Reply to reviewer

Dear Dr. Sugata Narsey,

We sincerely thank you for your constructive comments and insightful suggestions, which have significantly enhanced the quality of this manuscript. Below, we have provided a detailed, point-by-point response to each of your comments in **blue font** and correction made in the revised manuscript with line numbers in **red font**.

Title: Multiscale Assessment of Indian monsoon rainfall using ICON and CMIP6 model simulations

By: Pokhrel et al

Reviewer: Sugata Narsey, Australian Bureau of Meteorology

Recommendation: Accept with minor revisions

In this manuscript the authors systematically evaluate Indian monsoon rainfall at multiple time scales in the ICON non-hydrostatic model and 5 CMIP6 models under a prescribed SST set-up. They explore whether a non-hydrostatic model, and/or higher spatial resolution, improves the simulation of Indian monsoon rainfall across scales. Their findings are mixed – while ICON shows some improvements over the CMIP6 models it also has some larger deficiencies at some time scales and in geographic locations. Likewise, higher resolution alone was not a panacea for Indian monsoon rainfall, with relatively coarse models such as IITM-ESM outperforming or matching other models in some areas. The manuscript is well-written and presented, the analysis is of a high quality, the work is novel, and the conclusions are well-supported by evidence. I recommend that the manuscript is accepted with minor revisions.

Main comments:

1. While the multiscale error assessment is useful, I would find this manuscript to be more useful overall if there were a summary figure or analysis specifically showing the cross-scale relationships of errors, e.g. a heatmap of metrics for each model. I would cite this paper for that specifically! The abstract asks how scale-dependent biases shape monsoon variability, but I feel that this was not addressed clearly in the paper (but could be with one more figure).

Reply: We thank the reviewer for this valuable suggestion. Our existing Fig. 13 of the manuscript although quantified rainfall variance biases at individual timescales, however, it was unable to connect these biases in terms of cross-scale bias relationship. To address this, we have added two new summary figures consisting of complementary diagnostics.

The first figure (Fig-R1) quantifies the percentage contribution of each timescale (diurnal, synoptic, HF-MISO, LF-MISO, and seasonal) to the total variance bias over Central India (CI) and the Bay of Bengal (BoB) respectively. While Fig. 13 presents the magnitude of variance bias at each timescale separately, this new diagnostic (Fig-R1) identifies the dominant source of error within each model. For example, over CI, the largest error contribution arises from the diurnal scale in MPI, the synoptic scale in GFDL, and the HF-MISO band in CNRM, ICON, IITM-ESM, and MIROC. Over the BoB, diurnal-scale errors dominate in ICON and GFDL, synoptic-scale errors in MPI, HF-MISO errors in MIROC, and LF-MISO errors in IITM-ESM. Seasonal-scale errors in general contribute the least across nearly all models and regions. Thus, the figure provides a compact summary of the dominant timescales governing the overall variance bias in each model.

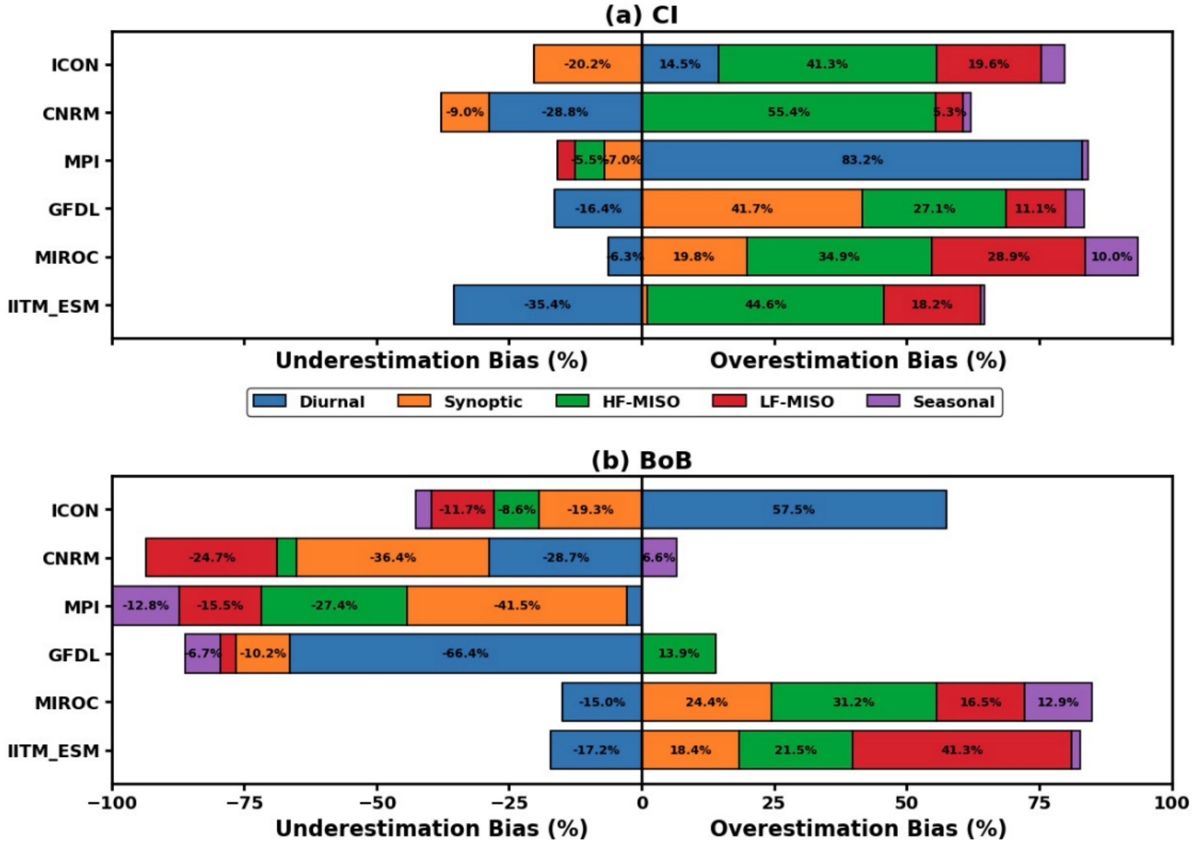

Figure R1: Percentage contribution of individual timescale variance biases relative to each model's total variance bias over (a) land (Central India) and (b) ocean (Bay of Bengal) domains.

The second figure (Fig-R2) presents a heatmap of spatial correlations between variance bias fields at different timescales. This heatmap explicitly examines whether bias patterns at one timescale are spatially related to those at another timescale. For physical interpretation, the heatmap is organized into three regimes: Convective Feedbacks, Synoptic Dynamics, and Intraseasonal Rectification.

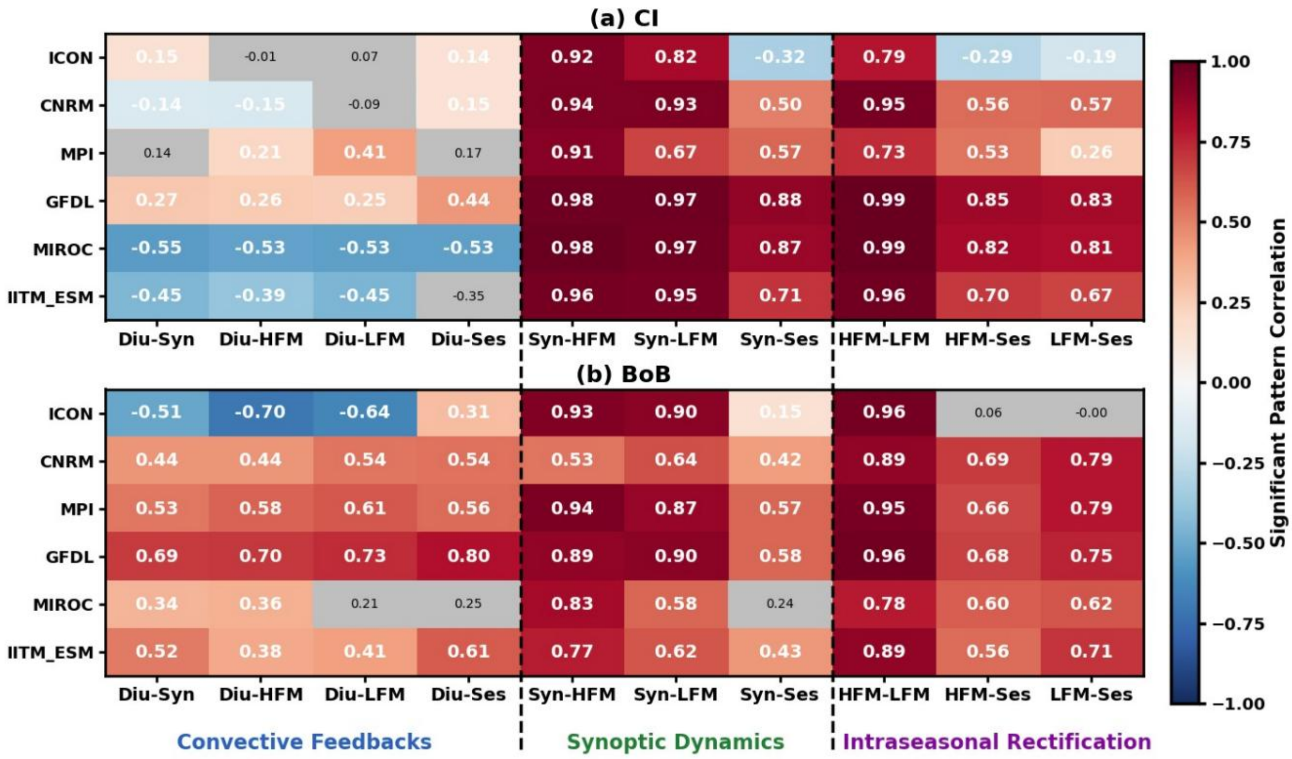

Figure R2. Heatmap of spatial pattern correlation coefficients among the relative bias fields of different scales. Correlations significant at the 95% confidence level are highlighted in bold white font, while non-significant cells are masked in gray (small black font). For clarity, the heatmap is divided into three scale-dependent regimes: convective, synoptic, and intraseasonal. The abbreviations Diu, Sys, HFM, LFM, and Ses represent the diurnal, synoptic, high-frequency MISO (10–20 days), low-frequency MISO (20–100 days), and seasonal scales, respectively.

In the **Convective Feedback regime**, the heat map connects sub-daily diurnal errors to macro-scales. Substantial inter-model differences emerge. Over the BoB, ICON exhibits negative correlations between diurnal and lower-frequency biases, indicating that regions with excessive diurnal variability tend to coincide with weaker synoptic and intraseasonal variability. This is consistent with the variance analysis in Fig. 13, where ICON shows a strong positive diurnal variance bias (~180%) but comparatively weaker biases at synoptic and intraseasonal scales. In contrast, both MIROC and IITM-ESM display the opposite tendency, with suppressed diurnal variability occurring alongside enhanced lower-frequency variance over BoB, reflecting their strong amplification of intraseasonal variability over the ocean. Over CI, ICON maintains a well-distributed error budget (+14.5% diurnal bias in Fig-R1) and shows statistically non-significant (greyed-out) correlations in Fig-R2, suggesting limited spatial correspondence between diurnal and lower-frequency error structures over land. These differences highlight the contrasting ways models partition rainfall variability between boundary-layer-driven diurnal convection and organized synoptic–intraseasonal disturbances.

The **Synoptic Dynamics regime** shows predominantly positive correlations across most models and regions. This indicates that locations with excessive (or deficient) synoptic variability generally exhibit similar intraseasonal and seasonal biases, suggesting coherent error structures across lower-frequency modes. A notable exception is ICON over CI, where correlations between synoptic, intraseasonal, and seasonal bias structures are weakly negative (e.g., Syn–Ses = -0.32). Likewise, over the BoB, ICON exhibits substantially weaker positive correlations than the CMIP6 models. These results suggest that lower-frequency errors in ICON are less spatially tied to seasonal rainfall errors than in the other models, contributing to its comparatively balanced multiscale bias structure

The **Intraseasonal Rectification regime** exhibits the strongest and most spatially coherent positive correlations. This suggests that intraseasonal variance biases and seasonal rainfall biases are often geographically co-located. Such behaviour is consistent with one of the central conclusions of this study: seasonal rainfall errors emerge as the cumulative imprint of variance redistribution across modes. Models such as IITM-ESM and MIROC, which exhibit amplified intraseasonal variability over the BoB, also show larger seasonal-scale deviations, whereas models with more balanced intraseasonal variability maintain better seasonal fidelity. A notable exception is ICON, particularly over CI, where weakly negative HFM–Ses and LFM–Ses correlations indicate a much weaker spatial correspondence between intraseasonal and seasonal bias structures than in the CMIP6 models.

Together, these diagnostics directly address the question raised in the abstract regarding how scale-dependent biases shape monsoon variability. Figure 13 quantifies the magnitude of variance redistribution at individual timescales, whereas the new figure identifies the dominant contributors to the total error budget and reveals whether errors at different timescales occur independently or as spatially coherent multiscale structures.

The corresponding **Figure R2 (Figure 15)** and **Figure R1** (**Figure S6 in revised supplementary file**) discussion have been added to **Section 3.3** of the revised manuscript at **line number 682-702**.

2. A little more clarity on the model choices and methodological choices earlier in the manuscript would improve readability. Why ICON? Will it be used for some other purpose soon? Why these 5 CMIP models? Yes they vary across scales, but is this an ensemble of convenience (data availability) or is it a model subset choice? If the latter then justify it. Harmonic analysis is ubiquitous in math and physics, but less common in atmospheric science than perhaps it should be. Add a few references showing examples of how it has usefully been applied to gain insights.

Reply: We thank the reviewer for this helpful suggestion. As per the reviewer's suggestions, we have now incorporated the short discussion in the revised manuscript regarding the selection of ICON **(line number 97-100)**, the CMIP6 models **(line number 170-176)**, and the use of harmonic analysis **(line number 179-183)**.

ICON was selected for two primary reasons. First, most previous evaluations of Indian Summer Monsoon Rainfall (ISMR) have focused on hydrostatic climate models, while assessments using non-hydrostatic global models remain limited. ICON, therefore, provides an opportunity to examine whether next-generation non-hydrostatic dynamical cores offer improvements in simulating multiscale monsoon rainfall variability relative to conventional CMIP6-class models. Second, ICON is increasingly being developed as a unified weather–climate modeling framework capable of kilometre-scale global simulations. Recent developments under projects such as EXCLAIM (Dipankar et al., 2026; Prein et al., 2026) highlight its growing role in future climate prediction systems, making a systematic baseline evaluation of its monsoon performance both timely and relevant.

The five CMIP6 models were chosen mainly based on the availability of 3-hourly AMIP simulations, which are necessary for resolving the diurnal cycle and are available for only a limited number of CMIP6 models through the ESGF archive. In addition, these models have been extensively evaluated in previous monsoon studies and collectively represent a broad range of horizontal resolutions (approximately 50–190 km), dynamical core formulations, and convection parameterization schemes (e.g., Chaudhari et al., 2019; Sperber et al., 2013; Rajendran et al., 2022). The selected set, therefore, provides a structurally diverse ensemble spanning major modeling centers across Europe, North America, Japan, and India, enabling a meaningful comparison of multiscale rainfall biases.

Regarding harmonic analysis, although it is more widely recognized in mathematics and physics, it has a long history of application in atmospheric science, particularly for diagnosing phase-locked variability such as the diurnal cycle of precipitation. Classical studies by Wallace (1975) and Dai et al. (1999) used harmonic decomposition to characterize large-scale diurnal rainfall patterns and the canonical land–ocean contrast. Over the Indian monsoon region, Sahany et al. (2010) and Johnson (2011) employed harmonic analysis to investigate diurnal rainfall propagation and coastal-orographic influences, while Utkarsh et al. (2026) recently applied the technique to examine the role of the low-level jet in modulating diurnal rainfall regimes over the Western Ghats. Tang et al. (2021, 2022) used harmonic decomposition to evaluate diurnal and semi-diurnal precipitation characteristics in CMIP6 models and demonstrated systematic timing biases in simulated rainfall. These studies illustrate the utility of harmonic analysis as a physically interpretable framework for quantifying rainfall amplitude and phase errors, which directly motivates its application in the present study.

References:

Dipankar, A., Bianco, M., Bukenberger, M., Ehrengruber, T., Farabullini, N., Fuhrer, O., Gopal, A., Hupp, D., Jocksch, A., Kellerhals, S.,Kroll, C. A., Lapillonne, X., Leclair, M., Luz, M., Müller, C., Ong, C. R., Osuna, C., Pothapakula, P., Prein, A., Röthlin, M., Sawyer, W., Schär, C., Schemm, S., Serafini, G., Vogt, H., Weber, B., Wills, R. C. J., Gruber, N., and Schulthess, T. C.: Toward exascale climate modelling: a python DSL approach to ICON's (icosahedral non-hydrostatic) dynamical core (icon-exclaim v0.2.0), Geoscientific Model Development, 19, 713–729, https://doi.org/10.5194/gmd-19-713-2026, 2026.

Prein, A. F., Pothapakula, P., Zeman, C., Lalonde, M., and Rixen, M.: From Single Storms to Global Waves: A Global 2.5 km ICON Simulation of Weather and Climate, EGUsphere, 2026, 1–34, https://doi.org/10.5194/egusphere-2025-6414, 2026.

Chaudhari, H. S., Hazra, A., Pokhrel, S., Saha, S. K., and Talluri, S. S.: Simulation of extreme Indian summer monsoon years in Coupled Model Intercomparison Project Phase 5 models: Role of cloud processes, International Journal of Climatology, 39, 901–920, https://doi.org/https://doi.org/10.1002/joc.5851, 2019.

Sperber, K. R., Annamalai, H., Kang, I. S., Kitoh, A., Moise, A., Turner, A., Wang, B., and Zhou, T.: The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century, Climate Dynamics, 41, 2711–2744, https://doi.org/10.1007/s00382-012-1607-6, 2013.

Wallace, J. M.: Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States, Monthly Weather Review, 103, 406–419, 1975.

Dai, A., Giorgi, F., and Trenberth, K. E.: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States, Journal of Geophysical Research, 104, 6377–6402, 1999.

Sahany, S., Venugopal, V., and Nanjundiah, R. S.: Diurnal-scale signatures of monsoon rainfall over the Indian region from TRMM satellite observations, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/https://doi.org/10.1029/2009JD012644, 2010.

Johnson, R. H.: Diurnal Cycle of Monsoon Convection, pp. 257–276, https://doi.org/10.1142/9789814343411_0015, 2011.

Utkarsh, V., Pokhrel, S., Sahoo, P. K., Choudhury, B. A., Yashas, S., Chaudhari, H. S., ... & Saha, S. K. (2026). Low level jet controlled dynamical and thermodynamical regimes of the diurnal cycle of rainfall over the western ghats. Environmental Research Communications, 8(4), 045026.

Tang, S., Gleckler, P., Xie, S., Lee, J., Ahn, M.-S., Covey, C., and Zhang, C.: Evaluating the Diurnal and Semidiurnal Cycle of Precipitation in CMIP6 Models Using Satellite- and Ground-Based Observations, Journal of Climate, 34, 3189 – 3210, https://doi.org/10.1175/JCLI-D-20-0639.1, 2021.

Tang, S., Xie, S., Guo, Z., Hong, S.-Y., Khouider, B., Klocke, D., Köhler, M., Koo, M.-S., Krishna, P. M., Larson, V. E., Park, S., Vaillancourt, P. A., Wang, Y.-C., Yang, J., Daleu, C. L., Homeyer, C. R., Jones, T. R., Malap, N., Neggers, R., Prabhakaran, T., Ramirez, E., Schumacher, C., Tao, C., Bechtold, P., Ma, H.-Y., Neelin, J. D., and Zeng, X.: Long-term single-column model intercomparison of diurnal cycle of precipitation over midlatitude and tropical land, Quarterly Journal of the Royal Meteorological Society, 148, 641–669, https://doi.org/https://doi.org/10.1002/qj.4222, 2022.

3. Well done on a nice manuscript! I enjoyed reading it and learned some interesting things about Indian monsoon rainfall.

Reply: We sincerely thank the reviewer for their kind and encouraging comments. We are pleased that the manuscript was enjoyable to read and provided useful insights into Indian monsoon rainfall. We greatly appreciate this positive feedback and the constructive suggestions that helped to improve the manuscript.

Specific comments:

L27: what does block level mean?

Reply: Here, "block level" refers to a sub-district administrative unit or region in India, typically covering an area of a few hundred square kilometres. Previous studies (Mukhopadhyay et al. 2019; Deshpande et al., 2020) have also used this terminology. For clarity, we have now replaced "block level" with "sub-district level" at **line number 27**.

Reference:

Mukhopadhyay, P., Prasad, V.S., Krishna, R.P.M. et al. Performance of a very high-resolution global forecast system model (GFS T1534) at 12.5 km over the Indian region during the 2016–2017 monsoon seasons. J Earth Syst Sci 128, 155 (2019). https://doi.org/10.1007/s12040-019-1186-6

M. Deshpande, C.J. Johny, R. Kanase, S. Tirkey, S. Sarkar, T. Goswami, K. Roy, M. Ganai, R.P.M. Krishna, V.S. Prasad, P. Mukhopadhyay, V.R. Durai, R.S. Nanjundiah, M. Rajeevan, Implementation of Global Ensemble Forecast System (GEFS) at 12 km Resolution, (2020), IITM Technical Report No.TR-06, ESSO/IITM/MM/TR/02(2020)/200

L31: N. 2005 is a strange way to reference.

Reply: The reference style "N. 2005" was a typographical error. This has been corrected in the revised manuscript and now appears as "(Goswami, 2005)" at **line number 31**.

L45: There is some evidence that monsoon depressions are barotropic: Diaz, M., & Boos, W. R. (2021). The influence of surface heat fluxes on the growth of idealized monsoon. *Journal of the Atmospheric Sciences*. https://doi.org/10.1175/JAS-D-20-0359.1

Reply: We thank the reviewer for this important reference. We agree that attributing monsoon depression growth solely to moist baroclinic instability oversimplifies the current state of understanding. Diaz and Boos (2021) showed that barotropic mechanisms driven by horizontal wind shear within the monsoon trough play an important role in depression intensification, with surface heat fluxes providing additional energetic support. We have revised the text at **line number 45-47** to acknowledge the potential roles of both baroclinic and barotropic processes and added the suggested reference.

L56: no need for etc

Reply: Now we have removed 'etc'.

L75: This is a key point that I think could be made much earlier!! Forgot full stop.

Reply: The text in L75 basically introduces the vertical momentum equation and the contrasting treatment of vertical accelerations in hydrostatic and non-hydrostatic dynamical cores, which provides the specific physical motivation for including ICON in this study.

Introducing this argument earlier, before the review of multiscale monsoon rainfall biases in existing models, would provide limited context for the reader. Its current placement, immediately before the introduction of ICON, allows the need for a non-hydrostatic framework to emerge naturally from the preceding discussion of model limitations and therefore provides a more logical flow of ideas. Therefore, it is retained here.

We have added a full stop in the revised version of the manuscript.

L102: why these 5 models? And what is the motivation for ICON specifically?

Reply: The five CMIP6 models were selected based on availability of 3-hourly AMIP output as well as considering the broad range of horizontal resolutions (50-190 km), dynamical core formulations, and convection parameterization schemes across major modelling centres of the globe.

ICON was selected primarily because non-hydrostatic global models have received limited attention in Indian monsoon studies. ICON provides an opportunity to examine whether non-hydrostatic dynamical cores offer improvements in simulating multiscale monsoon variability, and its ongoing development toward kilometre-scale global simulation under the EXCLAIM project (Dipankar et al., 2026; Prein et al., 2026) makes a systematic baseline evaluation both timely and necessary.

We have also explained it in detail in the 2nd major comment.

L170: can you provide some justification for using harmonic analysis? Some citations or examples of previous use. Pros and cons of this vs other potential methods.

Reply: As also discussed in our response to the Major comment-2 regarding harmonic analysis, here we provide additional justification for its use, together with representative applications from the atmospheric science literature.

Harmonic analysis was selected because it directly decomposes the diurnal rainfall cycle into physically interpretable amplitude and phase components, allowing straightforward quantification of timing and intensity biases across models and observations. This approach has been widely applied in studies of tropical and monsoon precipitation (e.g., Dai et al., 2001; Yang and Slingo, 2001; Nesbitt and Zipser, 2003; Sahany et al., 2010; Tang et al., 2021, 2022; Utkarsh et al., 2026).

Compared with alternative approaches such as EOF analysis, wavelet analysis, or empirical mode decomposition (EMD), harmonic analysis offers a simple and computationally efficient framework for diagnosing climatological diurnal variability. Its principal advantage is the direct estimation of the amplitude and phase of the dominant 24-hour cycle, which are the primary quantities of interest in the present study. We also acknowledge its limitations, including the assumption of stationarity and the inability to fully represent non-sinusoidal or transient variability. However, these limitations do not affect the objectives of the present work, which focuses on the climatological characteristics of the dominant diurnal harmonic and its representation across models.

The corresponding discussion and references have been added to the revised manuscript at **line number 179-183.**

Reference:

Dai, A., 2001: Global Precipitation and Thunderstorm Frequencies. Part II: Diurnal Variations. J. Climate, 14, 1112–1128, https://doi.org/10.1175/1520-0442(2001)014<1112:GPATFP>2.0.CO;2.

Yang, G., and J. Slingo, 2001: The Diurnal Cycle in the Tropics. Mon. Wea. Rev., 129, 784–801, https://doi.org/10.1175/1520-0493(2001)129<0784:TDCITT>2.0.CO;2.

Nesbitt, S. W., and E. J. Zipser, 2003: The Diurnal Cycle of Rainfall and Convective Intensity according to Three Years of TRMM Measurements. J. Climate, 16, 1456–1475, https://doi.org/10.1175/1520-0442(2003)016<1456:TDCORA>2.0.CO;2.

Zhou, T., R. Yu, H. Chen, A. Dai, and Y. Pan, 2008: Summer Precipitation Frequency, Intensity, and Diurnal Cycle over China: A Comparison of Satellite Data with Rain Gauge Observations. J. Climate, 21, 3997–4010, https://doi.org/10.1175/2008JCLI2028.1.

L173: throughout the manuscript, please use the acronyms from Table 1.

Reply: We thank the reviewer for pointing this out. We agree that model naming should be consistent throughout the manuscript. In the revised manuscript, the full CMIP6 model names are provided at first mention and in Table 1. For brevity and to avoid making the text and figure captions unnecessarily word-heavy, we use shortened model names thereafter: CNRM-CM6-1-HR, MPI-ESM1.2-HR, GFDL-CM4, and MIROC6 are referred to as CNRM, MPI, GFDL, and MIROC, respectively. IITM-ESM is retained as IITM-ESM.

We have revised the manuscript at **line number 163-166** and figure captions to ensure that these abbreviations are used consistently.

Fig 2: I wonder how these biases would look as a percentage... I suspect ICON would look better!

Reply: We thank the reviewer for this useful suggestion. Following the reviewer's recommendation, we calculated the mean rainfall biases as percentages relative to the observed climatology. The results indicate that the interpretation depends on the region considered. Over Central India (CI), ICON indeed exhibits the smallest positive percentage bias (~10%), whereas MPI shows the largest positive bias (~168%). However, over the Bay of Bengal (BoB), ICON exhibits the largest positive percentage bias (~66%), while several other models show smaller negative biases. Thus, expressing the biases as percentages does not consistently improve the relative ranking of ICON across all regions.

We have added the percentage bias values to **Table 2** of the revised manuscript. The description is added in **line 262-271.**

Model	**CI**	**BoB**
ICON	0.36 (10.10%)	3.91 (66.36%)
CNRM	-2.00 (-55.49%)	-1.32 (-22.48%)
MPI	6.16 (167.83%)	0.66 (10.90%)
GFDL	-1.52 (-42.76%)	-1.63 (-27.84%)
MIROC	-0.71 (-20.07%)	-1.54 (-25.12%)
IITM	-2.03 (-59.76%)	-2.54 (-42.89%)

L384: full stop.

Reply: Thanks for noticing this typo. We have added a full stop in the revised version of the manuscript.

L577: i can see how this is suggested by results, but its not clearly shown. here, and throughout the manuscript, the link between two or more evaluation metrics could be displayed through a figure (scatter plots for example).

Reply: We thank the reviewer for this suggestion. We agree that the original statement was primarily inferred from the collective results and was not explicitly demonstrated. This concern will be addressed by the already provided two new diagnostics (Fig-R1 and Fig-R2) as discussed in the main comments-1.

Fig-R1 quantifies the contribution of each temporal scale to the total variance bias, thereby identifying the dominant source of error in each model. Fig-R2 explicitly examines the spatial relationship between variance bias structures across timescales through a pattern correlation heatmap organized into Convective Feedback, Synoptic Dynamics, and Intraseasonal Rectification regimes.

Together, these diagnostics show that seasonal monsoon performance is not determined solely by the magnitude of errors at any individual timescale. While ICON exhibits higher standalone RMSE at some synoptic and intraseasonal scales (as shown in Supplementary Fig. S2 of the manuscript), its error does not propagate upscale. It maintains a comparatively balanced variance distribution across land and ocean regions (Fig-R1) and substantially weaker cross-scale error coherence than the CMIP6 models (Fig-R2). For example, over Central India, ICON shows weak or negative spatial correlations between its seasonal bias and higher-frequency structures (Syn–Ses = -0.32; HFM–Ses = -0.29; LFM–Ses = -0.19). Over the Bay of Bengal, these cross-scale correlations are even weaker (0.15, 0.06, and 0.00, respectively). In sharp contrast, CMIP6 models like IITM-ESM and MIROC exhibit strong, coherent positive correlations over CI (typically >0.67, with similarly high values over BoB). This proves that high-frequency errors in those models are structurally locked to and amplify the larger seasonal-scale biases. ICON successfully isolates its high-frequency errors within their respective bands, preventing an upscale cascade.

Now to address the reviewers concern we have integrated this cross-metric discussion into to the revised manuscript and the revised sentence at **L607-613**.

L640: what specifically do you mean by higher-frequency variance? these two statements seem circular, since higher amplitude surely means higher variance.

Reply: Thank you for this important observation. Our intention was not to relate diurnal amplitude to diurnal variance, but rather to discuss how biases at the phase-locked endpoints of the spectrum (diurnal and seasonal) are associated with the distribution of variance across the intermediate synoptic and intraseasonal timescales. To avoid ambiguity, we have revised the text accordingly at **L671-675**.

The revised sentence now reads:

" Taken together, the amplitude diagnostics confirm that errors at the phase-locked endpoints are associated with the distribution of variance across the intermediate synoptic and intraseasonal bands. Models with exaggerated diurnal amplitude often show distorted 2–7-day synoptic and 10–20-day high-frequency MISO variance, whereas amplified seasonal amplitude tends to accompany enhanced lower-frequency intraseasonal variability. Accurate representation of these temporal endpoints is therefore important for maintaining a physically consistent partitioning of monsoon rainfall variability across temporal scales.."

L645: this seems important. i would clarify this earlier, perhaps even in the methodology.

Reply: We thank the reviewer for this valuable observation. Following the reviewer's suggestion, we have now incorporated a brief discussion of the relationship between amplitude and variance in the Methodology section from **L201-205**.

L671: is this shown somewhere? I guess this indicates a larger-scale e.g. circulation (or finer scale e.g. microphysics) control on precipitation?

Reply: We acknowledge that this statement was an assertion not directly demonstrated within the present manuscript. It is, however, well supported by existing literature: Hertwig et al. (2015) showed that increasing horizontal resolution primarily redistributes rainfall between convective and large-scale components without improving precipitation timing or intensity, as the convective parameterisation remains the dominant control at resolutions where deep convection is unresolved. Such resolution-insensitive biases can reflect both larger-scale thermodynamic controls, such as premature convective triggering driven by insufficient CAPE buildup in mass-flux schemes (Bechtold et al., 2004; Rio et al., 2009), and finer-scale microphysical deficiencies in cloud and precipitation processes (Stevens et al., 2013). Isolating the relative contributions of these two pathways would require targeted parameterisation sensitivity experiments beyond the scope of the present study. The new changes are visible from **L723-730** in the revised manuscript.

Reference:

Bechtold, P., Chaboureau, J.-P., Beljaars, A., Betts, A. K., Köhler, M., Miller, M., and Redelsperger, J.-L.: The simulation of the diurnal cycle of convective precipitation over land in a global model, Quarterly Journal of the Royal Meteorological Society, 130, 3119–3137, https://doi.org/https://doi.org/10.1256/qj.03.103, 2004.

Hertwig, E., von Storch, J.-S., Handorf, D., et al.: Effect of horizontal resolution on ECHAM6-AMIP performance, Climate Dynamics, 45, 185–211, https://doi.org/10.1007/s00382-014-2396-x, 2015.

Rio, C., Hourdin, F., Grandpeix, J.-Y., and Lafore, J.-P.: Shifting the diurnal cycle of parameterized deep convection over land, Geophysical 975 Research Letters, 36, https://doi.org/https://doi.org/10.1029/2008GL036779, 2009.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth System Model: ECHAM6, Journal of Advances in Modeling Earth Systems, 5, 146–172, https://doi.org/https://doi.org/10.1002/jame.20015, 2013.

L677-681: this sentence is too long.

Reply: Following the reviewer's suggestion, the original sentence has been divided into shorter, logically distinct sentences to improve readability and clarity. The revised text in **L734-738** now reads:

" The similar early land convection in GFDL is consistent with the behaviour expected from quasi-equilibrium convective closure, which can trigger rainfall before sufficient boundary-layer moistening and instability have developed. Its nocturnal rainfall deficit further suggests limitations in representing cold-pool dynamics and mesoscale lifting processes that help sustain nighttime convection (Zhang et al., 2024). Over the ocean, additional biases may arise from boundary-layer and cloud-microphysical processes that affect radiative balance and marine convection (Donner et al., 2011)."

L694: though here you have prescribed SST so not really well-investigated?

Reply: Yes, reviewer has rightly pointed out that, in the AMIP framework with prescribed monthly SSTs, two-way air–sea feedbacks are suppressed. Therefore, our results do not directly investigate the role of air–sea coupling in shaping diurnal biases. The original phrasing has been revised to avoid the confusion in **L750-753**:

" Together, these results indicate that diurnal rainfall biases are not determined by the parameterization framework alone, but also by its interaction with land–atmosphere processes. Over oceanic regions such as the Bay of Bengal, however, the role of air–sea coupling in modulating diurnal rainfall biases cannot be fully assessed within the prescribed-SST AMIP framework and requires coupled simulations."

L718: credible for what purpose? No parameterised convection scheme will represent squall lines but may still be useful for other applications.

Reply: We thank the reviewer for this important clarification. We agree that the term "credible" was too broad in this context, as model suitability depends on the specific application being considered. Our intention was to emphasize that seasonal-mean rainfall alone does not fully characterize model performance across weather-to-climate timescales.

To clarify this point, we have revised the statement to emphasize the representation of monsoon variability rather than the general credibility of a model. The revised text now reads as in **L778-779**:

"This study demonstrates that realistic simulation of monsoon variability across weather-to-seasonal timescales depends not only on seasonal-mean rainfall accuracy but also on maintaining physically consistent cross-scale organization."

Reply to Reviewer-2

General Comments

The paper evaluates how well a medium-resolution (~40 km) non-hydrostatic global model (ICON, EXCLAIM/GT4Py version) reproduces Indian Summer Monsoon (ISM) rainfall variability compared with five hydrostatic CMIP6 AMIP models (CNRM-CM6-1-HR, MPI-ESM1.2-HR, GFDL-CM4, MIROC6, IITM-ESM; 50–190 km). The reference is GPM-IMERG V07. The topic is relevant and important to the model development community. However, the current version of the manuscript has some flaws that must be addressed before publication. I list my concerns below.

We sincerely thank the reviewer for the constructive comments and insightful suggestions, which have significantly improved the quality of this manuscript. Below, we have provided a detailed, point-by-point response to each of the comments in **blue font** and correction made in the revised manuscript with line numbers in **red font**.

Specific Comments

1. The abstract specifies 1998-2014 as the analysis period. However, in the caption of Figure 3, it is mentioned that the analysis period is 1979-2014. Is there an inconsistency in the analysis using different datasets?

Reply: We thank the reviewer for pointing out this ambiguity. There is no inconsistency in the analysis, but we agree that the earlier wording of the Figure 3 caption could create confusion.

The model bias analyses presented in the manuscript are calculated using the common period of availability between the model simulations and IMERG, i.e., 1998–2014. This common period is used to ensure a consistent comparison between observations and models.

Figure 3, however, was originally intended to show the climatological evening-minus-morning rainfall pattern as a background diagnostic. For this climatological figure, we used the full available period for each dataset: 1979–2014 for ICON and the CMIP6 models, and 1998 to 2023 for IMERG. The purpose of this figure was to present the broad observed and modelled climatological structure, not to quantify model bias.

To address the reviewer's concern, we have now also plotted the same diagnostic for the common period of 1998–2014 (Fig. R1). The overall spatial patterns remain largely unchanged compared with the full-period climatology, as also reflected in the pattern-correlation statistics. Although the amplitude is slightly enhanced or reduced in some specific regions and models, the broad evening-minus-morning rainfall structure remains consistent. This confirms that the main climatological interpretation is robust to the choice of period.

To avoid further confusion, we have revised the Figure 3 caption to clearly state the period used for each dataset. We have also modified the relevant text in the manuscript to clarify that all model bias calculations are based on the common IMERG–model overlap period of 1998–2014. The newly added common-period figure and its explanation are included to demonstrate that the climatological evening-minus-morning rainfall structure remains broadly consistent.

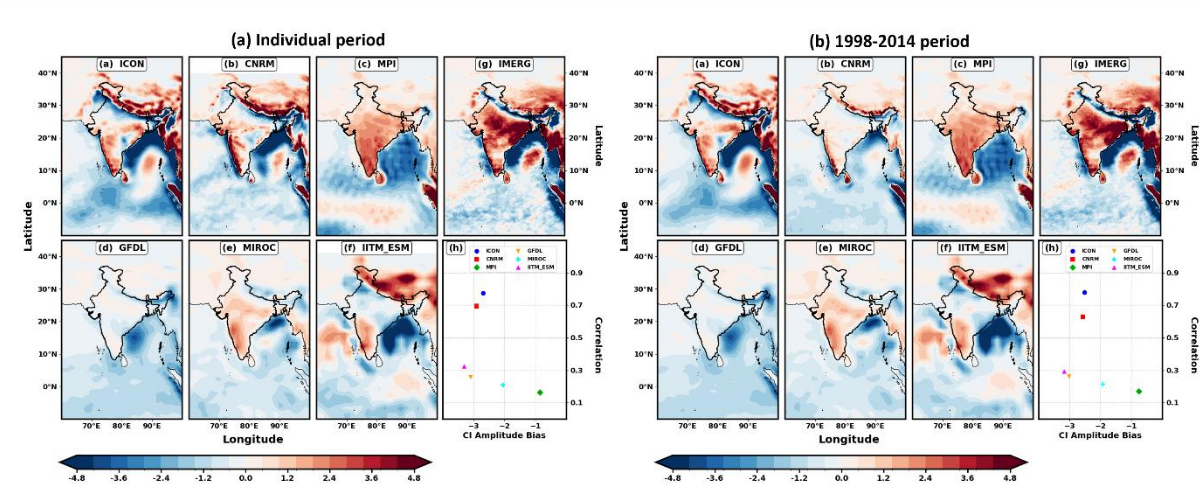

Figure R1. Climatological mean JJAS evening (1200–2330 LT) minus morning (0000–1130 LT) precipitation for the period (a) 1979–2014 (IMERG: 1998-2024) and (b) common period (1998-2014), normalized by the JJAS mean rainfall, derived from (a-f) ICON, CNRM, MPI, GFDL, MIROC, and IITM-ESM model simulations. Panel (g) shows IMERG observations, while panel (h) shows the pattern correlation of each model with IMERG, and the Central India (CI) mean amplitude difference between models and IMERG (Model-IMERG). Precipitation is expressed in mm/day.

The discussion has been added to the revised manuscript **at lines 173-176, 272-273** and in the **caption of Figure 3**

2. The major claim by the authors is that ICON (non-hydrostatic; 40 km resolution) is the best for the monsoon. The ICON model is compared with the AMIP simulations of a subset of CMIP6 models, which are much coarser in resolution. The gap between the resolutions of ICON and other models ranges from ~1.25 to ~5 times. This difference in resolution makes the comparison unfair. A fair comparison can be made with the hydrostatic version of the ICON model by keeping the resolution the same. Alternatively, the authors can use HighResMIP simulations.

Reply: We sincerely thank the reviewer for this important and constructive comment. We agree that differences in horizontal resolution can influence the model performance. However, the objective of the present study is not to isolate the effect of resolution alone, but to evaluate how different available model configurations represent the multiscale characteristics of Indian monsoon rainfall. The models analysed in this study differ not only in horizontal resolution, but also in dynamical cores, convective parameterization schemes, and other physical processes. Therefore, our comparison is intended as a model-configuration assessment rather than a controlled resolution-sensitivity experiment.

We also note that the models used in this study do not form a simple “high-resolution versus low-resolution” contrast. They include a range of resolutions, with ICON, CNRM, and MPI lying in the relatively finer-resolution group, while GFDL, MIROC6, and IITM-ESM are comparatively coarser. In particular, CNRM with 50 km grid resolution is close to ICON at 40 km. Thus, the comparison is not solely between ICON and much coarser models. Moreover, our results show that resolution alone does not fully explain model performance. Some relatively high-resolution models still show substantial errors in amplitude, phase, and spatial organization, whereas some coarser models perform reasonably well in selected diagnostics.

This interpretation is consistent with recent ICON sensitivity experiments over global monsoon regions (Pothapakula et al., 2026). They examined ICON simulations at 80, 40, and 10 km grid spacing and showed that the effect of horizontal resolution on monsoon simulation is region-dependent and non-systematic. Their study found that finer grid spacing does not uniformly improve monsoon rainfall simulations; while some aspects such as the diurnal cycle may improve, biases in mean precipitation, intense rainfall frequency, and intraseasonal variability can be amplified. This supports our view that horizontal resolution is important, but it is not the only factor controlling monsoon model performance. Model physics, convective parameterization, microphysical processes, and the representation of circulation–rainfall coupling are also critical.

We fully agree with the reviewer that the manuscript should avoid giving the impression that ICON is superior or that its better performance can be attributed only to its non-hydrostatic dynamical core. We have therefore revised the manuscript to soften this claim. In the revised manuscript section-4 (Discussions and summary section) it is explicitly mentioned that ICON is the best-performing model configuration among the models analysed for several diagnostics, particularly the spatial organization, phase, propagation of diurnal rainfall, and the annual cycle, rather than as “the best model for the monsoon simulation.”

We also acknowledge reviewer’s suggestion that a fully controlled comparison would require simulations with comparable resolution and model formulation, such as hydrostatic and non-hydrostatic versions of ICON at the same grid spacing, or a dedicated HighResMIP-based analysis. While such an investigation would provide valuable insights into the individual roles of resolution and dynamical core formulation, it is beyond the scope of the present study. Accordingly, the revised manuscript now presents ICON’s performance in a more balanced way and clearly states that the present analysis is not a controlled resolution-sensitivity experiment, but a comparative assessment of available model configurations.

References

Pothapakula, P. K., Prein, A. F., Sunkisala, A., and Dipankar, A.: Global monsoon in ICON: the scale-dependent response of Northern Hemisphere monsoons, Weather Clim. Dynam., 7, 979–1007, https://doi.org/10.5194/wcd-7-979-2026, 2026.

The discussion has been added to the revised manuscript **at lines 97-100, 228, 316 and 779-785.**

3. The PDFs are computed on the native grids of each model (section 3.2.2, Figure 9). The rainfall PDFs are heavily dependent on resolution. In line with the specific comment #2, this is also an unfair comparison and a misleading analysis (coarser grids smooth extremes).

Reply: We thank the reviewer for the important comment. The purpose of presenting PDFs on the native grids was to examine how each model configuration represents rainfall intensity distribution in its own effective resolution and formulation. Since one of the objectives of this analysis is to assess the ability of each model configuration to simulate rainfall intensity, including heavier rainfall events, regridding all datasets to a much coarser common grid could suppress the high-intensity tail and may mask part of the model's actual behaviour at its native resolution.

However, to address this concern, we have now repeated the PDF analysis after regridding IMERG and all model outputs to a common grid, corresponding to the IITM-ESM resolution of about 1.9°. The revised figure now shows PDFs both on the native grids and on the common IITM-ESM grid (Fig-R2).

The comparison shows that regridding modifies the rainfall distributions substantially. On the common grid, the IMERG PDF becomes broader than all models. IITM-ESM, which underestimated light rainfall at its native grid, now overestimates light rainfall, similar to the other models. Among the coarser-resolution models, the order of light-rainfall overestimation remains broadly consistent, with GFDL showing the largest overestimation, followed by MIROC and IITM-ESM. For the finer-resolution models, however, the order changes from CNRM–ICON–MPI on the native grid to MPI–CNRM–ICON on the common grid, indicating that part of the light-rainfall behaviour is resolution dependent.

For moderate rainfall, the common-grid analysis shows improved agreement with IMERG. In the 12.5–25 mm h^{-1} range, ICON closely matches the IMERG PDF, whereas it underestimated this range on the native grid. The coarser models also show better agreement with IMERG at different parts of the moderate rainfall range after regridding. However, MPI and CNRM continue to underestimate moderate rainfall even on the common grid. For heavy rainfall in the 60–100 mm h^{-1} range, all models still underestimate the IMERG PDF, although the magnitude of underestimation is reduced considerably after regridding.

These results confirm that rainfall PDFs are resolution sensitive. At the same time, the common-grid analysis shows that important inter-model differences remain even after resolution normalization. We have therefore revised the manuscript to interpret the native-grid PDFs as model-configuration-specific distributions and the common-grid PDFs as a more resolution-consistent comparison. The discussion has also been modified to avoid using native-grid PDFs alone as a direct ranking of model skill in simulating extreme rainfall.

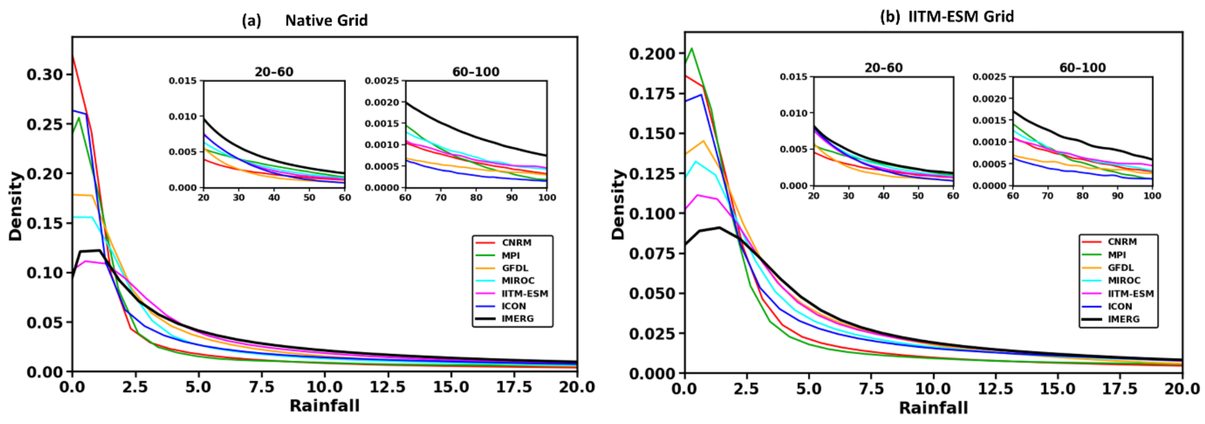

Figure R2: Probability density function (PDF) of JJAS daily rainfall over Central India from ICON and CMIP6 models (CNRM, MPI, GFDL, MIROC, and IITM-ESM), compared with IMERG (a) for native grid and (b) common grid. The main panel shows 0–20 mmday^{-1}, with inset panels extending to 20–60 mm/day and 60–100 mm/day.

The discussion has been added to the revised manuscript **at lines 466-472, 481, 482-484, 486, Figure 9 caption** and **Figure R2 (b)** is added in **the revised supplementary file** as **Figure S2**

4. The claim that 20–100-d ISO underestimation reflects missing air–sea coupling is plausible but unverified here; a coupled run of any of these models is not included.

Reply: We agree with the reviewer in this important observation. Our statement was intended to provide a possible physical explanation, since air–sea interaction is known to influence intraseasonal variability of the monsoon (Sperber et al 1997, Waliser et al 1999, Kemball-Cook and Wang 2001, Goswami 2005, etc.). However, we agree that this interpretation was too strong in the absence of coupled simulations.

We have therefore revised the manuscript to present this point more cautiously. In the revised text, we now state that the underestimation of 20–100-day ISO variability in AMIP simulations may be partly related to the absence of interactive air–sea coupling, but this cannot be confirmed within the present experimental framework. We have also clarified that testing this hypothesis would require comparison with coupled simulations, which is beyond the scope of the present study.

References

Goswami, B. N., 2005: South Asian monsoon. In: W. K. M. Lau and D. E. Waliser (Eds.), *Intraseasonal Variability in the Atmosphere–Ocean Climate System*, Springer/Praxis, Berlin/Heidelberg, pp. 19–61.

Kemball-Cook, S., and B. Wang, 2001: Equatorial Waves and Air–Sea Interaction in the Boreal Summer Intraseasonal Oscillation. *J. Climate*, **14**, 2923–2942, https://doi.org/10.1175/1520-0442(2001)014<2923:EWAASI>2.0.CO;2.

Sperber, K. R., J. M. Slingo, P. M. Inness, and K. M. Lau, 1997: On the maintenance and initiation of the intraseasonal oscillation in the NCEP/NCAR Reanalysis and the GLA and UKMO AMIP simulations. *Climate Dyn.,* **13,** 769–795.

Waliser, D. E., K. M. Lau, and J. Kim, 1999: The Influence of Coupled Sea Surface Temperatures on the Madden–Julian Oscillation: A Model Perturbation Experiment. *J. Atmos. Sci.*, **56**, 333–358, https://doi.org/10.1175/1520-0469(1999)056<0333:TIOCSS>2.0.CO;2.

The discussion has been added to the revised manuscript **at lines 763-766**

5. The discussion repeatedly highlights ICON's strengths while downplaying its weaknesses. E.g., a 180% variance bias and 60% amplitude bias over BoB (substantially worse than several CMIP6 models in absolute terms) is acknowledged but minimised as "over-responsive high-frequency convective adjustment."

Reply: We thank the reviewer for this important comment. In line with reviewer's suggestion, we have substantially modified the discussion in the revised manuscript to present a more balanced assessment of ICON's performance. We now explicitly state that although ICON performs well in several diagnostics related to the spatial organization, phase, and propagation of diurnal rainfall, it also shows large positive biases in diurnal amplitude and variance over the Bay of Bengal. The approximately 60% amplitude bias and 180% variance bias are now discussed as important model deficiencies rather than as minor limitations.

Now the revised discussion clearly distinguishes between different aspects of model performance: ICON shows better agreement with observations in the structure and timing of diurnal rainfall, but it does not consistently outperform all CMIP6 models in terms of amplitude and variance. In some diagnostics, particularly over the Bay of Bengal, several CMIP6 models show smaller absolute biases than ICON.

The discussion has been added to the revised manuscript **at lines 268-271**

Minor Comments

1. L324: Wrong figure reference. Fig. 5g is IMERG

Reply: Thanks for pointing this typo error, now the Figure referencing is corrected for MIROC (Fig 5e).

2. L31: Citation "(N., 2005)" appears malformed.

Reply: The reference style "N. 2005" was a typographical error. This has been corrected in the revised manuscript and now appears as "(Goswami, 2005)".