General overview

In this manuscript, the authors evaluate the performance of the RegCM5 regional climate model in simulating air–sea fluxes over the Southeast Asian seas. The model was run at a 25 km resolution for the year 2018 using 36 different combinations of physical parameterization schemes, selecting from multiple options for convection, microphysics, planetary boundary layer (PBL), radiation, and cloud fraction. Atmospheric forcing was provided by ERA5 reanalysis at 0.25° resolution, while sea surface temperatures were obtained from the high-resolution SYMPHONIE ocean model running at approximately 0.083° resolution. Model outputs, such as precipitation, surface radiation, latent and sensible heat fluxes, and sea surface wind speed, were evaluated against satellite and reanalysis data. A multi-criteria decision-making framework, incorporating 180 performance metrics across eight oceanic subregions, was used to rank the experiments. The results indicate that the top-performing configuration is a combination of the RRTM radiative transfer scheme, UW-PBL planetary boundary layer, Tiedtke cumulus convection, SUBEX resolved-scale microphysics, and Xu–Randall cloud fraction (identified as 12511, i.e. RRTM/UW-PBL/Tiedtke/SUBEX/Xu–Randall), with the Tiedtke cumulus convection scheme consistently outperforming others, particularly in simulating precipitation and wind. The findings highlight cumulus convection as the primary driver of model performance and suggest that the optimal physical parameterizations may vary depending on the variable of interest (e.g., precipitation vs. shortwave radiation). The manuscript is well written, logically structured, and easy to follow, making it a worthy candidate for publication in Geoscientific Model Development. However, there are some points that need to be clarified.

Thank you for your careful reading, understanding and appreciation of our proposed manuscript. We carefully read your comments and answer them down below (line numbers refer to the updated manuscript).

Comments

1. First, the authors' use of simulation results from ***only one neutral year*** *(2018) to evaluate the model's performance is not sufficiently convincing. A single-year simulation provides only one monthly and annual value per grid point for each variable, which introduces substantial uncertainty into the performance assessment due to the lack of statistical robustness. Furthermore, by excluding years influenced by major climate variability phenomena such as ENSO and IOD, the evaluation overlooks the model's capacity to simulate responses under extreme conditions, one of the key strengths of dynamical models. As a result, the findings may be overfitted to neutral conditions and may not adequately reflect the model's robustness or broader applicability across different climate regimes.*

Thank you for bringing this up. As a first answer, we want to highlight that this article represents for us one early brick of a broader initiative relating to evaluating and tuning RegCM over SEA, and that in this context, conducting experiments over several years including both neutral and non neutral conditions is definitely something we have in mind and will apply in the future. Yet, we think that this article's assessment based on a single year was necessary to reduce the ensemble of candidate configurations before going further with more experiments. Those future experiments should be conducted with a greater simulation period indeed, but also potentially address other aspects of performance (as also mentioned when addressing your second point below). In this study, we can only highlight the uncertainties bound to our protocol and the implications in terms of how to interpret and use our results. This is done in the conclusions of the new manuscript version:

L599 "Several limitations are yet to be mentioned. Indeed, our strategy targeted the seasonal cycle, assessed over a single year selected for its neutrality with respect to large-scale oscillations. Consequently, our findings are specific to that context, and no conclusions can be drawn about the performance of this study's top-performing configurations under non-neutral conditions, in terms of intraseasonal/inter-annual variability or more. While this work provides a basis for identifying a subset of promising configurations, additional experiments are needed to further refine the selection, notably involving longer simulations to conduct more comprehensive diagnostics."

2. This study is highly valuable for advancing our understanding of air–sea coupling and for supporting the development of coupled models. However, in many practical applications, the accurate simulation of ***precipitation and temperature over land is even more critical****. In fact, coupled models are still relatively uncommon, and most studies continue to rely on standalone RegCM without ocean coupling. Therefore, I suggest that the authors conduct a parallel analysis using the same model configurations over terrestrial subregions where high-quality observational data are available.*

We cannot disagree with you here, as precipitation and temperature over land have been the focus of the great majority of the regional climate community for the past two decades (e.g. Juneng et al., 2016; Ngo-Duc et al., 2017; Ngo-Duc et al., 2024). In this article, we aim to fill a gap in the literature about the performance of RegCM in SEA over the oceans, precisely considering the exclusive attention of the literature on land performance so far, hence largely disregarding ocean variables. Studying how configurations identified as optimal over the oceans actually perform over land is now legitimate, but this will be addressed in a standalone paper with more comprehensive performance evaluation of the model applied to a smaller subset of configurations (and with a longer simulation period). Nevertheless, at the moment, the parallel analysis you suggest can somehow be found with the last sensitivity experiments of

CORDEX-SEA by Ngo-Duc et al. (2024). Relating the results of the two studies (ours and that of Ngo-Duc et al., 2024) is done in the updated conclusion to give the big picture to interested readers:

L626 "Another approach within the same constraints would have been to fine-tune one single configuration we would have chosen based on previous research. However, previous studies featured significantly less physical options than those tested here, and focused almost exclusively on land performance. For example, Ngo-Duc et al. (2024) recently employed 0***0 configurations to assess land temperature and precipitation (and to our knowledge, only 01*10 experiments were tested in earlier works). As a result, our understanding of how RegCM performs across the full SEA domain was incomplete, and some recent options were never assessed despite yielding good results in the present study (e.g. RRTM and Xu–Randall). After the current paper, assessing RegCM's most updated schemes over land would be a valuable follow-up. Nonetheless, in order to guide modelers seeking homogenous RegCM performance over the region, we can conduct as of now a brief comparison of our ocean-focused results with the land-only ones of Ngo-Duc et al. (2024). They notably identified four configurations with equivalent aggregate scores, including three using Kain–Fritsch and one using Tiedtke. Our results indicate that Kain–Fritsch tends to overestimate oceanic monsoon signals in terms of precipitation, sea surface wind and latent heat flux, such that Kain–Fritsch configurations generally ranked in the bottom third of the ensemble. This supports favoring their top experiment that used Tiedtke instead. The Tiedke configuration highlighted in their study (i.e. 02510 using our notation) ranked third overall in ours while sharing the same PBL (UW-PBL), cumulus convection (Tiedke) and microphysics (SUBEX; i.e. *251*) as in the first and second ranks. This suggests that a balanced configuration may lie among these *251* combinations. Our work thus serves as a prerequisite before embarking on any fine-tuning efforts from a relevant configuration. According to our findings, future fine-tuning efforts should first target the cumulus convection scheme, which was the primary driver of performance. Radiative transfer, PBL, and microphysics should follow as secondary priorities, while the cloud fraction algorithm warrants lower focus."

*3. **<u>Using ERA5</u>** at the same resolution (0.25°) to force RegCM5 is valid and appropriate for a controlled physics sensitivity study, as done by the authors. However, in this setup, the added value of high-resolution spatial detail from the regional model cannot be fully realized.*

We do agree with you on this point. 25 km was chosen in agreement with the experiments made in the CORDEX-SEA community, who eventually conducts regional climate projections thereby limiting the possibilities of high resolution for computing resource concerns. A refinement of resolution should be opted for in the next phase of CORDEX-SEA (or in other future coordinated regional experiments), in particular alongside the upcoming CMIP7. As of now, indeed, we cannot (and do not claim to at this stage) evaluate finer spatial climate patterns in the region, relative to ERA5 (although relative to most CMIP6 models, the resolution is still refined). This will have to wait for future studies following the regional community's plans.

A new paragraph opening up the perspectives of the study now include this consideration (together with ideas of the other anonymous referee):

L646 "The research proposed in this article also invites further exploration. For example, while we chose to force RegCM with a high-resolution SST field from SYMPHONIE in place of traditional, smoother SST datasets, we did not address the impact of this choice on the outputs of the model. How oceanic mesoscale eddies and meander impact the formation of clouds and precipitation in the area? We employed a 25 km horizontal resolution, so this may limit the impact of oceanic mesoscale in the atmosphere, but will this influence of SST become more critical with future resolution improvements? Indeed, with the upcoming seventh phase of the Coupled Model Intercomparison Project (CMIP; Dunne et al., 2024), resolution should increase in both global and regional climate models (including those we employ). More generally, how will our performance ranking evolve with those new resolutions?"

References

Juneng, L., Tangang, F., Chung, J., Ngai, S., Tay, T., Narisma, G., Cruz, F., Phan-Van, T., Ngo-Duc, T., Santisirisomboon, J., Singhruck, P., Gunawan, D., and Aldrian, E.: Sensitivity of Southeast Asia rainfall simulations to cumulus and air-sea flux parameterizations in RegCM4, Clim. Res., 69, 59–77, https://doi.org/10.3354/cr01386, 2016.

Ngo‐Duc, T., Tangang, F. T., Santisirisomboon, J., Cruz, F., Trinh‐Tuan, L., Nguyen‐Xuan, T., Phan‐Van, T., Juneng, L., Narisma, G., Singhruck, P., Gunawan, D., and Aldrian, E.: Performance evaluation of RegCM4 in simulating extreme rainfall and temperature indices over the CORDEX-Southeast Asia region, International Journal of Climatology, 37, 1634–1647, https://doi.org/10.1002/joc.4803, 2017.

Ngo-Duc, T., Nguyen-Duy, T., Desmet, Q., Trinh-Tuan, L., Ramu, L., Cruz, F., Dado, J. M., Chung, J. X., Phan-Van, T., Pham-Thanh, H., Truong-Ba, K., Tangang, F. T., Juneng, L., Santisirisomboon, J., Srisawadwong, R., Permana, D., Linarka, U. A., and Gunawan, D.: Performance ranking of multiple CORDEX-SEA sensitivity experiments: towards an optimum choice of physical schemes for RegCM over Southeast Asia, Clim Dyn, https://doi.org/10.1007/s00382-024-07353-5, 2024.