RC2: 'Comment on egusphere-2025-4453', Anonymous Referee #2, 01 Nov 2025 reply

This study is evaluating the relationship between MJO and modes of climate variability in the tropics such as ENSO, IOD, IOBM, and stratospheric QBO during two periods: 1981-1998 and 1999-2018. The analysis indicates different relationships between the MJO and the other modes of variability from one period to another. Assuming that climate modes provide a source of predictability for the MJO, the second objective of the study is to test if models show a change in the forecast skill of the MJO for the two periods. While the first part is robust, the approach chosen for the second part has limitations because the POAMA2 model has a poor representation of stratospheric dynamics and the CESM2 and GEOS-S2S-2 models do not have data for the first period. There are other concerns that can be addressed and they are listed below.

We thank Reviewer 2 for their detailed feedback on our study. We acknowledge the critical limitations regarding model configuration and data availability and address each concern below. Please see the attached file for our responses to all comments.

POAMA2 Stratospheric Dynamics:

We recognise the limited representation of the stratosphere in POAMA2 (model top at 9–10 hPa), particularly when assessing the QBO–MJO relationship. However, we utilise this constraint diagnostically (also see Abhik & Hendon (2019) and Marshall & Hendon (2017)). A key point is that POAMA2 is an initialised forecast system: while it does not have sufficient vertical resolution to realistically simulate QBO dynamics in free-running mode, the QBO state is well-initialised at the forecast start, allowing QBO and its influence on the troposphere to be skilfully predicted (e.g., Table 2). Furthermore, the upper troposphere—where QBO–MJO interaction is thought to occur—has considerably better vertical resolution in POAMA2 than the stratosphere.

Importantly, POAMA2 nonetheless tracks observed decadal skill shifts despite its poor stratospheric resolution, highlighting the likely role of the tropospheric background state and large-scale forcing in modulating MJO predictability, often overcoming model deficiencies. In the revised manuscript, we have explicitly discussed these limitations (Section 3.4) and clarified that POAMA2's inclusion is motivated by experimental design and interdecadal continuity, and it’s relatively poor stratospheric resolution has limited impact on the results. We have also:

- Given greater emphasis to high-top models (CESM2 and GEOS-S2S-2) in our physical interpretations
- Softened claims relating to QBO–MJO mechanisms in POAMA2
- Explicitly noted that POAMA2 skill likely reflects statistical or coherent tropospheric response rather than dynamically realistic stratospheric pathways

Data Availability (CESM2 and GEOS-S2S-2):

We confirm that CESM2 and GEOS-S2S-2 data were only available for 1999–2018. To ensure robust interdecadal comparison, we relied primarily on models spanning both periods (ACCESS-S2 and POAMA2) and incorporated CESM2 and GEOS-S2S-2 for the second period analysis to broaden the assessment to additional modern, state-of-the-art systems. These constraints were necessary to conduct a multi-model decadal comparison using currently available hindcast data. Our methodology is designed to diagnose non-stationarity in MJO prediction skill and MJO–climate mode relationships within this framework.

We will now address the other specific concerns listed below.

> L236: The three models used in the study (ACCESS-S2, CESM2 and GEOS-S2S-2) are not part of what is known in the community as the S2S data base: https://apps.ecmwf.int/datasets/data/s2s/levtype=sfc/type=cf/

We apologise for the incorrect statement regarding the S2S database. We have removed references to ACCESS-S2, CESM2, and GEOS-S2S-2 being part of the official S2S database, as they do not meet the formal classification criteria.

> Section 3.1: Please provide a table showing which years have been used for each of the phases of the climate modes shown in Fig. 1. The table can be in the supplement file.

We have included a supplementary table listing the specific years used for each climate mode phase (ENSO, IOD, IOBM, QBO) for the two periods.

> L238: Please discuss the source of initial conditions used for POAMA2 and ACCESS-S2. If they use the same initial conditions the difference in skill will be solely due to models' differences. If there was any change in the DA system used to generate the initial conditions between the two periods, that should also be discussed.

This is a crucial point regarding model independence. We have added clarification that POAMA2 and ACCESS-S2, while both are the Bureau of Meteorology system's, use different initial conditions and data assimilation (DA) systems. POAMA2 employs its own internal DA system, whereas ACCESS-S2 uses a more modern, separate system. Consequently, differences in skill arise from a combination of model physics/dynamics and initialisation differences. Critically, there was no change in the DA system within each model between the two periods (1981–1998 and 1999–2018), ensuring a consistent framework for interdecadal comparison.

> L389-394 and L419-422: Coincidently, the two periods considered in the study correspond to two phases of the Pacific Decadal Oscillation (PDO). 1981-1998 is mostly dominated by positive values of the PDO index whereas the 1999-2018 is dominated by negative values of the PDO index. This is also another factor affecting the mean state and should be mentioned when describing the shift in the background state.

This is an insightful observation and we should have clarified this in the introduction. We have now added a discussion acknowledging that the two study periods approximately correspond to the Pacific Decadal Oscillation (PDO) phase shift: 1981–1998 was dominated by positive PDO values, while 1999–2018 experienced predominantly negative PDO values. We have incorporated this as another large-scale factor contributing to changing the background dynamical environment and non-stationarity in MJO predictability.

> L396: Please explain how 'the mean DJF duration and total yearly event count for DJF' are calculated.

We apologise for not including a statement on the mean DJF duration and total yearly event count for DJF. We have now updated the MJO event estimation section in the Supplementary material. An event is defined as an occurrence of MJO amplitude >1 for at least 7 consecutive days and progression to at least 2 subsequent phases (adapted from Wei & Ren (2019)). Total yearly event count is the total number of MJO events for a given DJF, and the mean duration is the average duration for these MJO events. We have also updated the text around Line 321 to point to this new section.

> L399-400: Figure S1 shows the phases grouped as 4, 5, 6, 7 and 8, 1, 2, 3. One cannot see that 'the MJO spends more days in phases 3–6.' If this is the message that the figure is intended to convey, then the grouping of phases should be 3,4,5,6 and 7,8,1,2.

We apologies for this typographical error which was also noted by Reviewer 1. We have corrected the legend for Figure S1 to reflect the actual phase groupings used in the plot: Phases 3, 4, 5, 6 and Phases 7, 8, 1, 2, ensuring the figure clearly conveys that the MJO spends more days in phases 3–6.

> L405-407: If the negative correlation is explained by the enhanced frequency of N-IOD years the reversal of sign means an increase frequency of positive IOD years? Weakening means a lower value of r?

The interpretation is correct. We have clarified the text to explicitly address the shift: The negative correlation in 1981–1998 ($r = -0.48$) reflects an increase in MJO event frequency during Negative Indian Ocean Dipole (N-IOD) years. The subsequent sign reversal to a positive but weak correlation in 1999–2018 ($r = 0.29$) indicates a shift toward slightly higher MJO event frequency during Positive IOD years in the later period, contrasting the substantial N-IOD enhancement observed earlier. Weakening is precisely defined by correlation values moving closer to zero.

> L405-422, L579-590: I suggest summarizing all correlation coefficient values into a table.

We agree that this information is best presented in a centralised, easily digestible format, as also suggested by Reviewer 1. We have created a comprehensive supplementary table that summarises all the linear correlation coefficients discussed in **Sections 3.1 and 3.4** (MJO event characteristics vs. climate indices, and multi-model skill index vs. climate indices) for both periods.

> Figure 2: Panel B ‘S2’ should be ACCESS-S2

We appreciate the attention to detail. We have corrected the legend in Figure 2B to display **ACCESS-S2** instead of 'S2'

> L512-514: These results should be connected to the findings of Jiang et al. (2015, https://doi.org/10.1002/2014JD022375). They also show that feedbacks between moist convection and circulation are critical for simulation of the MJO.

We agree that drawing this connection strengthens the discussion on model biases. We have added a reference and discussion to Jiang et al. (2015) to support the point that poor representation of convection schemes and their coupling to large-scale dynamics (highlighted by the SST-moisture decoupling in Figure 3) is a recognised limitation in MJO simulation.

> Please explain the interpretation of regression analysis. The idea of identifying patterns associated with high/low MJO skill depends on how the low skill is defined. For example, if the correlation coefficient has a large negative value, the skill is low, but the regression coefficient will have a large value. Second, what is the reason for regressing observations onto the model skill? And lastly, the regression coefficients in Fig. 4 show very limited statistical significance, which raise the question of how robust this analysis is.

We appreciate the request for clarification. We have expanded the explanation in Section 3.4:

Linear Regression Method: Regressing observed fields (e.g., OLR) onto the multi-model MJO skill index identifies large-scale patterns associated with years of high model MJO skill. Positive/negative regression coefficients signify that positive/negative anomalies of the state variable tend to occur during years when models show high MJO forecast skill. This links model performance directly to real-world large-scale forcing. Regressing **observed** fields onto the **model's MJO skill index** determines whether the years a model finds easy to predict (high skill) correspond to a specific, observed background climate state. This links model performance directly to real-world large-scale forcing.

Significance Level: We note that the significance threshold is set at $p < 0.10$, a standard for decadal variability studies due to limited sample sizes. While individual regions show limited stippling, the discussion emphasises consistent dipole structures in the Indian Ocean (Figure 4, Period 1) and overall pattern shifts, suggesting robust structural changes across epochs despite limited local significance.

The usage of POAMA2 model for the evaluation of MJO-QBO relationship raises some questions about this model ability to resolve the stratosphere more than what is acknowledged in the study. The model top is located at 10 hPa, meaning that the model does not have a full stratosphere. Compared to the QBO lifecycle, these forecasts are relatively short, and the model might be tuned to have a 'good QBO' but miss the QBO dynamics.

We fully acknowledge and agree with the reviewer that POAMA2's relatively low model top (≈9–10 hPa) is a limitation for faithfully simulating the QBO. Models with such a low upper boundary cannot fully represent the vertical structure, wave–mean flow interaction, or downward influence of the QBO in the same way as high-top models.

However, we note that, as these are initialised experiments, the QBO state is well represented at the beginning of the forecast, thus allowing the QBO's impact on the MJO to persist. Furthermore, the vertical resolution of POAMA 2 around the upper troposphere, where the QBO-MJO interaction is thought to occur, is much better than its resolution in the stratosphere. Therefore, there is still some utility in using POAMA2 to study the impact of the QBO on the MJO.

In the revised manuscript, we now explicitly discuss this in Section 3.4 and the Conclusions.

Importantly, POAMA2 is **not used here as a benchmark for realistic stratospheric representation**, but rather for **consistency in the decadal comparison**. This continuity enables a clean assessment of temporal changes in forecast skill under a constant model framework.

To ensure that physical interpretations of the QBO–MJO coupling are not disproportionately influenced by a low-top model, we have:

- Given greater emphasis in the discussion on the **high-top models** (ACCESS-S2, CESM2 and GEOS-S2S-2),
- **Softened claims** relating to the QBO–MJO mechanism in POAMA2,
- Explicitly noted that any POAMA2 "skill" is likely a **statistical persistence or coherent tropospheric response**, rather than a dynamically realistic stratospheric pathway.

These changes clarify that POAMA2's inclusion is motivated by experimental design and interdecadal continuity, not by its capacity to resolve the full stratospheric dynamics of the QBO.

Fig. 6 Caption: Please explain why some boxes are filled with color. On the x-axis please use font of different colors for denoting the two periods and draw a thick vertical line between the left side of the plot with event count and the right side of the plot with event duration.

We thank the reviewer for this helpful suggestion. These changes have been fully implemented:

- Coloured boxes: The figure caption now explicitly states that filled/coloured boxes indicate statistically significant differences between periods at the 90% confidence level ($p < 0.10$), while unfilled boxes denote non-significant values.
- Visual Improvements:
    - Used different font colours on the x-axis to distinguish Period 1 (1981–1998) and Period 2 (1999–2018)
    - Added a thick vertical dividing line separating event count panels (left) from event duration panels (right)

These modifications significantly improve figure readability and clarify the intended comparisons between periods and MJO characteristics.