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**. We have diligently incorporated all the suggestions and comments in the revised manuscript.

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 2024 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.

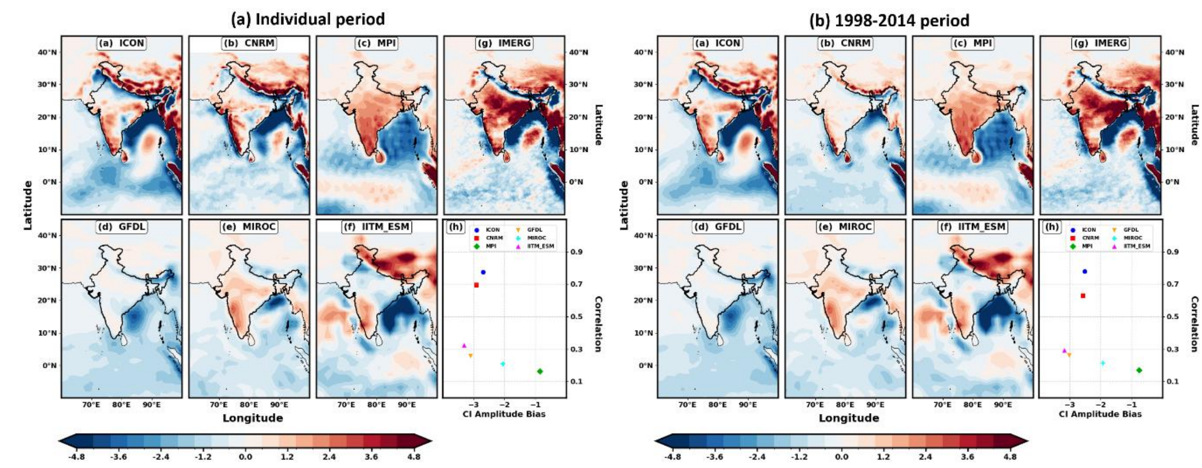

Figure R1. Climatological mean JJAS evening (1200–2330 LT) minus morning (0000–1130 LT) precipitation for the period (a) 1979–2014 (IMERG: 1998-2024) and (b) common period (1998-2014), normalized by the JJAS mean rainfall, derived from (a-f) ICON, CNRM, MPI, GFDL, MIROC, and IITM-ESM model simulations. 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.

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.

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.

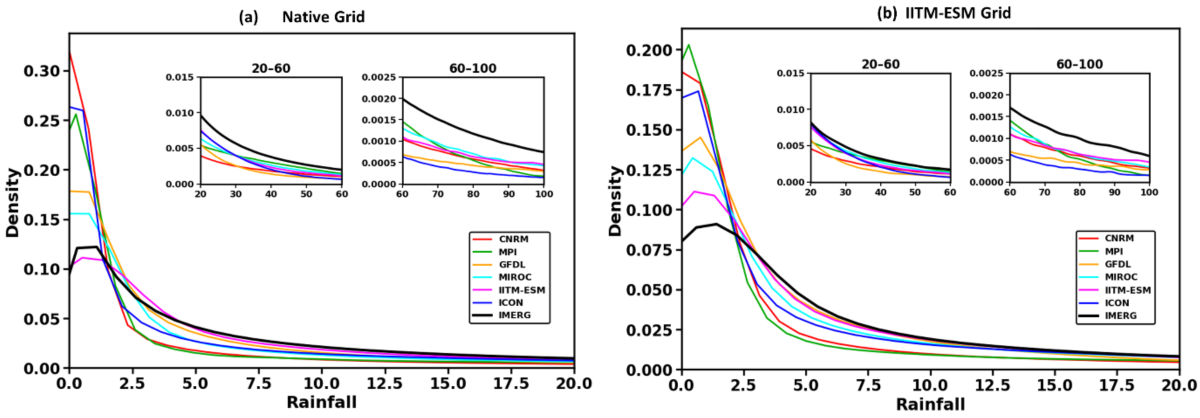

Figure R2: Probability density function (PDF) of JJAS daily rainfall over Central India from ICON and CMIP6 models (CNRM, MPI,GFDL, MIROC, and IITM-ESM), compared with IMERG (a) for native grid and (b) common grid. The main panel shows 0–20 mmday−1, with inset panels extending to 20–60 mm/day and 60–100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

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.

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.

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)".