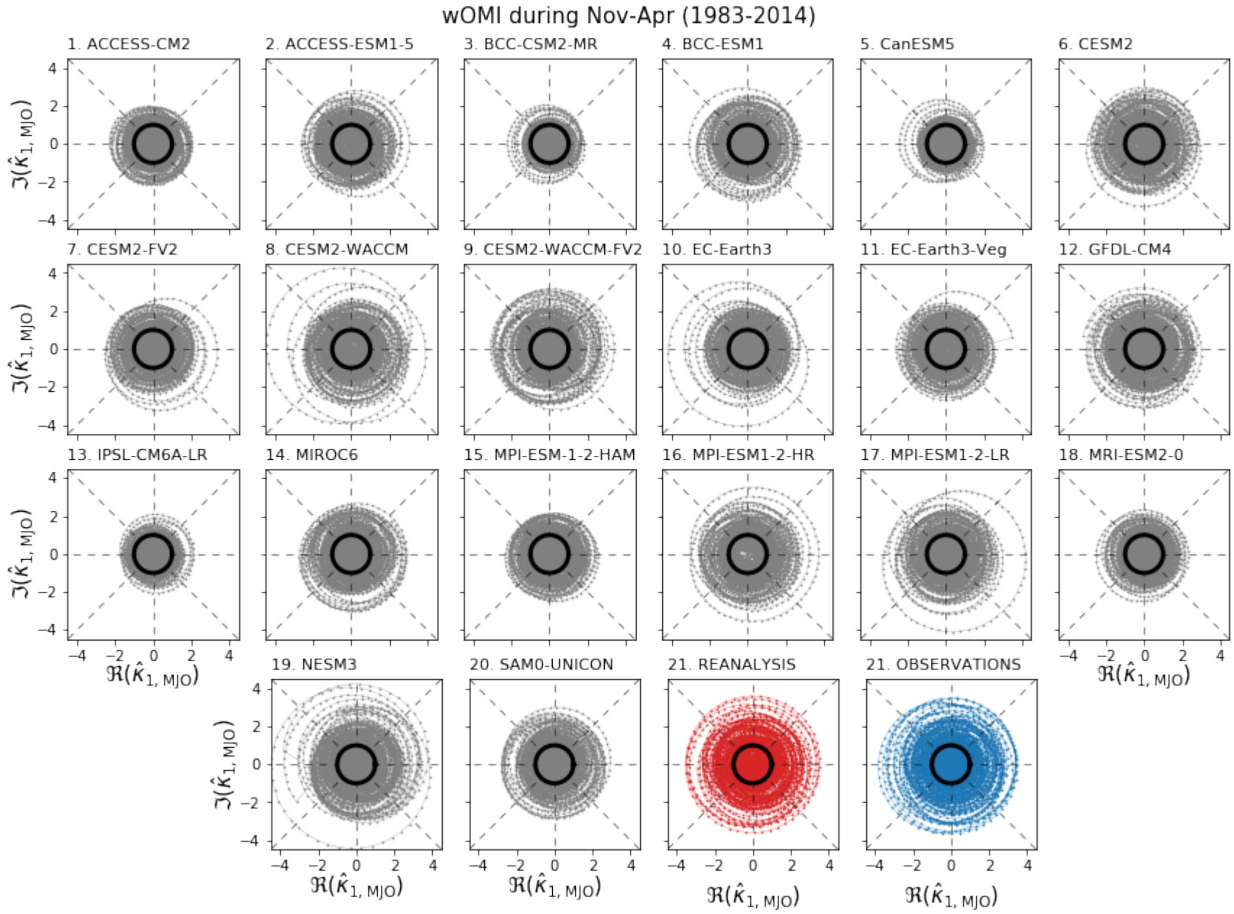

Figure S7. Comparison of wsPCA-based OLR MJO index (wOMI) $\hat{\boldsymbol{\kappa}}_{1,\mathrm{MJO}}$ reproduced by 20 CMIP6 models, reanalysis products, and observations. The wOMI plots are shown during boreal winter season (Nov-Apr) from 1983-2014. It can be seen that a large number of models underestimate the amplitude of MJO.

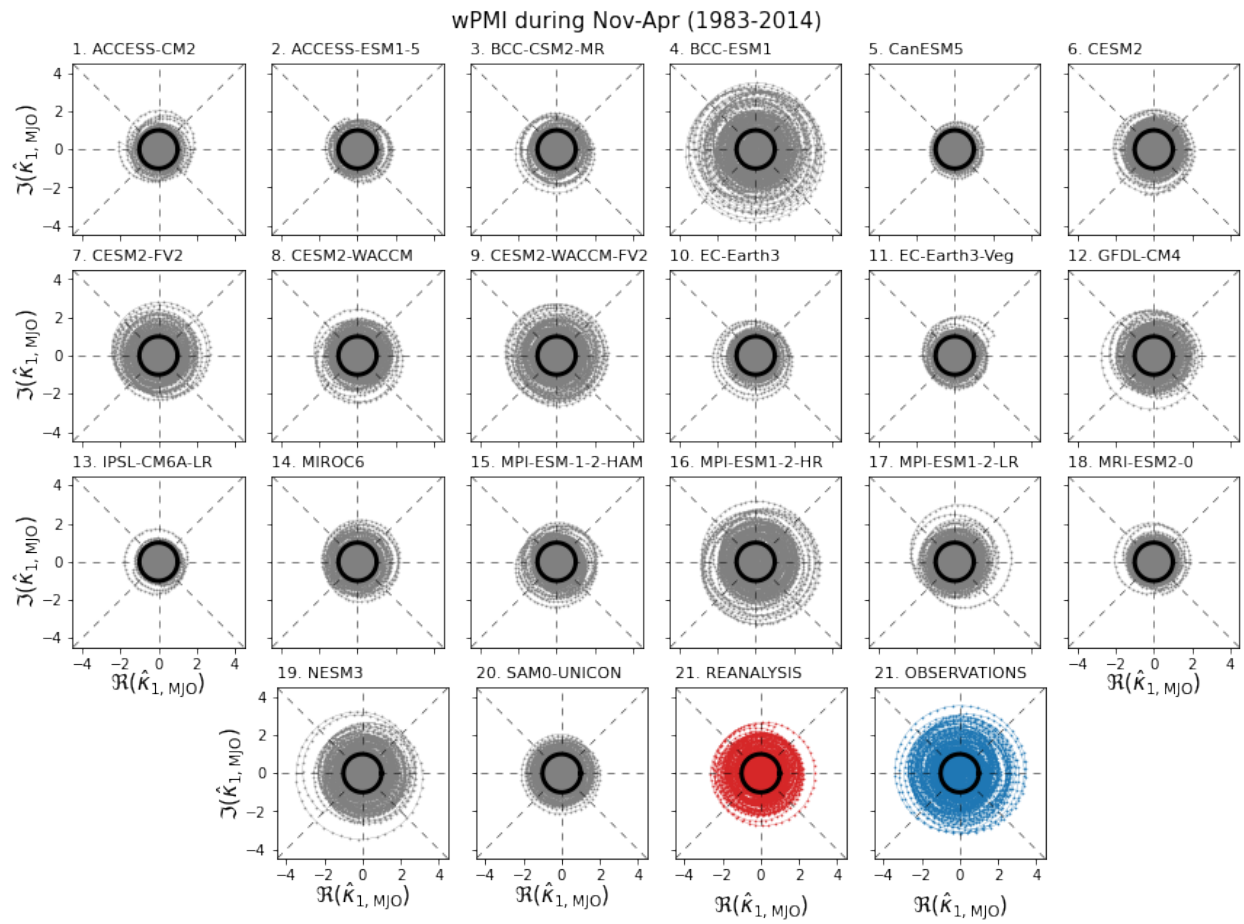

Figure S8. Comparison of wsPCA-based PPT MJO index (wPMI) $\hat{\boldsymbol{\kappa}}_{1,\mathrm{MJO}}$ reproduced by 20 CMIP6 models, reanalysis products, and observations. The wPMI plots are shown during boreal winter season (Nov-Apr) from 1983-2014. Similar to wOMI, it can be seen that a large number of models underestimate the amplitude of MJO using the wPMI.

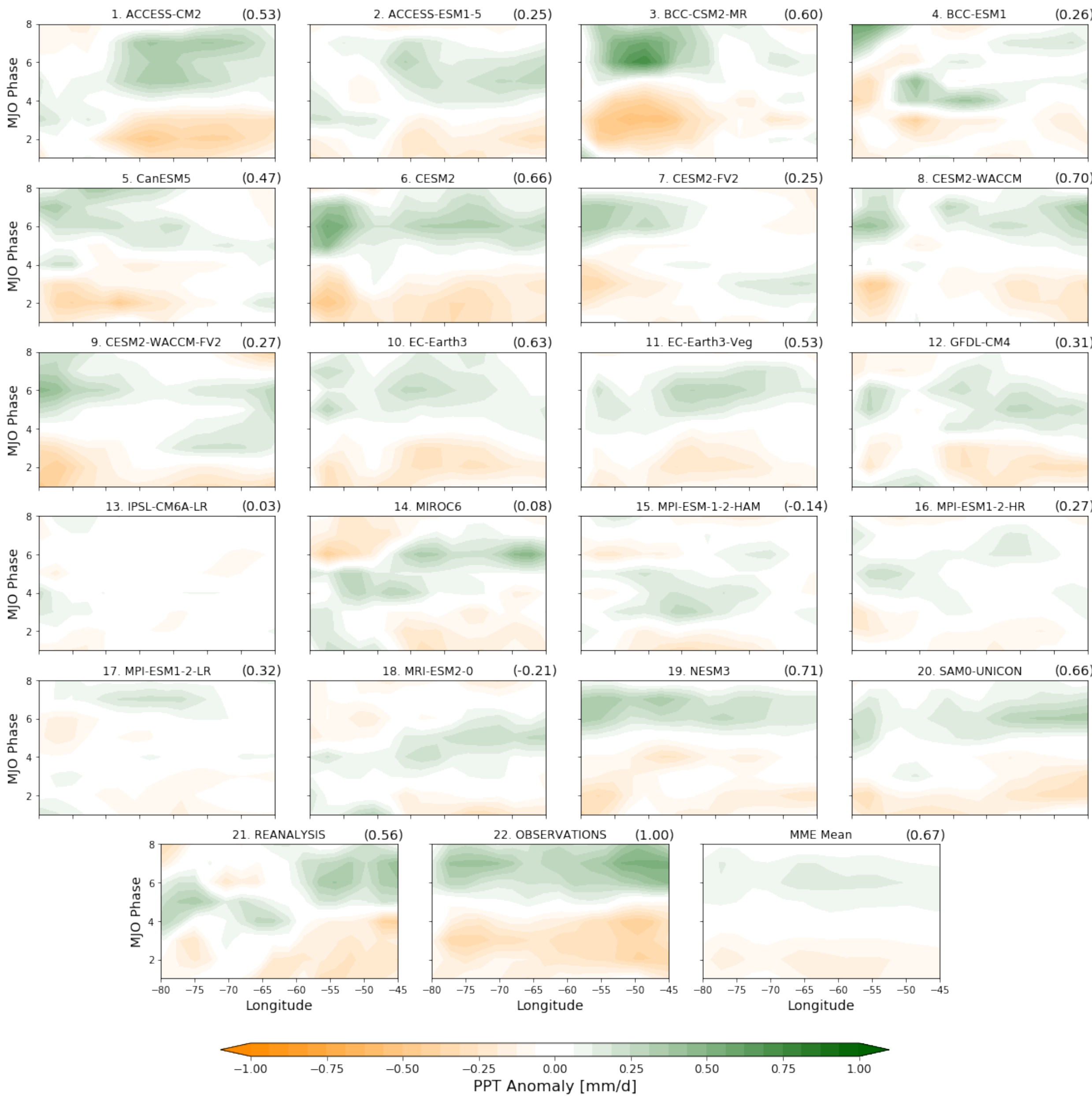

Figure S9. Hovmöller phase-longitude diagrams of PPT anomalies in the Amazonia (10°N-20°S, 45°W-80°W) for 20 CMIP6 models, reanalysis, and observations. MME Mean represents the mean of models. Numbers in parentheses represent the correlation coefficients of the phase-longitude patterns with the pattern obtained from the observations.

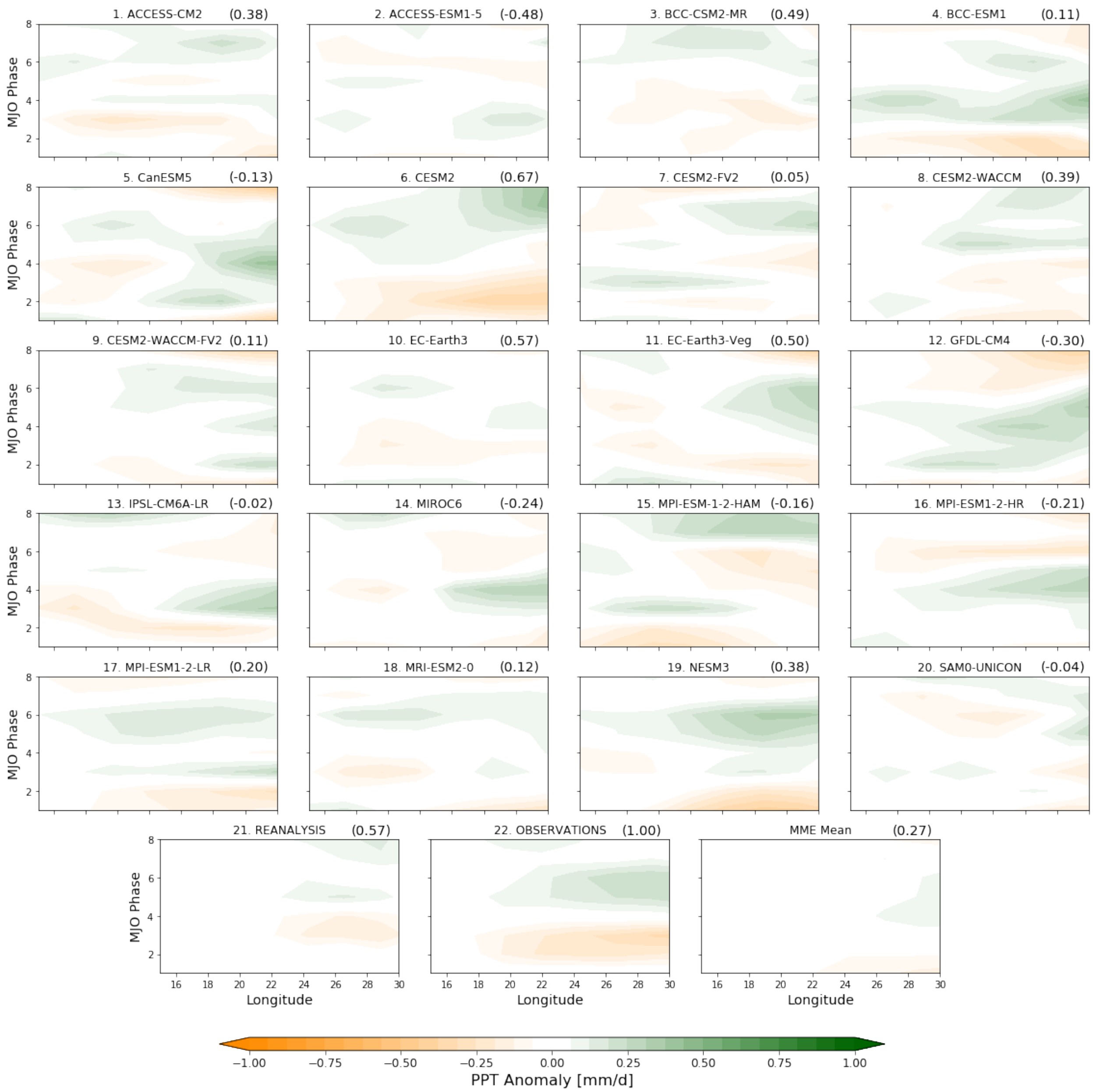

Figure S10. Hovmöller phase-longitude diagram of PPT anomalies in the Southwest Africa (10ºS-30ºS, 15ºE-30ºE) for 20 CMIP6 models, reanalysis, and observations. MME Mean represents the mean of models. Numbers in parentheses represent the correlation coefficients of the phase-longitude patterns with the pattern obtained from the observations.

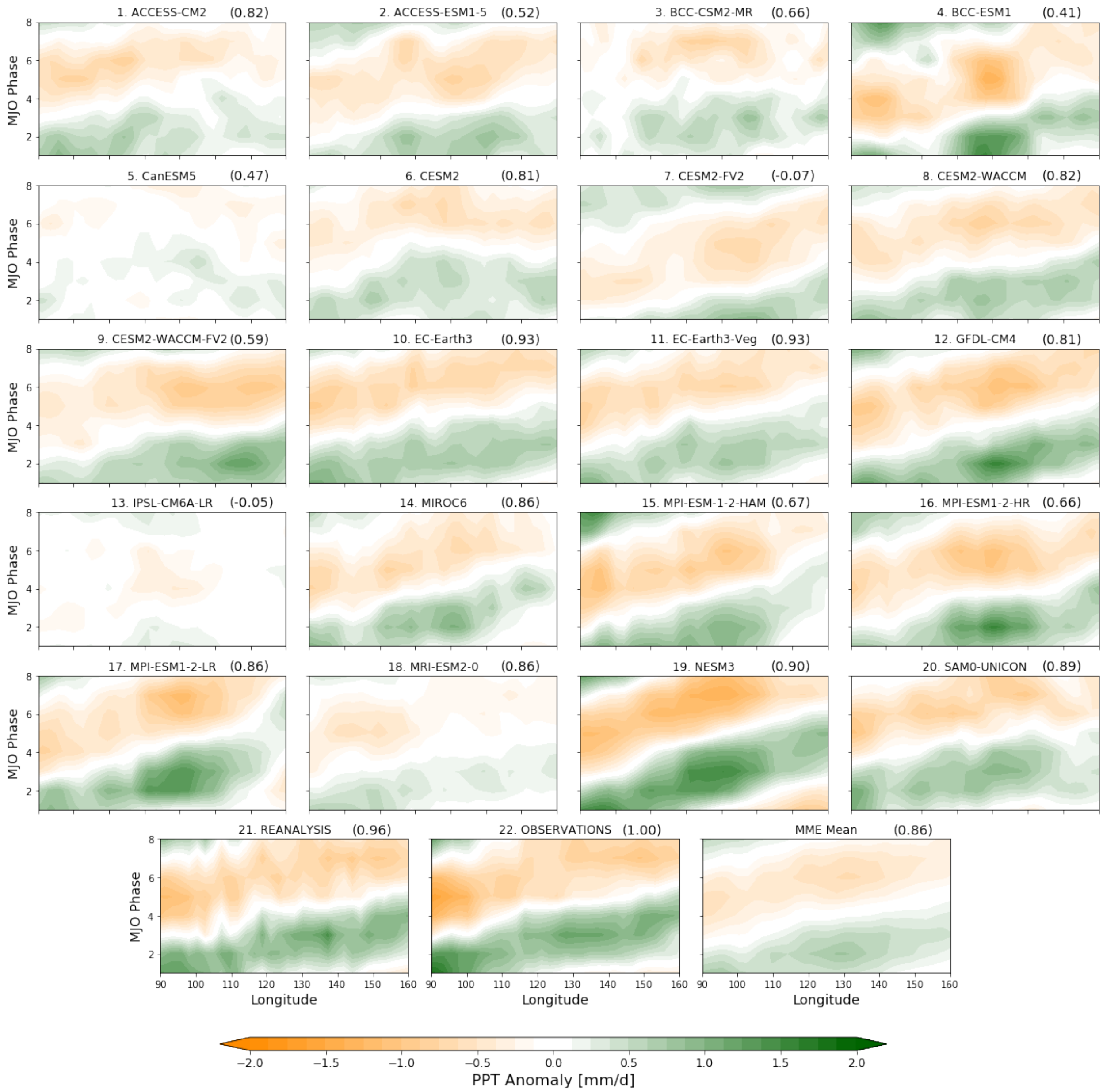

Figure S11. Hovmöller phase-longitude diagram of PPT anomalies in the Maritime Continent (20°S-20°N, 90°E-160°E) for 20 CMIP6 models, reanalysis, and observations. MME Mean represents the mean of models. Numbers in parentheses represent the correlation coefficients of the phase-longitude patterns with the pattern obtained from the observations.