Dear Editor and Reviewer,

We sincerely thank the Reviewer for the constructive comments. Your recommendations are valuable and helped enhance the analysis and sharpen the argumentation in the manuscript. In the following, we address each comment. The reviewers' comments are presented in black, and our responses are **in blue**.

Overview of the paper

The paper presents a hydrological simulation in a data-scarce region, utilizing a convection-permitting climate model and a lake-integrated hydrological model. It highlights and quantifies the improvements brought by convection-permitting WRF simulations and the inclusion of lake and reservoir modeling. A key contribution of the paper is its identification of specific areas of improvement, particularly at peak and low-flow points, and a detailed explanation of the underlying causes and attributions of these improvements.

The study addresses the added value of convection-permitting modeling in hydrological simulations and the integration of the lake/reservoir module within WRF-Hydro, both of which are topics that have been rarely explored. The findings provide fresh insights into the benefits of using a convection-permitting climate model and the lake/reservoir module, offering valuable benchmarks for optimizing hydrological modeling, especially in regions with limited data availability. As noted by the authors, this enhanced hydrological simulation has potential future applications in forecasting extreme hydrological events, such as floods and droughts.

The paper aligns well with the scope of HESS, as it focuses on hydrological simulation, especially in a data-scarce region. The use of the convection-permitting climate model and lake-integrated hydrological model addresses critical concerns in the climate-hydrology field. Furthermore, the study area, East Africa, is particularly relevant due to the scarcity of data, the complexity of hydrological simulations, and the frequent occurrence of extreme flood and drought events.

In conclusion, this paper offers significant value and is suitable for publication in HESS. However, I have several suggestions and comments for consideration, as outlined below.

Major comments

Although this manuscript is well-structured and comprehensive, with some well-written sections, it requires careful editing by professional English editors. Special attention should be given to sentence structure, as well as minor spelling and grammatical errors, to ensure that the study's goals and results are clear to the reader.

Reply: Thanks for your comment. We have revised the English language throughout the manuscript to enhance clarity and readability.

The conclusions should be presented with caution. The sensitivity of the simulated peak flows to the spin-up time, based on a single event in 2011. The conclusion may vary if different regions or other peak flow events are considered. I recommend adding further discussion on this point.

Reply: Thank you for the suggestion. We added further discussion on this point in Sect. 5.4. The added discussion now reads as follows.

"Also, uncertainty may exist in the sensitivity analysis of the simulated peak flow to spin-up time, which was based on a single event (the largest observed peak from 2010 to 2014) at a specific discharge station (i.e., Garissa or Rukanga). The conclusion, especially about the spin-up time required for model stabilization, may vary when different regions or other peak flow events are considered. For example, a WRF/WRF-Hydro simulation (Lu et al., 2020) exhibits that initialization times needed for soil moisture stabilization differ for different basins in Western Norway. The varying spin-up periods required for flow stabilization between the dry and rainy seasons (Sect. 4.2.2 and Fig. 5 d-e) indicate the possible sensitivity of peak flow to spin-up duration across different peak flow events. The sensitivity of different regions and other peak flow events to spin-up time will be further investigated."

The authors should verify the model configuration. Figure 1 indicates that WRF is directly driven by ERA5, but lines 156-157 suggest a nested simulation domain. Please double-check this for accuracy.

Reply: Thanks for your careful review. We have corrected it in the revised manuscript (Line 163) as below.

"The convection parameterization was turned off for the WRF simulation,…".

When evaluating the WRF simulation, the authors focus primarily on bias. However, a smaller bias does not necessarily indicate a better simulation. I suggest using additional indices, such as a Taylor diagram, to support the conclusions.

Reply: Thanks for your valuable suggestion. We acknowledge that focusing only on bias may not fully capture the simulation's accuracy. We have added additional evaluation indices, including the Taylor diagram, to provide a more comprehensive assessment of the WRF simulation's performance. To ensure the continuity of the context, we have revised Section 4.1 as follows.

"4.1. WRF Precipitation refinement

Using IMERG precipitation as a benchmark, we assessed the performance of convection-permitting WRF precipitation at a 5 km resolution in TRB, through the comparison to ERA5 reanalysis (the forcing driver of our WRF simulation). The evaluation focused on seasonal precipitation averaged over 2010-2014 (Fig. 2) and daily precipitation distribution (Fig. 3). Taylor diagram (Taylor, 2001), with spatial correlation (r, correlation coefficient) and spatial variance (normalized standardized deviation), is also applied for evaluation (Fig. S2).

The WRF model captures the spatial pattern of precipitation and its seasonal variations over TRB, as presented in IMERG (Fig. 2 and Table 7). The spatial distribution of WRF simulation reveals that the precipitation is primarily concentrated in mountainous regions (such as Mount Kenya and Aberdare Range, and the surroundings in Fig. 1 b), with significantly less precipitation in the plain area (Fig. 2 b, g, l, and q). The annual mean precipitation is approximately 1500 mm in the mountainous areas compared to less than 500 mm in the plain area (Table 7). During the rain seasons (MAM and OND), the total precipitation is 976 mm over the terrain area and 327 mm over the plain area, in contrast to 417 mm and 33 mm during the dry season (JF and JJAS). This spatial and seasonal pattern is consistent with that in IMERG data (Figs 2 a, f, k, and p), indicating a distinct orographic and seasonal dominance.

Compared to ERA5, WRF precipitation generally shows better performance, indicated by the Taylor diagram (Fig. S1 e-f) in terms of spatial correlation (correlation coefficient) and spatial variance (normalized standardized deviation) although the advantage is not obvious. The median correlation coefficient of WRF precipitation against IMERG is 0.80, higher than ERA5's value of 0.66 (Fig. S1e). Similarly, the median normalized standardized deviation of WRF precipitation is 1.1, closer to 1, compared to ERA5's value of 1.7 (Fig. S1f). The improved performance of WRF is also evident from the model-data bias comparison. WRF simulation shows a smaller area with larger biases (exceeding 60%) compared to ERA5. During MAM, OND, JF, and JJAS, the areas with larger biases are 618.2 km^2 (1.9%), 711.0 km^2 (2.2%), 680.0 km^2 (2.1%), and 3431.0 km^2 (10.4%), respectively. In contrast, ERA5 shows corresponding areas of 1545.5 km^2 (4.7%), 1545.5 km^2 (4.7%), 10818.3 km^2 (32.9%), and 8500.1 km^2 (25.9%), respectively.

Spatially, the superior performance of WRF precipitation compared to ERA5 is evident in the mountainous regions (mainly Mount Kenya and its surroundings), as demonstrated by the spatial distribution of the model-data bias (Fig. 2 d-e, i-j, n-o, and s-t) and AV (Fig. S2) result. Specifically, the model-data bias from WRF is 210 mm (18 %) per year over the mountainous areas, whereas ERA5 shows a bias of 681 mm (58 %) (Table 7). Additionally, over the mountainous areas, WRF adds value to ERA5 (Fig. S2 a-e), with a positive AV of 0.14 averaged across the four seasons and this area. Such improvement over the mountainous areas is more pronounced in JF season. Model-data bias in JF season is -5 mm (-5%) from WRF and 139 mm (152%) from ERA5. In contrast, during the MAM, OND and JJAS, the bias is 29 mm (7 %), 48 mm (10 %) and 138 mm (72%) from WRF with the value of 161 mm (37 %), 100 mm (22 %) and 281 mm (146%) from ERA5. The improvement over the mountainous areas during JF season is highlighted in the Taylor diagram (Fig. S1 c). The spatial correlation or normalized standardized deviation, calculated over the JF averaged precipitation in the mountainous areas, is 0.56 or 2.18 for WRF, in contrast with -0.14 or 5.46 for ERA5.

Also, the probability distribution of regionally averaged daily precipitation from WRF result exhibits better alignment with the benchmark than from ERA5 (Fig. 3). WRF aligns more closely with IMERG for both small (0-20 mm/day) and extreme (>20 mm/day) rainfall events compared to ERA5, as shown in Fig. 3 and Table 8. The cumulative probability of the small (or extreme) rainfall is 0.991 (0.009) from CPWRF and 0.981 (0.019) from IMERG, whereas 0.995 (0.005) from ERA5. Among these, the better alignment of light rainfall (1-15 mm day-1) probability between WRF and IMERG is pronounced (Fig. 3 and Table 8). The probability of light rainfall from WRF is 0.255, and 0.242 from IMERG, whereas 0.489 from ERA5. Consistently, WRF adds value to ERA5 over the probability of light rainfall (Fig. S2 f-k), with a positive AV of 0.21 averaged across the basin during the four seasons. The better alignment of light rainfall from WRF than ERA5 is particularly evident during the dry season (Fig. 3). The probability of 1-15 mm day-1 events during the dry season from WRF is 0.15 and 0.13 from IMERG, whereas 0.32 from ERA5."

Additionally, we have updated the methodology in Sect. 3.4 to include details on the Taylor Diagram (Line 312-314), which reads as follows.

"To fully evaluate the simulated precipitation by CPWRF, we also employed Taylor diagrams (Taylor, 2001), which present a concise statistical summary in terms of spatial correlation (indicated by correlation coefficient), and spatial variance (indicated by normalized standardized deviation). A higher spatial correlation and a spatial variance closer to one indicate better simulation skills."

We also add discussion about the performance of the model in the Uncertainty, which reads as follows.

"Different metrics (r, bias, and normalized standardized deviation) were used to provide a more comprehensive assessment of the CPWRF's performance, which may cause contradictory or different evaluations of its skill. Each metric emphasizes different aspects of model performance and leads to divergent conclusions about the model's strengths or weaknesses. For instance, seasonal precipitation from the CPWRF result exhibits apparent added value to the forcing data over mountainous areas (Fig. S2 a-e), which however is not distinct in the Taylor figure (Fig. S1). This discrepancy arises because the region with apparent added value is mainly centered on Mount Kenya, whereas the mountainous region in the Taylor diagram analysis includes areas above 1600 m, extending beyond Mount Kenya. Therefore, further in-depth research is needed to fully assess the performance of CPWRF with these different metrics and explain the possible discrepancy."

The manuscript emphasizes the importance of CPM in East Africa. If the authors add a discussion of the added values of CPM with respect to its driving forces, this manuscript would be more informative.

Reply: Thank you for this insightful suggestion. To address this, we have expanded Section 5.2 including the added value of Convection-Permitting Models (CPM) in East Africa, particularly in relation to their driving forces. This additional content emphasizes CPM enhances the representation of convective processes compared to coarser-resolution models, over Mount Kenya and the surrounding area. We believe this provides a more thorough understanding of CPM's benefits in this region. To ensure the continuity of the context, we have revised Section 5.2 as follows.

"5.2. Hydrological modeling improvement from precipitation simulated by convection-permitting WRF – Effect of precipitation forcing

Dynamic downscaling at convection-permitting resolution allows for a more accurate representation of precipitation processes. The convection-permitting WRF simulation enhances local (e.g., mesoscale) processes and interactions between local and large-scales, especially over complex terrain (Kendon et al., 2021; Guevara Luna et al., 2020; Schmidli et al., 2006; Schumacher et al., 2020; Li et al., 2020). As a result, CPWRF potentially contributes to improving precipitation simulation in our study (Sect. 4.1), especially reducing bias in seasonal precipitation over mountainous areas, and light rainfall (1-15mm day-1) probability in the dry season compared to ERA5 (Fig. 3 and Table 8).

The improvement in the seasonal precipitation over mountainous regions and rainfall probability can be supported by the spatial distribution of the added value (AV) in seasonal precipitation with respect to the driving forces (Fig. S2). The WRF simulation adds consistent value to ERA5 over the mountainous areas across all four seasons (MAM, OND, JF, and JJAS). The area with positive AV is mainly over Mount Kenya and its surrounding areas, with the positive AV being particularly distinct during the dry season. WRF also adds value to ERA5 in the light rainfall probability (Fig. S2 f-k), as demonstrated in Sect. 4.1. The basin averaged AV of CPWRF over the probability of light precipitation events are 0.32, 0.26, 0.30, and 0.07 in MAM, OND, JF, and JJAS, respectively. The positive AV of WRF with respect to ERA5 over the extreme rainfall probability, also concentrates around Mount Kenya, consistently across all four seasons (Fig. S2 l-p). Previous studies (Giorgi et al., 2022) have demonstrated that the added value of WRF simulations is influenced by various factors, including timescale, variables, regions, and uncertainty of the benchmark. Therefore, further in-depth research is required for a more reliable AV assessment.

Due to the precipitation improvement from WRF, hydrological simulation with WRF precipitation as a driving force (LakeCal), showed significant improvements, compared to simulations driven by ERA5 (LakeCal-ERA5) (Fig. 10a). These improvements are particularly notable in reducing false peak simulations, likely due to the reduction in the overestimation of light rainfall probability. The enhancement in peak flow simulation is also observed in the WRF-Hydro model without the lake/reservoir module (Fig. 10b).”

Besides, we have also added a description of the AV (added value) methodology in Sect. 3.4, which reads as follows.

“3.4. Evaluation of simulated precipitation from CPWRF

To assess whether the CPWAR has advantages over their driving forces (ERA5), added value (AV) proposed by Dosio et al. (2015) was applied, expressed as follows.

$$AV = \frac{(X_{ERA5}-X_{IMERG})^2-(X_{CPWRF}-X_{IMERG})^2}{\max\left((X_{ERA5}-X_{IMERG})^2,(X_{CPWRF}-X_{IMERG})^2\right)} \tag{3}$$

X_{ERA5}, X_{CPWRF}, *and* X_{IMERG} indicate precipitation from the driving forces (ERA5), CPWRF simulation, and benchmark (IMERG), respectively. The added value (AV) from CPWRF is defined as the performance difference between itself and the driving forces for precipitation in a specific region and period. If the CPWRF adds value to the driving forces (ERA5), the AV is positive, whereas a negative AV suggests no adding value.

To fully evaluate the simulated precipitation by CPWRF, we also employed Taylor diagrams (Taylor, 2001), which present a concise statistical summary in terms of spatial correlation (indicated by correlation coefficient), and spatial variance (indicated by normalized standardized deviation). A higher spatial correlation and a spatial variance closer to one indicate better simulation skills.”

Thank you once again for your suggestion. We believe these additions enhance the manuscript by providing a more comprehensive view of CPM's contributions in East Africa.

Minor comments

Line17-18: Replace “the upper and middle stream of the Tana River basin was” with “the upper and middle streams of the Tana River basin were ”.

Reply: Thanks and done.

Line19-20: This sentence is ambiguous. Please revise it.

Reply: Thank you for pointing it out. We have revised the sentence and changed it to “We performed convection-permitting (CP) simulations using the Weather Research and Forecasting (WRF) model and conducted lake/reservoir-integrated WRF Hydrological modeling (WRF-Hydro) driven by CPWRF output.”.

Line21: Replace “ using IMERG as the benchmark” with “ when benchmarked against IMERG

Reply: Thanks and done.

Line22-23 Change “alleviates the peak false” to “alleviates the false peak simulation”.

Reply: Thanks and done.

Line24: Change “NSE” with “NSE (Nash-Sutcliffe Efficiency)”.

Reply: Thanks and done.

Lin29-30: There are two terms “lake” and “lake/reservoir” which seem to represent the same thing. It would be better to unify them for consistency throughout the document.

Reply: Thanks and done.

Line29-30: Replace “highlight the enhanced hydrological modelling capability with” with “highlight the enhanced capability of hydrological modelling using”.

Reply: Thanks and done.

Line123: Replace “resulting” with “which results”.

Reply: Thanks and done.

Line129: Please replace “S 1.25°~N 0.50°” with “S 1.25°-N 0.50°”.

Reply: Thanks and done.

Line139: Replace “ the upper and middle stream of the Tana River Basi” with “ the upper and middle streams of the Tana River basin”.

Reply: Thanks and done.

Line:154: Usually, one month spin-up is sufficient for WRF downscaling.

Reply: Thanks for pointing it out. We have revised Sect. 3.3 as the follows.

“To obtain convection-permitting modeling precipitation, we used the Advanced Research WRF (WRF-ARW) model of version 4.4 (Skamarock et al., 2019) with the designed domain of 5 km spatial resolution (Fig. 1). The lateral boundaries were forced with the 6-hourly ERA5 reanalysis with a spatial resolution of 0.25 degrees (Hersbach et al., 2020). The model was set with 50 vertical levels up to 10hPa. The convection parameterization was turned off for the WRF simulation, the Mellor-Yamada Nakanishi Niino Level 2.5 (MYNN2.5) Scheme

(Nakanishi and Niino, 2006) for the planetary boundary layer, the RRTM scheme for longwave radiation (Mlawer et al., 1997), and the Dudhia Shortwave scheme for shortwave radiation (Jimy Dudhia, 1989). The Noah-MP Land Surface model ('Noah-MP LSM', Yang et al., 2011) was used for the land surface scheme.

The model runs from 1 January 2010 to 31 December 2014. Typically, WRF simulations require a spin-up of about one month, which should ideally be excluded from precipitation evaluation. However, given the limited length of simulated precipitation, the subsequent analysis is based on full precipitation simulation from January 2010 to December 2014."

Line182-183: Please add "(with the lake/reservoir module inactive)" behind "without the lake/reservoir module".

Reply: Thank you for the suggestion. We have revised it as requested.

Line185: Replace "of the Garissa discharge." with "of simulated discharge against the observation at Garissa".

Reply: Thank you for the suggestion. We have revised it.

Line208-210: Replace "which may affect the subsequent sensitivity analysis and hydrological modelling assessments." with "which may potentially affect the result of subsequent sensitivity analyses and the performance of the hydrological simulation."

Reply: Thank you. We have revised it.

Line260: Replace "For each lake test" with "For each test".

Reply: Thanks and done.

Line274: Replace "Each lake" with "Each".

Reply: Thanks and done.

Line306-307: Please change the unit "mm a^{-1}" to "mm". The unit should be corrected in the whole text.

Reply: Thanks and done.

Line368: Replace "GIS pre-41.processing" with "GIS pre-processing".

Reply: Thanks and done.

Line408: Replace "demonstrate" with "demonstrates".

Reply: Thanks and done.

Line411: The view "a larger lake seems to require more time to reach equilibrium." depends.

You should add some references.

Reply: Thanks for the helpful suggestions. We did not find a related reference. So, we have removed this sentence and revised the related paragraph as follows.

“The water levels from the lake/reservoir-integrated model show a consistent spin-up period of 4 years across nearly all five lakes for the entire period, as well as during both the rainy and dry seasons (Fig. S2). Although KIAMBERE (one of the five lakes) exhibits a spin-up period of 3 years during the rainy season (Fig. S2e), it can be considered nearly 4 years due to the uncertainty in determining the spin-up time required for the stabilization of specific variables. Since the lakes are interconnected, the stabilization time is governed by the longest spin-up period. This may result in nearly the same initialization time for all five lakes (Table 1).”

Line499: Replace “-53” with “-53%”.

Reply: Thanks for pointing out our careless, and done.

Line527: Replace “under estimation” with “underestimation”.

Reply: Thanks and done.

Line529: “WRF-refined precipitation” is not idiomatic. Please revise it.

Reply: Thanks and done.

Line544: Replace “Woodhams et al.'s research (2018) demonstrates” with “Woodhams et al. (2018) demonstrates”.

Reply: Thanks and done.

Line554: Delete “. Factors”.

Reply: Thanks and done.

Line565: Replace “Arnault et al.,” with “Arnault et al.”

Reply: Thanks and done.

Line574-575: Replace “ lake water levels may be not well presented” with “it shows a limited skill for simulating water level”.

Reply: Thanks and done.

Line577: Please replace “ with r2 ranging from near zero (0.005) to 0.25 for the five lakes.” with “ with r2 of the simulated discharge against the observation at Garissa less than 0.25 for all the five lakes.”

Reply: Thanks and done.

Line599: Replace “a seamless, consistent meteorological-hydrological modelling system” with “a seamless and consistent meteorological-hydrological modelling system”.

Reply: Thanks and done.

Line601-602: The sentence “which makes an NSE increase of 0.53 when comparing LakeCal to LakeCal-ERA5” is not clear. Please correct it.

Reply: Thanks for pointing it out. We have revised the text as “which result in an increase of 0.53 in NSE between the well-calibrated lake-integrated WRF-Hydro simulation driven by CRWRF output (LakeCal) and that driven by ERA5 precipitation (LakeCal-ERA5)”.

We thank the reviewer for the valuable comments, which have helped improve our manuscript. We hope the revisions meet your expectations and strengthen our study.