Response to Reviewer-2

Review of: Global Monsoon in ICON: The scale-dependent response of Northern Hemisphere Monsoons

Authors: Praveen Kumar Pothapakula*, Andreas F. Prein, Anusha Sunkisala, and Anurag Dipankar

Journal: Weather and Climate Dynamics

Review: This contribution by Pothapakula et al. is a comprehensive evaluation study of ICON in a new version. It focuses on important aspects of the global monsoon system and concludes, that grid-spacing still influences the climate modeling performance. It is an interesting and important contribution, which I find well worth publication after revisions.

We thank the reviewer for the positive comments on our manuscript and for suggesting important points and clarifications. All the comments are addressed in detail, and furthermore, sensitivity tests were conducted to answer comments where necessary. Please find our point-by-point reply below

Still, I have a few comments and questions as detailed below.

1. The abstract is quite long and it is difficult to identify the main findings. Is it a goal to evaluate the new ICON core (line 7, but not mentioned below) or just the grid-spacing dependence of global/regional monsoon representation? What is "organized moderate variability"? What is an "excellent representation"? Line 23: increased → decreased. Line 25: What "underscores" the region-dependent sensitivity?

We agree with the reviewer, and now we have made the abstract shorter and focused, highlighting important findings of our manuscript.

Modified Abstract: The global monsoon system is a lifeline for two-thirds of the world's population, as it is essential for tropical water security, food, and agriculture. However, its complex multiscale interactions challenge weather and climate models. This study investigates how horizontal grid spacing (80 km, 40 km and 10 km) in the ICOsahedral Non-hydrostatic (ICON) model affects both the mean and the variability of Northern Hemisphere monsoons across diurnal, intraseasonal, and interannual timescales.

All simulations show substantial skill in capturing the global monsoon system domain and its mean annual range of precipitation with a pattern correlation of > 0.7 and RMSE < 3 mm/day. For the key Northern Hemisphere regional monsoons, South Asia (SAsiaM), West Africa (WAfriM) and North America (NAmerM), ICON achieves an accuracy > 80% in capturing the observed monsoon domain. Crucially, the impact of grid spacing is strongly region-dependent and non-systematic. The finer grid spacing induces higher mean precipitation biases over continental SAsiaM, and WAfriM. Some of these biases are related to the intensity and location of moist monsoonal low-level jets, as well as their sensitivity to grid spacing. Furthermore, the fine grid spacing overestimates monsoon precipitation variability at interannual and intraseasonal scales, including intense precipitation frequency (> 10 mm/day). This amplification stems primarily from enhanced grid-scale precipitation resulting from efficient microphysical processes, while convective precipitation exhibits limited sensitivity to grid spacing. Over NAmerM, biases are smaller and show minimal sensitivity to model grid spacing.

Increased intraseasonal variance (2-30 day band) in the 10 km simulation is linked to more intense low-pressure synoptic systems over SAsiaM and intense African easterly wave activity over WAfriM. All simulations agree on the diurnal precipitation peak timing, with the 10 km simulation marginally performing better over continents.

Our results demonstrate that fine grid spacing alone does not uniformly improve monsoon simulations. Some features, such as the precipitation diurnal cycle, are improved while existing biases in mean precipitation and variability are enhanced. This underscores the role of region-dependent sensitivity of grid spacing governing monsoon dynamics.

Questions:

(a) Is it a goal to evaluate the new ICON core (line 7, but not mentioned below) or just the grid-spacing dependence of global/regional monsoon representation?

This work aims to: (a) validate the performance of the ICON model for global and regional monsoon systems, and (b) identify the sensitivity of monsoon representation to horizontal grid spacing.

The refactoring of the dynamical core to GT4py is merely a code reimplementation for computational efficiency and does not include any physical or dynamical changes from the original Fortran-based ICON version documented in Dipankar et al. (2026). Therefore, evaluating the GT4py core *per se* is not a goal of this study. Rather, we use the GT4py version as the computational platform to conduct our monsoon validation and grid-spacing sensitivity experiments.

To the authors' knowledge, this is the first extensive validation of the ICON model (in any configuration) for global and regional monsoon studies. We have now removed the "ICON GT4py" statement from the abstract to avoid confusion and to focus clearly on the validation and grid-spacing analysis. A description of the GT4py ICON code is provided in the methodology section for completeness, but the scientific conclusions do not depend on the specific code implementation.

(b) What is organized moderate variability: Our study shows that the variability associated with intraseasonal scale (30-90 day and 2-30 day) bands increasees with fine grid spacing. To determine if intense precipitation in fine grid spacing simulations disproportionately influences the variability estimates, we applied a log(1+P) transformation to all precipitation (P) datasets prior to band-pass filtering. This transformation reduces the statistical leverage of intense rainfall, while preserving the timing of precipitation events.

After transformation, the systematic increase of variance persists across both the 30--90 and 2--30 day bands over SAsiaM and WAfriM domains. Crucially, the transformed variance ratios are below unity for 2--30 day, revealing that all simulations underestimate organized precipitation variability when intense precipitation is statistically controlled, while the ratio of 30--90 day in fine grid spacing remains near unity.

We define **"organized moderate variability"** as the precipitation variability associated with statistically controlling (i.e., down-weighting) intense precipitation events, which largely originate from convective systems. This distinction allows us to separate variability contributed by extreme events from more moderate, organized precipitation processes.

For the abstract to remain focused and concise, we have removed the statement about statistically controlled precipitation transformation. A detailed discussion is provided in Section 3.4 of the manuscript.

(c) Excellent representation of diurnal cycle?: Now the statement in the abstract reads as "All simulations agree on the diurnal precipitation peak timing, with the 10 km simulation marginally performing better over continents. (L24-L25)

(d) We have corrected the increasing grid spacing to fine grid spacing (L25).

2. Has the simple linear interpolation (line 120) of monthly SST data to daily data some impact on monsoon onset/offset? In the pre-monsoon phase this should slightly underestimate SST warming and this slightly increases the land-sea contrast. Perhaps this can explain a few days of monsoon onset/offset bias?

Thank you for this very important suggestion. We completely agree with the reviewer, and accordingly conducted a sensitivity experiment to address this point.

In the manuscript, we used sea surface temperatures from the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for the AMIP simulations. For a sensitivity experiment, we used the ESA Climate Change Initiative (ESA-CCI) daily SST product, whereas the AMIP SST is available at a monthly resolution. Figure 1 shows the absolute SST magnitude over the extended SasiaM (Fig. 1a-c), WAfriM (Fig. 1d-f), and NAmerM (Fig. 1g-i) regions, along with the difference between AMIP and ESA datasets during the pre-monsoon season (MAM). Over the Indian Ocean, AMIP SST is approximately 0.5–1°C colder compared to ESA, with similar cooling observed over the tropical Atlantic off the coast of NAfriM. Over the Pacific, the differences are mixed, while AMIP is colder in the western Atlantic.

As the reviewer pointed out, the underestimation of AMIP SST compared to ESA during the pre-monsoon period may be further enhanced by the interpolation from the monthly to daily scale. This would likely increase the land-sea contrast during the monsoon initiation phase and could potentially lead to an early onset. This is particularly relevant over SasiaM and NafriM due to their strong dependence on surrounding oceans, specifically the land-ocean temperature contrast. Figure 2 confirms this expected behavior. Over the SasiaM domain, especially over the eastern Indian subcontinent, the daily-updated (ESA SST) ICON simulation shows a delay in monsoon onset of approximately 6–10 days (Fig. 2d) compared to the monthly-updated (AMIP SST) simulation. This is also true for the WafriM domain, where the daily-updated SST

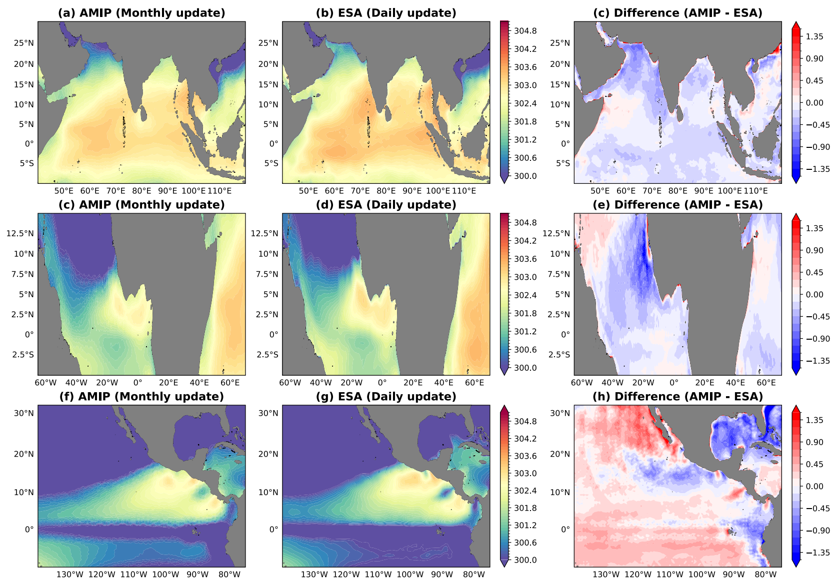

Figure 1: Sea surface temperature (SST) boundary conditions during the pre-monsoon season used in the ICON reference simulation and a sensitivity experiment. (a) Monthly AMIP SST (reference simulation) over South Asia. (b) Daily ESA SST (sensitivity experiment) over South Asia. (c) Spatial difference in SST (AMIP minus ESA) over South Asia. (d-f) Same as (a-c), but for the Africa region. (g-i) Same as (a-c), but for the North America region.

simulation shows a delayed south-to-north propagation, particularly over the western Sahel region. We would like to bring to the attention of the reviewer that, due to computational constraints, our sensitivity experiment only includes 4 simulated years.

These results confirm that our ICON monsoon onset simulations in the manuscript are sensitive to the monthly-updated SST boundary conditions. However, since our comparison of grid spacing sensitivity (10 km, 40 km, and 80 km) uses the same SST boundary dataset, our results remain valid for quantifying the role of grid spacing. We have noted this sensitivity in the manuscript in L364-L366.

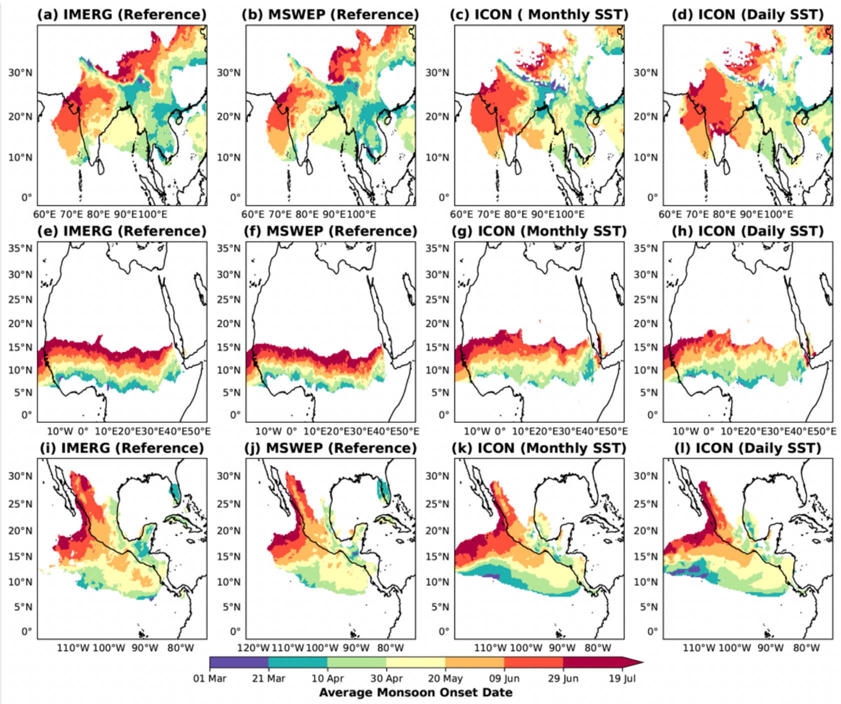

Figure 2: Spatial patterns of monsoon onset dates over the (a-d) SAsiaM , (e-h) WAfriM and the (i-l) NAmerM domain in IMERG, MSWEP and two ICON simulations. The Monthly SST update is the reference ICON simulation used in the manuscript, while the daily SST update is a sensitive experiment with ICON using the ESA high resolution dataset (10 km).

3. Significance: (line 330) “indicates that 40 km performs best”. This statement is based on a relativly small number of simulation years and the distributions shown in the violin plots (Fig. 4) are well overlapping. I assume here there is no conclusion about best grid-spacing possible.

Thank you for bringing this to our attention. The reviewer is correct that this statement is based on a relatively short simulation period (10 years) for quantifying monsoon onset bias. However, while acknowledging the sample size, we conducted significance tests, and the results shown in the violin plot are statistically significant.

Nevertheless, we agree with the reviewer that concluding the ICON 40 km run performs "best" is too strong a statement, given that the onset median (bias) for the 10 km, 40 km, and 80 km runs are 7 (7), 3 (5), and 12 (12) days, respectively, differences that are relatively small. We have therefore rephrased this sentence as follows:

“This indicates that the 10 km and 40 km run marginally improve the onset bias compared to 80 km. The difference in the onset bias is very small between 10 km and 40 km grid spacings, highlighting a non-monotonic relationship with grid spacing”. (L339-L341)

Furthermore, in the manuscript, we have highlighted that the grid spacing sensitivity is non-monotonic, and thus we did not intend to identify a clear winner among the ICON simulations.

4. Why do you exclude the southern hemispheric monsoons from the evaluation, but show them in the Figs?

We have acknowledged in the manuscript (L280) that we focus only on the tropical monsoon systems in the Northern Hemisphere (NH) for this study. We agree with the reviewer that Fig. 1 and Fig. 2 show the performance of Southern Hemisphere (SH) monsoons in addition to NH monsoons.

However, we would like to emphasize that our inclusion of both NH and SH domains is solely to demonstrate the representation of monsoon domains in terms of spatial extent and to quantify spatial correlation, RMSE, and skill scores. Including a detailed discussion of SH monsoon characteristics, such as onset, variability, and process-oriented understanding, would significantly complicate the manuscript and make it excessively lengthy. As the reviewer can see, the manuscript already provides a very detailed analysis of the NH monsoon with extensive supporting details.

While we sincerely wanted to include SH discussions, the manuscript already struggles with maintaining a focused and concise narrative. In our next detailed study, we will provide results and descriptions for the SH monsoon. For the benefit of readers, we have emphasized in L273-L275 that:

"For the purposes of this article, we focus on tropical monsoon systems in the Northern Hemisphere, but we show a preliminary investigation of monsoon domain spatial extent from various ICON simulations for Southern Hemisphere monsoons. A detailed investigation of the characteristics of SH monsoons is, however, beyond the scope of this article." (L281-L283)

5. l632: I guess, here decreased not increased grid-spacing is meant? Does this worsening performance of precipitation simulation over ocean imply that the there is a tuning issues in the model? Later (l647) you state that the precipitation contribution by the convection parameterization independent of the chosen grid-spacing. The parameterization should contribute less at 10 km grid-spacing.

Yes, the reviewer is correct to point out that in we wrote "increased grid spacing" instead of "decreased grid spacing." We have now corrected this sentence to state that fine grid spacing leads to precipitation biases in oceanic regions (L661).

Furthermore, we agree with the reviewer regarding model tuning issues. The current setup is built on the previous expertise of the authors' model setup, mainly for regional simulations (Pothapakula et al., 2020; Soerlund et al., 2020). The authors have not performed any tuning experiments for the simulations used in the current manuscript. Systematic tuning requires a large consortium effort, and unfortunately, we have limited resources. Given the decent performance of ICON, the authors have documented its skill in capturing global monsoon systems.

Nevertheless, we plan to adopt objective tuning using a Linear Meta-Model optimisation (LiMMo) to derive an optimized model configuration in future studies. Specifically, we aim to tune the most important parameters for the surface energy budget, which affect turbulent heat fluxes. These parameters include the resistances to the fluxes in the laminar boundary layer: rlam_heat, cr_bsmin, rat_lam, rat_sea, and rsmin_fac in ICON. (We kindly refer the reviewer to lines 193–195 in the manuscript by the ICON-CLM community: https://doi.org/10.5194/egusphere-2025-4726)

Regarding the contribution from the deep and shallow convection parameterization scheme, we agree with the reviewer that the 10 km grid spacing should ideally contribute less to the total precipitation. In response to Reviewer 1, we performed a detailed analysis of precipitation partitioning between convective and grid-scale (resolved) precipitation (we kindly refer the reviewer to Response to Reviewer 1, Point 5). Over the monsoon core regions, we notice that the 10 km convective scheme contributes relatively less compared to the coarse grid simulation; however, it is important to note that the rate of increase in grid-scale precipitation is not proportional to the rate of decrease in convective precipitation. This is also very true over oceanic regions, where the contribution from stratiform precipitation dominates while the rate of decrease from the convection scheme is slow. Hence, we see such strong biases over the oceans.

Furthermore, rain conversion, especially in the lower troposphere, is very efficient, indicating efficient microphysical processes in the fine grid spacing simulation (10 km) compared to the coarser simulation (80 km) (detailed analysis in response to reviewer 1, point 5). We believe these factors contribute to excessive precipitation in fine grid spacing simulations, both in core regions and over oceans. This indicates a need for tuning efforts in turbulence, microphysics, and boundary layer process.

Minor issues:

l133: Tergen → Tegen

We apologize for this typo. Now this issue is fixed. (L137)

l165: "V" is horizontal (add) wind vector (and thus should be bold). Also here and in the following "kg/m-s" is uncommon. Correct would be kg/(ms) but we are used to kg/ms too. At line 441,2 you use inconsistently but correctly kg m-1 s-1.

Thank you for assisting us in getting the units correct. We have now kg/ms throughout the manuscript, and corrected the units. Also, the V is turned to bold to represent a vector.

l289: "skill scores" – usually a reference prediction/simulation (like a random, persistent ... simulation) is used in the definition of skill scores. In this sense you do not consider skill scores, just scores.

We thank the reviewer for this comment. While we acknowledge the strict predictive definition of "skill score," our study is a climatological evaluation (10-year simulation vs. observed climatology), not an operational prediction. In the climatological literature, metrics such as POD, FAR, and Bias Score are routinely referred to as "skill scores" when evaluating model performance against observed climatology. To avoid any potential confusion, we have clarified in L289 that we use *verification metrics* (POD, FAR, Bias Score, Accuracy) to evaluate model skill in capturing monsoon domain extent. (updated in L162)

l325: "common bias" – references?

Thank you, we have now added this essential reference.

Senan, Retish, et al. "Impact of springtime Himalayan–Tibetan Plateau snowpack on the onset of the Indian summer monsoon in coupled seasonal forecasts." *Climate Dynamics*47.9 (2016):

l389: dipole → dipole bias?

Thank you, corrrected.

l656: "... variability .... show_s"

The sentence is corrected.

l837: incomplete reference

Now the reference is complete.