Reviewer #2

This study performed a regional-scale assessment of climate change impacts on flood for Western Africa, using two large-scale hydrological models with the bias-corrected CMIP6 climate projections. I think the study has high potential to form valuable knowledge base of the likely changes in flood in Western Africa, but I have some major concerns on the approach taken in hydrologically modelling, which limited my capacity to assess the results presented. As such, I'd like to seek further clarification and justification from the authors on their chosen approach, or reconsideration of alternative approach, before proceeding to further review of the results.

We thank the reviewer for his encouraging feedback on the potential contribution of our study. We acknowledge the concern regarding the hydrological modelling approach. We have provided additional clarification and justification for our methodology in the revised manuscript.

General comments:

1. Given the substantial lack of data in the study region, I'm wondering about the value of using rather complicated hydrological models (distributed and semi-physical) rather than simpler models (e.g., lumped conceptual models)? There is a lack of 1) motivation for exploring distributed and semi-physical modeling approach within the study objective (in the Introduction); 2) justification of modelling approach within Section 2.3 of the Materials and Methods. The start of the Introduction also touched on the challenge with data scarcity for the study region, which seems to suggest that uncertainties from input data might affect modelling (especially for more complex models which has higher data requirement) to some large degree - some assessments and/or discussion on this aspect would be useful.

We agree that model complexity must be balanced against data availability, particularly in data-scarce regions such as West Africa. Nevertheless, our primary motivation for utilizing distributed models is to account for the spatial heterogeneity of runoff-generating processes such as variations in land use, soil properties, and rainfall patterns, which cannot be adequately captured by simple lumped models. It should also be noted that in these regions, many studies are based on simple models that rely exclusively on calibration, without explicitly taking basin properties into account. Our study proposes an important step forward by using process-based models, and in the future these models could also provide a better understanding of the complex interactions between climate and land-use changes. We have added in the introduction, pages 4-5, lines 127-138: "... Due to their simplicity and computational efficiency, lumped hydrological models have been widely applied in West Africa (Niel et al., 2003; Bodian et al., 2016; 2018; Kwakye & Bárdossy, 2020; Koubodana et al., 2021). However, because runoff generation is an inherently spatial and temporally dynamic process, changing environmental conditions may impact flood frequencies and water availability (Wilson et al., 1979; Haddeland et al., 2002; Descroix et al., 2018). Although lumped models often perform comparably or even better than distributed models at the catchment outlet (Reed et al., 2004), their main limitation lies in evaluating the overall catchment response simply at the outlet, without accounting for the contributions of upstream individual sub-basins (Cunderlik, 2003; Pokhrel et al., 2008; Jajarmizad et al., 2012). The main advantage of distributed models is not necessarily a higher accuracy of runoff simulations at specific points (e.g., outlet or gauge stations), but rather their broader applicability and ability to simulate the impacts of spatially varying drivers and scenarios (Gebremeskel et al., 2005; Tang et al., 2007; Thielen et al., 2009; Chu et al., 2010; Tran et al., 2018). ..."

.

2. In the current analyses, the HMF-WA model has not been calibrated, while calibration for LISFLOOD seems to be done previously which are not part of this study. This attracts several major questions on the modelling approach:

- The disadvantage of not calibrating HMF-WA is clearly demonstrated in the results (Figure 2), the largely unsatisfactory performance of the model suggests that we have low confidence that it could even well represent the historical flood events. Although the results is accompanied by brief discussion on this issue that 'projections of climate change impacts on African hydrological trends were produced using…' – the decision to use an uncalibrated model is generally not standard in the international literature and require much more justification.

We partly agree with the reviewer that using an uncalibrated model is uncommon in the international literature to provide hydrological scenarios of climate change impacts. One of the most striking examples is the use of ISIMIP (Frieler et al., 2017) simulations (large-scale global models that are mostly uncalibrated) for numerous hydrological impact studies, including in the journal Science (Gudmundsson et al. 2021). Furthermore, most land-surface models used to provide hydrological scenarios are also not calibrated.

One of the core objectives of this study is to evaluate the consistency of climate change signals across hydrological models that differ in structure and complexity, and notably to identify whether the use of a calibrated model could provide different hydrological projections under climate scenarios. By including both a calibrated model (LISFLOOD) and an uncalibrated model (HMF-WA), we aim to investigate how model calibration influences the projection of flood trends under changing climatic conditions, and to evaluate the potential of uncalibrated hydrological models as a practical alternative for addressing data scarcity in poorly gauged regions for climate impact studies. We have added a justification in the description of selected hydrological models (Section 2.3, lines 312-319): "... Nevertheless, while calibration can enhance the accuracy of discharge simulations, several studies have highlighted that uncalibrated global hydrological models often exhibit comparable sensitivity to climate variability as the regional calibrated hydrological models, particularly when assessing relative changes in extreme events between future and historical periods (Gosling et al., 2017; Zhao et al., 2025). Therefore, whether a calibrated hydrological model offers different climate change projections than an uncalibrated model needs further investigation (Pechlivanidis et al., 2017). …". Moreover, our findings (line 803 in the revised manuscript) suggest that: "... using both models, the climate forcing has more importance than the hydrological representation itself."

- On the LISFLOOD model, further justifications and details are required on the calibration process, including the input data, objective function, and cross-validation (if any). Such information on calibration is necessary for the reviewers/readers to assess the suitability of these models for the purpose of the study.

We thank the reviewer for raising this important point. We have added a detailed description of the calibration process of the LISFLOOD model in the revised manuscript, at page 10, lines 291-319): "... The LISFLOOD version used in this study (OS LISFLOOD v4.1.3) was regionally calibrated with a 0.05° (~5 km) resolution, using in-situ discharge gauge stations with at least four years of daily measurements recorded after 1 January 1982. In this setup, model parameters are linked to global geospatial datasets describing catchment morphology and river networks, land use, vegetation characteristics, soil properties, lake distribution, and water demand (Salamon et al., 2024; Choulga et al., 2024). The Distributed Evolutionary Algorithms in Python (DEAP; Fortin et al., 2012) framework

.

was applied to optimize parameters in gauged catchments, with the modified Kling-Gupta Efficiency (KGE; (Gupta et al., 2009) utilized as the objective function. Calibration was performed over a continuous simulation period using ERA5 reanalysis meteorological forcing. Due to the varying length and temporal coverage of the discharge records used for calibration, model performance was assessed using all available observational data at each station, rather than splitting the records into separate calibration and validation periods. The LISFLOOD calibration tool is freely available at https://github.com/ec-jrc/lisflood-calibration."

- Further, it is not clear whether the LISFLOOD models have been explicitly calibrated/evaluated to a flood context. Please see an example of calibration of hydrological models tailored to rarer floods, would the models used in this study benefit from a flood-centered calibration? Wasko et al., 2023. https://doi.org/10.1016/j.jhydrol.2023.129403

We thank the reviewer for the relevant reference. Unlike in Wasko et al. (2023), where the GR4J lumped rainfall-runoff model was locally calibrated to rare floods using a flood-centered objective function, the LISFLOOD model calibration relies on the modified Kling-Gupta Efficiency (KGE; Gupta et al., 2009), which is not explicitly designed to prioritize rare or extreme events. Nevertheless, LISFLOOD has demonstrated robust performance in simulating daily river discharge across a large number of calibration sites worldwide, with a global median KGE of 0.70 (https://confluence.ecmwf.int/display/CEMS/GloFAS+v4+calibration+hydrological+model+performance). Importantly, LISFLOOD is the core hydrological model used in both the Global Flood Awareness System (GLOFAS) which provides an overview on upcoming floods in large world river basins (Alfieri et al., 2013; Hirpa et al., 2018; Harrigan et al., 2020; Prudhomme et al., 2024; Silva Peixoto et al., 2024) and the European Flood Awareness System (EFAS; Thielen et al., 2009; Matthews et al., 2024) which operates on a pan-European scale to provide short-to medium-range flood forecasts (Smith et al., 2016; Zábori et al., 2024), under the Copernicus Emergency Management Service (CEMS). Its proven applicability to large-scale hydrological forecasting and flood monitoring confirms its suitability for simulating floods, especially in large river basins. This is a key reason for its use in the present study. In addition, LISFLOOD GloFAS and EFAS set-ups are also used by the CEMS Global and European Drought Observatories (GDO, EDO, respectively) for low flow index and soil moisture anomaly estimation (e.g. Toreti et al., 2025). LISFLOOD GloFAS set-up also proved to enable adequate assessment of total water storage (e.g. Jensen et al., 2025) and is used for various purposes. Therefore, a flood-centered objective function was not used in the calibration process.

Moreover, we have assessed the LISFLOOD model's ability to simulate flood behavior, by comparing the distributions of observed and simulated annual maximum flows using the Anderson-Darling (AD) test at the 0.05 significance level (Scholz & Stephens, 1986). The results indicate that LISFLOOD reproduces the statistical properties of extreme flows reasonably well at a majority of the gauged stations, providing further confidence in its application for flood frequency analysis under historical and projected climate conditions. We have added this point into the discussion of the hydrological model evaluation results in the revised manuscript (lines 635-639): "... In addition, the satisfactory performance of the LISFLOOD model indicates that, although a flood-centered calibration approach could potentially improve its ability to capture extreme flows and their trends (Wasko et al., 2021), the current model setup provides a satisfactory basis for regional-scale flood trend assessments. …"

3. Section 2.2 on data: given the substantial challenges in data availability for the study region, I think specific attention should be paid to the representativeness of the data to ensure they are not

.

biased towards a specific type of catchment, and/or specific time periods. I think this can be achieved by adding the following details:

- A summary table (possibly in the Supplementary) of the selected study sites, with information on their catchment areas, mean annual catchment-averaged rainfall, mean annual streamflow, and the range of years over which streamflow data is available.

We thank the reviewer for this helpful suggestion. We have compiled a summary table of the selected gauge stations, providing key characteristics including catchment area, mean annual catchment-averaged rainfall, mean annual streamflow, and the range of years for which streamflow data are available. This table has been added to the Supplementary Material (Table S1).

- The study site selection criteria mentioned 'a minimum of 10 years streamflow time series between 1950 and 2018' – does this allow for data gaps (i.e., days with missing or low-quality streamflow data), and if so, what is the maximum length of gaps allowed?

As stated, our selection criterion required a minimum of 10 years of continuous available streamflow data between 1950 and 2018. This criterion allows for data gaps provided that there are no missing values near the potential annual peak flood). We have clarified this aspect in the revised manuscript (section 2.2, lines 213-218): "... To address the challenges associated with missing data in the database, we conducted a year-by-year visual inspection of hydrographs at each station as illustrated by Supplementary Figure S2. Years with data gaps near the flood peak were excluded from the analysis to avoid the risk of missing the true annual peak flood (Wilcox et al., 2018). Through this screening process, we ensured that no AMF values were derived from periods characterized by a lot of missing data."

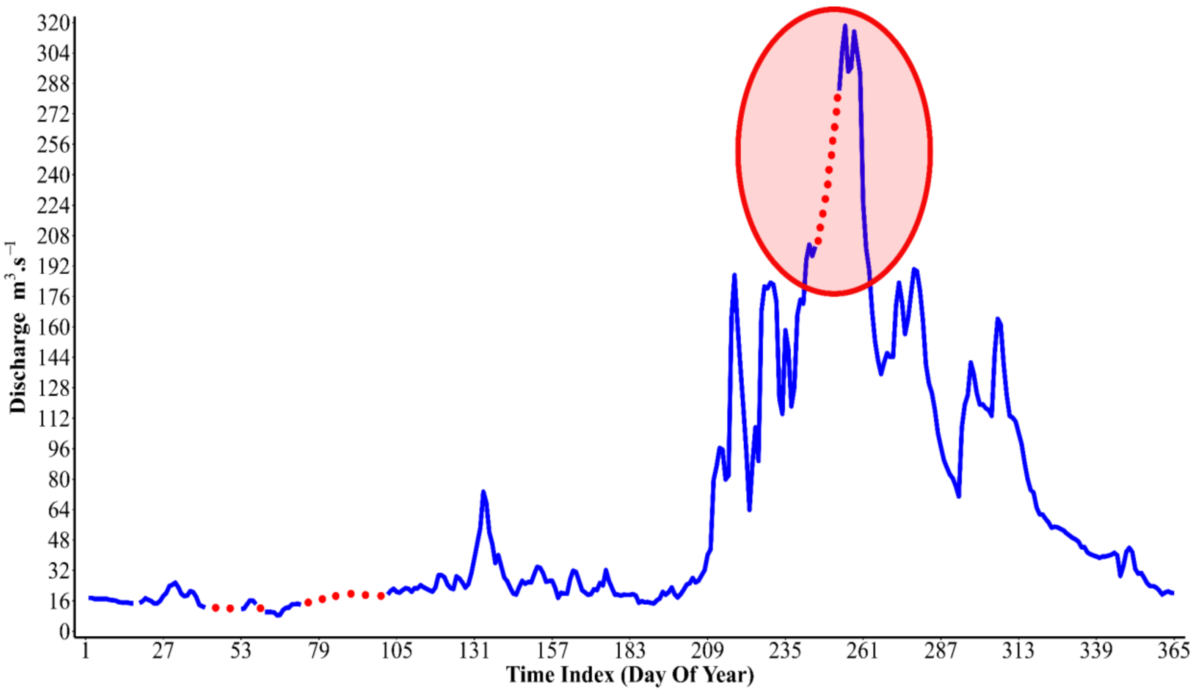

Figure S2: Illustration showing the handling of missing data in an annual hydrograph of daily discharge measurements. A significant portion of data, particularly around the peak discharge period, is missing (highlighted by the red circle). Such a year is excluded from the analysis to ensure the accuracy of the annual peak flood sampling.

.

- How variable is the land uses in this study region? If they are rather heterogeneous, a summary of key land use types in each catchment would also be useful.

The study area is characterized by a heterogeneous landscape across the different catchments, with considerable variability in land use. We have added to the Supplementary Material (Table S2) a summary of key land cover types (forest, urban, crop, irrigated crops, grass, shrub, sparse, and bare) in each catchment, detailing the dominant land uses and their respective proportions.

Table S2: Land use distribution across different catchments in the study area. The table shows the proportion of each land use type (Forest, Urban, Crop, Irrigated Crops, Grass, Shrub, Sparse, and Bare) within the catchments identified by their unique IDs. Each value represents the proportion (percentage) of the respective land use type within a given catchment.

ID	Forest	Urban	Crop	Crop Irrig	Grass	Shrub	Sparse	Bare
ADHI_114	0.06	0	0.46	0	0	0.48	0	0
ADHI_121	0.27	0	0.29	0	0	0.44	0	0
ADHI_123	0.83	0	0.11	0	0	0.06	0	0
ADHI_131	0.74	0	0.17	0	0	0.09	0	0
ADHI_144	0.49	0	0.49	0	0	0.02	0	0
ADHI_163	0.96	0	0.01	0	0	0.03	0	0
ADHI_172	0.05	0	0.78	0.01	0.02	0.14	0	0
ADHI_179	0.27	0.01	0.72	0	0	0	0	0
ADHI_180	0.31	0	0.67	0	0	0	0	0
ADHI_183	0.56	0	0.44	0	0	0	0	0
ADHI_187	0.33	0	0.39	0	0	0.27	0	0
ADHI_198	0.52	0	0.25	0	0	0.23	0	0
ADHI_270	0.76	0	0.17	0	0	0.07	0	0
ADHI_276	0.74	0	0.14	0	0	0.12	0	0
ADHI_304	0.02	0	0.04	0.03	0.25	0.37	0.26	0.07
ADHI_315	0.05	0	0.74	0	0	0.2	0	0
ADHI_316	0.09	0	0.41	0.01	0.11	0.08	0.05	0.25
ADHI_319	0.14	0	0.38	0.01	0.09	0.11	0.03	0.25
ADHI_320	0	0	0.62	0.02	0.21	0	0.12	0.03
ADHI_321	0.33	0	0.51	0	0.01	0.13	0	0
ADHI_324	0.37	0	0.48	0.01	0	0.14	0	0
ADHI_325	0.83	0	0.09	0	0.04	0.03	0	0
ADHI_332	0.39	0	0.02	0	0.58	0.01	0	0
ADHI_372	0.02	0	0.8	0	0	0.17	0	0
ADHI_390	0.7	0	0.23	0	0	0.07	0	0
ADHI_394	0.62	0	0.33	0	0	0.04	0	0
ADHI_507	0.58	0	0.18	0	0	0.24	0	0
ADHI_510	0.8	0	0.08	0	0	0.11	0	0
ADHI_511	0.43	0	0.44	0	0	0.13	0	0
ADHI_515	0.32	0	0.54	0	0	0.14	0	0
ADHI_519	0.05	0	0.74	0	0.09	0.11	0	0
ADHI_531	0.5	0	0.44	0	0	0.05	0	0
ADHI_548	0.78	0	0.01	0	0	0.2	0	0
ADHI_550	0	0	0.21	0.01	0.69	0.01	0.05	0.03
ADHI_560	0.08	0	0.38	0.02	0.23	0.2	0.04	0.06
ADHI_571	0.11	0	0.49	0	0	0.39	0	0
ADHI_585	0.63	0	0.29	0	0	0.07	0	0
ADHI_587	0.56	0	0.23	0	0	0.21	0	0
ADHI_592	0.1	0.01	0.87	0	0	0.02	0	0

.

ADHI_595	0.19	0	0.58	0.01	0.01	0.21	0	0
ADHI_596	0.75	0	0.15	0	0	0.09	0	0
ADHI_597	0.96	0	0.01	0	0	0.03	0	0
ADHI_605	0.02	0	0.93	0.01	0	0.05	0	0
ADHI_607	0.43	0	0.09	0	0	0.48	0	0
ADHI_612	0.39	0	0.09	0	0	0.52	0	0
ADHI_613	0.71	0	0.07	0	0	0.21	0	0
ADHI_617	0.61	0	0.21	0	0	0.18	0	0
ADHI_639	0.41	0	0.55	0	0	0.04	0	0
ADHI_640	0.3	0	0.57	0.01	0.01	0.11	0	0
ADHI_649	0.91	0	0.07	0	0	0.02	0	0
ADHI_650	0.92	0	0.01	0	0	0.07	0	0
ADHI_651	0.72	0	0.16	0	0	0.11	0	0
ADHI_678	0.76	0	0.1	0	0	0.14	0	0
ADHI_692	0.1	0	0.45	0.03	0.19	0.12	0.04	0.08
ADHI_1183	0.19	0	0.4	0	0	0.4	0	0
ADHI_1269	0.71	0.01	0.28	0	0	0.01	0	0
ADHI_1400	0.07	0	0.8	0	0	0.13	0	0
ADHI_1401	0.1	0	0.75	0	0	0.15	0	0

- Sources of input data for the hydrological models e.g., rainfall, temperature – it is unclear where they are from, it is implied from the later Section 2.5 that rainfall and temperature were from GCM rather than observed, but it would be helpful to clarify this earlier in the data section.

We appreciate the helpful suggestion from the reviewer. In response, we have revised Section 2.2 (lines 201–226) to clarify the sources of input data for the hydrological models. Now Section 2.2 reads: "**2.2 Observational data and climate forcings for hydrological experiments:** Daily streamflow data for the period 1950-2018 were obtained from the African Database of Hydrometric Indices (ADHI) recently developed by Tramblay et al. (2021). This database provides hydrometric indices computed from different data sources, with daily discharge time series that span at least 10 years. In the ADHI database, the size of the 441 West African catchments ranges from 95 to 2,150,000 km^2, and some stations have daily discharge data spanning over 44 years. Figure 1 shows the spatial distribution of the ADHI stations used in this study. We only selected watersheds from the ADHI database that met the following three criteria: (i) low regulation, determined through visual inspection of dam locations relative to watershed outlets (see Supplementary Figure S1), combined with a year-by-year analysis of annual hydrographs to assess the impact of dam operations on streamflow, (ii) surface area of less than 150,000 km^2, and (iii) a daily streamflow time series covering a minimum of 10 years between the 1950 and 2018. To address the challenges associated with missing data in the database, we conducted a visual inspection of hydrographs at each station as illustrated by Supplementary Figure S2. Years with data gaps near the flood peak were excluded from the analysis to avoid the risk of missing the true annual peak flood (Wilcox et al., 2018). Through this careful screening process, we ensured that no AMF values were derived from periods characterized by a lot of missing data. It is important to note that the observational streamflow data are not used to calibrate or drive the hydrological models. Instead, these observations serve as an independent benchmark to evaluate the ability of the hydrological models to reproduce key flood statistics during the historical period. The LISFLOOD model was calibrated using the ERA5 reanalysis dataset, which provides consistent and high-resolution precipitation and temperature fields. Moreover, ERA5 was also used as a reference for the bias correction of the five climate models from the CMIP6 ensemble that were used to drive the hydrological simulations for both the historical and future periods (see Section 2.4)."

.

4. Section 2.3: I understand that the details of the two hydrological models are presented in the corresponding papers cited, but I think the readers could benefit from some additional background on these models, at least covering the key processes represented in each model on converting rainfall to runoff. This information is currently only partly available for the HWF-WA model (with only the recently added process representations listed) and not communicated for the LISFLOOD model. After presenting these, I'd also love to see a quick summary of the key differences between the models to justify your point in the Abstract that the two models 'differ in their hydrological process representation'.

We appreciate the reviewer's suggestion. We have expanded Section 2.3 (lines 261-319) to clarify the key hydrological processes in each model and the key differences between the two models. The updated section now reads: "The HMF-WA model is adapted from the modular HMF model, and is designed for large-scale applications across West Africa (Rameshwaran et al., 2021). It employs a vertically integrated soil moisture scheme to simulate runoff production, driven by rainfall and potential evaporation inputs. Runoff generation considers soil drainage and a spatial probability distribution of soil moisture. Routing is based on a kinematic wave approach (Bell et al., 2007), with parallel pathways for surface and subsurface flow. Key enhancements over the classical HMF model include modules to simulate wetland inundation, endorheic basins, and anthropogenic water withdrawals, making it well-suited for semi-arid environments with complex hydrology (Rameshwaran et al., 2021). HMF-WA simulates spatially consistent river flows across West Africa at a 0.1° × 0.1° spatial resolution. Although it has not yet been specifically calibrated to individual West African catchments using observed streamflow data where the model hydrology is configured to local conditions using spatial datasets of physical and soil properties, HMF-WA model evaluation against observational data indicates that it performs reasonably well in simulating both daily high and low river flows across most catchments. The median values of NSE (Nash-Sutcliffe efficiency), NSElog, and BIAS are 0.62, 0.82, and 0.06 (6 %), respectively (Rameshwaran et al., 2021).

The LISFLOOD model, developed by the Joint Research Centre (JRC) of the European Commission (https://ec-jrc.github.io/lisflood/), is a physical, spatially distributed hydrological model, designed for simulating several hydrological processes that occur in a catchment (Van Der Knijff et al., 2010). The LISFLOOD model simulates water processes using a three-layer soil water balance, along with groundwater and subsurface flow models. It accounts for several processes such as snow accumulation/melt, infiltration, evapotranspiration, groundwater flow, surface runoff, etc. Moreover, it supports the integration of human influences such as reservoirs and water abstraction. The numerical LISFLOOD simulation is driven by meteorological forcing (precipitation, temperature, and evapotranspiration) combined with high-resolution spatial data on terrain morphology, soil characteristics, land use, and water demand. This integrated setup allows the model to simulate runoff processes under diverse climatic and socio-economic conditions, capturing both natural and anthropogenic influences across heterogeneous landscapes. The runoff produced at every grid cell within the model domain is routed through the river network using a kinematic wave approach. The LISFLOOD version used in this study (OS LISFLOOD v4.1.3) was regionally calibrated with a 0.05° (~5 km) resolution, using in-situ discharge gauge stations with at least four years of daily measurements recorded after 1 January 1982. In this setup, model parameters are linked to global geospatial datasets describing catchment morphology and river networks, land use, vegetation characteristics, soil properties, lake distribution, and water demand (Salamon et al., 2024; Choulga et al., 2024). The Distributed Evolutionary Algorithms in Python (DEAP; Fortin et al., 2012) framework was applied to optimize parameters in gauged catchments, with the modified Kling-Gupta Efficiency (KGE; Gupta et al., 2009) utilized as the objective function. Calibration was performed over a continuous simulation

.

period using ERA5 reanalysis meteorological forcing. Due to the varying length and temporal coverage of the discharge records used for calibration, model performance was assessed using all available observational data at each station, rather than splitting the records into separate calibration and validation periods. The LISFLOOD calibration tool is freely available at https://github.com/ec-jrc/lisflood-calibration.

Globally, while both models use a kinematic wave routing scheme, HMF-WA and LISFLOOD differ significantly in their hydrological process representation. HMF-WA applies a vertically integrated soil moisture scheme with simplified runoff generation based on spatial soil moisture distribution. In contrast, LISFLOOD features a more detailed, physically-based three-layer soil model with an explicit representation of groundwater, snow processes, and anthropogenic influences. Furthermore, LISFLOOD has been calibrated using in-situ discharge data. Nevertheless, while calibration can enhance the accuracy of discharge simulations, several studies have highlighted that uncalibrated global hydrological models often exhibit comparable sensitivity to climate variability as the regional calibrated hydrological models, particularly when assessing relative changes in extreme events between future and historical periods (Gosling et al., 2017; Zhao et al., 2025). Therefore, whether a calibrated hydrological model offers different climate change projections than an uncalibrated model needs further investigation (Pechlivanidis et al., 2017)."

Specific comments:

1. Line 51 – it will be clearer if the change in flood magnitude can be summarized specific to the flood return period(s) investigated.

We appreciate the reviewer's suggestion. We now specify the change in flood magnitude by return period and future horizon in the abstract (lines 52-55): "... Flood magnitudes are projected to increase at 94% (96%) of stations for the 2-year (20-year) event in the near-term future, and at 88% (93%) of stations for the 2-year (20-year) event in the long-term future, with some locations expected to experience increases exceeding 45%. …"

2. Figure 1 caption: 'grey lines' instead of 'white lines'?

We have changed the word "white" by "grey" in the wording caption of Figure 1 (page 8): "Figure 1: Spatial distribution of the ADHI stations used in this study, covering the three climatic zones in the West African region, as delimited by the blue isohyets (600 mm and 1200 mm annual rainfall) on the map. The color ramp of the circles indicates the record lengths of flood data (in years). The blue lines represent isohyets delimiting West African climatic regions, and the grey lines indicate the borders of West African countries."

3. Line 177 – decision on 'low regulation' catchments: the Supplementary Fig. 1 suggested that this is based on whether there is a dam located near the watershed outlet, with no information what defines a dam 'near' or 'far from' the outlet – was this based on visual inspection, or a threshold distance used? If the latter, how was the threshold distance determined?

We appreciate the reviewer's constructive comment regarding the decision on 'low regulation' catchments. The identification of "low regulation" catchments was based on visual inspection of the dam locations relative to the watershed outlet, using both the GRanD database (https://www.globaldamwatch.org/grand) and Google maps, combined with a year-by-year analysis of the annual hydrographs. This allowed us to verify whether the dam's construction or operational start

.

date caused noticeable changes in the streamflow regime. No fixed distance threshold was applied. We have clarified this aspect in the description of the observational data, Section 2.2, lines 208-213: "... We only selected watersheds from the ADHI database that met the following three criteria: (i) low regulation, determined through visual inspection of dam locations relative to watershed outlets (see Supplementary Figure S1), combined with a year-by-year analysis of annual hydrographs to assess the impact of dam operations on streamflow, (ii) surface area of less than 150,000 km², and (iii) a daily streamflow time series covering a minimum of 10 years between the 1950 and 2018. …"

4. Section 2.6.1 – the introductory section for GEV is very informative, however, I think it could benefit from additional information on what positive and negative shape parameters mean, which seem to be useful context to the subsequent discussion on the plausible values of the shape parameter.

We thank the reviewer for the insightful suggestion. We have now added information about the meaning of shape parameter values, in Section 2.6.1, lines 418-426: "... The shape parameter (ξ) governs the tail behaviour of the GEV distribution, which encompasses three types of extreme value distributions (Coles, 2001): (i) a positive shape parameter ($\xi > 0$) indicates a heavy-tailed Fréchet case (Fréchet, 1927), suggesting an increased probability of extreme flooding events, (ii) a null shape parameter ($\xi = 0$) suggests a light-tailed Gumbel class (Gumbel, 1958), and (iii) a negative shape parameter ($\xi < 0$) indicates a short-tailed or (bounded) negative-Weibull distribution (Weibull, 1951). This parameter is crucial for assessing the risk of rare floods and informing the design infrastructure to withstand such extremes. ..."

5. Line 293: '…estimate the GEV parameters in a non-stationary context' – can you elaborate a bit on what exactly this refers to – is it about fitting multiple GEVs to different periods of the data to represent non-stationary conditions?

We appreciate the reviewer's suggestion regarding the clarification of the wording "estimate the GEV parameters in a non-stationary context." This refers to allowing the GEV distribution parameters to vary with time, in order to capture temporal changes in the statistical behavior of annual peak flood time series. We have added this clarification in the revised manuscript, at page 15, lines 436-439: "... We have used the Generalized (Penalized) Maximum Likelihood Estimation (GMLE) method (Martins & Stedinger, 2000) to estimate the GEV parameters in a non-stationary context, by allowing the model parameters to vary with time (Coles, 2001). …"

.

References

1. Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., and Pappenberger, F.: GloFAS – global ensemble streamflow forecasting and flood early warning, Hydrol. Earth Syst. Sci., 17, 1161–1175, https://doi.org/10.5194/hess-17-1161-2013, 2013.
2. Beck, H. E., Van Dijk, A. I. J. M., De Roo, A., Miralles, D. G., McVicar, T. R., Schellekens, J., & Bruijnzeel, L. A. (2016). Global-scale regionalization of hydrologic model parameters. Water Resources Research, 52(5), 3599–3622. https://doi.org/10.1002/2015WR018247
3. Bell, V. A., Kay, A. L., Jones, R. G., & Moore, R. J. (2007). Development of a high resolution grid-based river flow model for use with regional climate model output. Hydrology and Earth System Sciences, 11(1), 532–549. https://doi.org/10.5194/hess-11-532-2007
4. Bodian, A., Dezetter, A., Deme, A., & Diop, L. (2016). Hydrological Evaluation of TRMM Rainfall over the Upper Senegal River Basin. Hydrology, 3(2), 15. https://doi.org/10.3390/hydrology3020015
5. Bodian, A., Dezetter, A., Diop, L., Deme, A., Djaman, K., & Diop, A. (2018). Future Climate Change Impacts on Streamflows of Two Main West Africa River Basins: Senegal and Gambia. Hydrology, 5(1), 21. https://doi.org/10.3390/hydrology5010021
6. Choulga, M., Moschini, F., Mazzetti, C., Grimaldi, S., Disperati, J., Beck, H., Salamon, P., & Prudhomme, C. (2024). Technical note: Surface fields for global environmental modelling. Hydrology and Earth System Sciences, 28(13), 2991–3036. https://doi.org/10.5194/hess-28-2991-2024
7. Chu, H., Lin, Y., Huang, C., Hsu, C., & Chen, H. (2010). Modelling the hydrologic effects of dynamic land-use change using a distributed hydrologic model and a spatial land-use allocation model. Hydrological Processes, 24(18), 2538–2554. https://doi.org/10.1002/hyp.7667
8. Coles, S. (2001). An introduction to statistical modeling of extreme values. Springer.
9. Cunderlik, J. (2003). Hydrologic Model Selection for the CFCAS Project: Assessment of Water Resources Risk and Vulnerability to Changing Climatic Conditions. Water Resources Research Report. https://ir.lib.uwo.ca/wrrr/9
10. Descroix, L., Guichard, F., Grippa, M., Lambert, L. A., Panthou, G., Mahé, G., Gal, L., Dardel, C., Quantin, G., Kergoat, L., Bouaïta, Y., Hiernaux, P., Vischel, T., Pellarin, T., Faty, B., Wilcox, C., Malam Abdou, M., Mamadou, I., Vandervaere, J.-P., ... Paturel, J.-E. (2018). Evolution of Surface Hydrology in the Sahelo-Sudanian Strip: An Updated Review. Water, 10(6), 748. https://doi.org/10.3390/w10060748
11. Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A. G., Parizeau, M., & Gagné, C. (2012). DEAP: Evolutionary algorithms made easy. J. Mach. Learn. Res., 13(1), 2171–2175.
12. Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. Annales de la Société Polonaise de Mathématique. 6.
13. Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., Geiger, T., Halladay, K., Hurtt, G., Mengel, M., Murakami, D., Ostberg, S., Popp, A., Riva, R., Stevanovic, M., Suzuki, T., Volkholz, J., Burke, E., Ciais, P., Ebi, K., Eddy, T. D., Elliott, J., Galbraith, E., Gosling, S. N., Hattermann, F., Hickler, T., Hinkel, J., Hof, C., Huber, V., Jägermeyr, J., Krysanova, V., Marcé, R., Müller Schmied, H., Mouratiadou, I., Pierson, D., Tittensor, D. P., Vautard, R., van Vliet, M., Biber, M. F., Betts, R. A., Bodirsky, B. L., Deryng, D., Frolking, S., Jones, C. D., Lotze, H. K., Lotze-Campen, H., Sahajpal, R., Thonicke, K., Tian, H., and Yamagata, Y.: Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b), Geosci. Model Dev., 10, 4321–4345, https://doi.org/10.5194/gmd-10-4321-2017, 2017.
14. Gebremeskel, S., Liu, Y. B., De Smedt, F., Hoffmann, L., & Pfister, L. (2005). Assessing the hydrological effects of Landuse changes using distributed hydrological modelling and GIS. International Journal of River Basin Management, 3(4), 261–271. https://doi.org/10.1080/15715124.2005.9635266
15. Gosling, S. N., Zaherpour, J., Mount, N. J., Hattermann, F. F., Dankers, R., Arheimer, B., Breuer, L., Ding, J., Haddeland, I., Kumar, R., Kundu, D., Liu, J., Van Griensven, A., Veldkamp, T. I. E., Vetter, T., Wang, X., & Zhang, X. (2017). A comparison of changes in river runoff from multiple global and catchment-scale hydrological models under global warming scenarios of 1 °C, 2 °C and 3 °C. Climatic Change, 141(3), 577–595. https://doi.org/10.1007/s10584-016-1773-3

.

16. Gudmundsson, L., Boulange, J., Do, H. X., Gosling, S. N., Grillakis, M. G., Koutroulis, A. G., Leonard, M., Liu, J., Müller Schmied, H., Papadimitriou, L., Pokhrel, Y., Seneviratne, S. I., Satoh, Y., Thiery, W., Westra, S., Zhang, X., & Zhao, F. (2021). Globally observed trends in mean and extreme river flow attributed to climate change. Science (Vol. 371, Issue 6534, pp. 1159–1162). American Association for the Advancement of Science (AAAS). https://doi.org/10.1126/science.aba3996
17. Gumbel, E. J. (1958). Statistics of Extremes. Columbia University Press. https://doi.org/10.7312/gumb92958
18. Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 377(1), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003
19. Haddeland, I., Matheussen, B. V., & Lettenmaier, D. P. (2002). Influence of spatial resolution on simulated streamflow in a macroscale hydrologic model. Water Resources Research, 38(7). https://doi.org/10.1029/2001WR000854
20. Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., Barnard, C., Cloke, H., and Pappenberger, F.: GloFAS-ERA5 operational global river discharge reanalysis 1979–present, Earth Syst. Sci. Data, 12, 2043–2060, https://doi.org/10.5194/essd-12-2043-2020, 2020.
21. Hirpa, F. A., Salamon, P., Beck, H. E., Lorini, V., Alfieri, L., Zsoter, E., & Dadson, S. J. (2018). Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data. Journal of Hydrology, 566, 595–606. https://doi.org/10.1016/j.jhydrol.2018.09.052
22. Jajarmizad, M., Harun, S., & Salarpour, M. (2012). A Review on Theoretical Consideration and Types of Models in Hydrology. Journal of Environmental Science and Technology, 5(5), 249–261. https://doi.org/10.3923/jest.2012.249.261
23. Jensen, L., Dill, R., Balidakis, K., Grimaldi, S., Salamon, P., & Dobslaw, H. (2025). Global 0.05° water storage simulations with the OS LISFLOOD hydrological model for geodetic applications. Geophysical Journal International, 241(3), 1840–1852. https://doi.org/10.1093/gji/ggaf129
24. Koubodana, H. D., Atchonouglo, K., Adounkpe, J. G., Amoussou, E., Kodja, D. J., Koungbanane, D., Afoudji, K. Y., Lombo, Y., & Kpemoua, K. E. (2021). Surface runoff prediction and comparison using IHACRES and GR4J lumped models in the Mono catchment, West Africa. Proceedings of the International Association of Hydrological Sciences, 384, 63–68. https://doi.org/10.5194/piahs-384-63-2021
25. Kwakye, S. O., & Bárdossy, A. (2020). Hydrological modelling in data-scarce catchments: Black Volta basin in West Africa. SN Applied Sciences, 2(4), 628. https://doi.org/10.1007/s42452-020-2454-4
26. Martins, E. S., & Stedinger, J. R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Water Resources Research, 36(3), 737–744. https://doi.org/10.1029/1999WR900330
27. Matthews, G., Baugh, C., Barnard, C., Carton De Wiart, C., Colonese, J., Grimaldi, S., Ham, D., Hansford, E., Harrigan, S., Heiselberg, S., Hooker, H., Hossain, S., Mazzetti, C., Milano, L., Moschini, F., O'Regan, K., Pappenberger, F., Pfister, D., Rajbhandari, R. M., Salamon, P., Ramos, A., Shelton, K., Stephens, E., Tasev, D., Turner, M., van den Homberg, M., Wittig, J., Zsótér, E., & Prudhomme, C. (2025). Chapter 15 - On the operational implementation of the Global Flood Awareness System (GloFAS). In T. E. Adams, C. Gangodagamage, & T. C. Pagano (Eds.), Flood Forecasting (2nd ed., pp. 299–350). Academic Press. https://doi.org/10.1016/B978-0-443-14009-9.00014-6
28. Niel, H., Paturel, J.-E., & Servat, E. (2003). Study of parameter stability of a lumped hydrologic model in a context of climatic variability. Journal of Hydrology, 278(1–4), 213–230. https://doi.org/10.1016/S0022-1694(03)00158-6
29. Parajka, J., Merz, R., & Blöschl, G. (2005). A comparison of regionalisation methods for catchment model parameters. Hydrology and Earth System Sciences, 9(3), 157–171. https://doi.org/10.5194/hess-9-157-2005
30. Pechlivanidis, I. G., Arheimer, B., Donnelly, C., Hundecha, Y., Huang, S., Aich, V., Samaniego, L., Eisner, S., & Shi, P. (2017). Analysis of hydrological extremes at different hydro-climatic regimes under present and future conditions. Climatic Change, 141(3), 467–481. https://doi.org/10.1007/s10584-016-1723-0
31. Pokhrel, P., Gupta, H. V., & Wagener, T. (2008). A spatial regularization approach to parameter estimation for a distributed watershed model. Water Resources Research, 44(12), 2007WR006615. https://doi.org/10.1029/2007WR006615

.

32. Prudhomme, C., Zsótér, E., Matthews, G., Melet, A., Grimaldi, S., Zuo, H., Hansford, E., Harrigan, S., Mazzetti, C., de Boisseson, E., Salamon, P., & Garric, G. (2024). Global hydrological reanalyses: The value of river discharge information for world-wide downstream applications – The example of the Global Flood Awareness System GloFAS. Meteorological Applications, 31(2), e2192. https://doi.org/10.1002/met.2192
33. Rameshwaran, P., Bell, V. A., Davies, H. N., & Kay, A. L. (2021). How might climate change affect river flows across West Africa? Climatic Change, 169(3), 21. https://doi.org/10.1007/s10584-021-03256-0
34. Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., & Dmip Participants, A. (2004). Overall distributed model intercomparison project results. Journal of Hydrology, 298(1–4), 27–60. https://doi.org/10.1016/j.jhydrol.2004.03.031
35. Salamon, P., Grimaldi, S., Disperati, J., Prudhomme, C., Choulga, M., Moschini, F., & Mazzetti, C. (2023). LISFLOOD static and parameter maps for GloFAS. European Commission, JRC132801.
36. Smith, P. J., Pappenberger, F., Wetterhall, F., Thielen del Pozo, J., Krzeminski, B., Salamon, P., Muraro, D., Kalas, M., & Baugh, C. (2016). On the operational implementation of the European Flood Awareness System (EFAS). In T. E. Adams & T. C. Pagano (Eds.), Flood forecasting (pp. 313–348). Academic Press. https://doi.org/10.1016/B978-0-12-801884-2.00011-6
37. Silva Peixoto, J. da, Ernesto de Moraes, M. A., Garcia, K., Broedel, E., Cuartas, A., & Nova da Cruz, P. P. (2024). Performance analysis of the LISFLOOD hydrological model in a flood event in the Madeira River basin. International Journal of Hydrology, 8(2), 38–42. https://doi.org/10.15406/ijh.2024.08.00372jcar-atrace.eu
38. Tang, Q., Oki, T., Kanae, S., & Hu, H. (2007). The Influence of Precipitation Variability and Partial Irrigation within Grid Cells on a Hydrological Simulation. Journal of Hydrometeorology, 8(3), 499–512. https://doi.org/10.1175/JHM589.1
39. Thielen, J., Bartholmes, J., Ramos, M.-H., & De Roo, A. (2009). The European Flood Alert System – Part 1: Concept and development. Hydrology and Earth System Sciences, 13(2), 125–140. https://doi.org/10.5194/hess-13-125-2009
40. Toreti, A., Bavera, D., Acosta Navarro, J., Acquafresca, L., Barbosa, P., De Jager, A., Ficchì, A., Fioravanti, G., Grimaldi, S., Hrast Essenfelder, A., Magni, D., Mazzeschi, M., McCormick, N., Moutia, S., Otieno, V., Salamon, P., Nunes Santos, S., & Volpi, D. (2025). Drought in Africa – April 2025. Publications Office of the European Union. https://data.europa.eu/doi/10.2760/2135988
41. Tramblay, Y., Rouché, N., Paturel, J.-E., Mahé, G., Boyer, J.-F., Amoussou, E., Bodian, A., Dacosta, H., Dakhlaoui, H., Dezetter, A., Hughes, D., Hanich, L., Peugeot, C., Tshimanga, R., & Lachassagne, P. (2021). ADHI: The African Database of Hydrometric Indices (1950–2018). Earth System Science Data, 13(4), 1547–1560. https://doi.org/10.5194/essd-13-1547-2021
42. Tran, Q. Q., De Niel, J., & Willems, P. (2018). Spatially Distributed Conceptual Hydrological Model Building: A Generic Top-Down Approach Starting From Lumped Models. Water Resources Research, 54(10), 8064–8085. https://doi.org/10.1029/2018WR023566
43. Van Der Knijff, J. M., Younis, J., & De Roo, A. P. J. (2010). LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. International Journal of Geographical Information Science, 24(2), 189–212. https://doi.org/10.1080/13658810802549154
44. Wasko, C., Guo, D., Ho, M., Nathan, R., & Vogel, E. (2023). Diverging projections for flood and rainfall frequency curves. Journal of Hydrology, 620, 129403. https://doi.org/10.1016/j.jhydrol.2023.129403
45. Weibull, W. (1951). A Statistical Distribution Function of Wide Applicability. Journal of Applied Mechanics, 18(3), 293–297. https://doi.org/10.1115/1.4010337
46. Wilcox, C., Vischel, T., Panthou, G., Bodian, A., Blanchet, J., Descroix, L., Quantin, G., Cassé, C., Tanimoun, B., & Kone, S. (2018). Trends in hydrological extremes in the Senegal and Niger Rivers. Journal of Hydrology, 566, 531–545. https://doi.org/10.1016/j.jhydrol.2018.07.063
47. Wilson, C. B., Valdes, J. B., & Rodriguez-Iturbe, I. (1979). On the influence of the spatial distribution of rainfall on storm runoff. Water Resources Research, 15(2), 321–328. https://doi.org/10.1029/WR015i002p00321
48. Zábori, J., Betterle, A., Corzo Toscano, M., D´Angelo, C., Garcia Padilla, M., Lemke, C.-D., Arroyo, M. M., Pechlivanidis, I., Sperzel, T., Ziese, M., Grimaldi, S., & Salamon, P. (2024). European flood awareness system

.

: a technical assessment of CEMS EFAS performance during the floods in Northern Germany in December 2023/January 2024, Publications Office of the European Union. https://data.europa.eu/doi/10.2760/1959574

49. Zhao, F., Nie, N., Liu, Y., Yi, C., Guillaumot, L., Wada, Y., Burek, P., Smilovic, M., Frieler, K., Buechner, M., Schewe, J., & Gosling, S. N. (2025). Benefits of Calibrating a Global Hydrological Model for Regional Analyses of Flood and Drought Projections: A Case Study of the Yangtze River Basin. Water Resources Research, 61(3), e2024WR037153. https://doi.org/10.1029/2024WR037153

.