Note: The original comments by Referee 2 (R) are in regular text. Replies by the authors

(A) are colored in green and changes in the text are in italics.

A: We thank the Referee very much for the helpful comments. The suggestions for clarifications in the respective parts of the text have all been implemented and they help to make the study easier to understand and more consistent.

R: General Comments: The paper presents an analysis of the impacts of climate change on the streamflow of the Niger, the Upper Blue Nile, the Ubangi and the Limpopo basins using the SWIM hydrological model and an ensemble of 5 CMIP5 models. This work is one of few examples which draws comparisons between sub-Saharan African river basins. The methodology used and the conclusions drawn are sound, and the manuscript is well structured. However, clarification is required surrounding the reasoning for the selection of the basins and the on both the spatial and temporal modelling scale.

Specific Comments:

P13008 & P13009: The reasoning behind the choice of the four river basins needs to be clarified. The authors state that the river basins chosen cover the "main Sub-Saharan climate zones". Which climate zones are these and on which cassification are they based? The statement by the authors that "they are all highly dependent on the weather conditions, as their economy is mainly based on the primary sector" is confusing and needs to be clarified. Was the climate and the dependence thereon the sole reason for the selection of these basins?

A: We agree that the criteria for the choice of the basins are not clearly communicated in the text. This was also noted by referee 1 and we revised the whole paragraph, avoiding also the confusing statement about the primary sector. Please see response on comment of referee 1 on section 2.1.

R: P13009/13010: Differing catchment rainfall and discharge do not make the catchments "extremely heterogeneous".

A: We agree, and in the revised text the statements about "extreme heterogeneity" are excluded:

P 13009, L15-18: *"However, the diverse climates, topographical and geological conditions, soils, and vegetation types result in characteristic hydrological conditions in each of the basins. This can exemplarily be seen in the broad spectrum of runoff coefficients in the catchments, ranging from about 2% in the Limpopo catchment to 21% in the Oubangui (Fig. 1 and Table 1)."*

P13009.L26-P13010,L2: *"Besides this broad range of climates, the regime of the Niger is substantially influenced by the Inner Niger Delta by delaying the peak runoff and smoothening the hydrograph."*

R: P13011, L4: "This geographical setting results not only in a typical subtropical intraannual, but also a very distinct inter-annual variability of flow". This statement contradicts an early statement that all for

four catchments were “classified as tropical” (P13009). Following this, the next paragraph refers to the Limpopo basin as arid to semi-arid.

A: The classification of all catchment regimes as tropical was meant as being located in the tropics and therefore being mainly influenced by the alteration of dry and wet seasons, i.e. also the subtropics. However, we agree that this is confusing and modified the sentence:

P13009, L13-14: *“The hydrological regimes of all four rivers are characterized by the alternation of dry and wet seasons.”*

The reference Latrubesse et al. 2005 was removed.

R: P13011, L19: Sentence needs revising and clarification - “part of them are considered to have extensive impacts on water resources”.

A: We agree that the sentence sounds confusing and modified it:

P13011, L17-20: *“In addition, there are a lot of mining activities in the Limpopo River Basin with about 1900 mines over the years (Ashton et al., 2001b), and many of them have extensive impacts on water resources (Ashton et al., 2001a).”*

P13011, L21: Sentence needs revising and clarification – “However, water management does not play a major role and currently there are only five major reservoirs in the catchment with volumes of over 1000Mm3, mainly built for irrigation and hydropower: Selingué (Mali), Kainji, Jebba, Shiroro (all three Nigeria) and Lagdo (Cameroon).”

A: We agree that the sentence needs clarification and changed it:

P13011, L21 “In the Niger, water management infrastructure does not influence the streamflow fundamentally on the basin scale. Currently there are only five major reservoirs in the catchment with volumes over 1000 Mm^3, mainly built for irrigation and hydropower: Selingué (Mali), Kainji, Jebba, Shiroro (all three in Nigeria) and Lagdo (Cameroon). These influence the streamflow locally and are included in the model.”

R: P13011: Section 2.2 Water management in the basins. The title of this section is misleading. The section focuses on the number and size of reservoirs and irrigation in the catchments, it does not discuss the how the water is managed per se.

A: We agree that the title is misleading and changed it:

P13011, Section 2.2: *“Human influence on discharge dynamics in the basins”*

R: P13013: The discussion surrounding the modelling scale needs clarification. How does the scale at which the model simulates relate to the objectives? What are the typical sizes of the “subbasins and

hydrotopes"? Are hydrotopes homogeneous hydrological response units? On what basis is the subdivision into these units done?

A: We agree with referee 2, and also referee 1 commented on this topic. Therefore we revised the model description and explained the use of hydrotopes, subbasins and subcatchments in the model, and included information on their average size. The term hydrological response unit is not used anymore (please see response to referee 1, 13015 L1 and 13015 L6-7.

R: P13014, L21: "The climate scenarios have been downscaled" – is this referring to the ensemble of 19 or the selected 5 ESMs? The authors state that both the RCP 2.6 and 8.5 scenarios were used, for each of the 19 or 5 models were both RCP's used or only one per model? Please clarify.

A: We agree that this is not clear in the text. The downscaling refers only to the 5 ESMs and for each of these both RCPs have been used. In order to clarify this we changed two sentences:

P13014, L21-26: *"The five chosen ESMs have been downscaled using a trend-preserving bias correction method with the WFD reanalysis data, and have been resampled on a 0.5 ×0.5 grid for the time period 1950–2099 (Hempel et al., 2013). "Representative Concentration Pathways" (RCP) cover different emission concentrations, and in this study the RCP 2.6 and 8.5 scenarios were used for all 5 ESMs to cover the low and high ends of possible future climatic projections."*

R: P13015, L1: Table 3 does not show the period for calibration

A: We agree that additional information on this is important. Therefore we added the calibration period and results for all basins into table 3, including the additional stations in the Limpopo and Niger basins which have been used to calibrate model for subcatchments. Since in the Niger basin 18 stations have been used for calibration, we provide another table in the supplementary material including all results of the calibration in this basin.

R: P13015, L6: The differences between subcatchments, subbasins and hydrotopes needs to be clarified. Was the modelling still undertaken at the hydrotope scale and the flows accumulated to subcatchment scale? The results for the calibration/validation at the 18 gauge points is not shown in Table 3 (similarly only results for one of the gauging points in the Limpopo is shown) – why was results from only one gauging point for the Niger and Limpopo basins shown?

A: We agree with referee 2 and add all the recommended information in table 3 and the additional table, mentioned in the response to the comment of referee 2 on P13015, L1. In addition we revised the model description completely and the relevant processes are explained in more detail now. Please see the response to comment of referee 2 on 13015 L1.

R: P13017, L10: The authors discuss the model representation of high and low flows. How were these assessed? How are high and low flows defined? Related to this, what is the time-scale of the modelling? Is the model run daily and the results aggregated to monthly? Are the statistics on validation based on monthly or daily results?

A: We agree with referee 2 that the use of high and low flow in the text is not completely consistent as we define high and low flow in section 4.5 explicitly as Q10 and Q90. In the context of validation we referred to low and high flow as difference between the discharge during the rainy and the dry season. Therefore we explain the use of the terms now in section 4.1 and 4.5 in this context:

P13017, 3-5: "The focus of the calibration and model set-up for all four basins was to achieve adequate efficiency for streamflow simulations with daily time step, for mean as well as high and low flows. The high flows refer in this context for discharge peaks during the rainy season and low flows as the minimum discharge during the dry season, and are quantified as Q10 and Q90, correspondingly."

P13023, L1: *"The Q10 value is a robust indicator for high flows and designates a value of river discharge which is only exceeded 10% of the time."*

The related questions about the time scale and aggregation of daily to monthly data and its evaluation has also been noted by referee 1. We agree with both referees that this was not clear in the text so far and additional information is necessary. Therefore we revised section 4.1 and extended Figure 3 with validation results with the daily model outputs. Please see response to comment of referee 1 on P13017, L5.

R: P13017, Sect. 4.2: I am unsure of the reasoning behind the presentation of the biascorrected vs uncorrected scenarios. Please clarify.

A: This was also unclear to referee 1 and we agree that we did not state clearly the intention of the figure. Therefore we added an explanation in the text. Please see response to comment of referee 1 on section 4.2.

R: P13026, Sect. 5.2: Comments on whether changes in flows correspond with previous studies. The seasonality in the title indicates a discussion on changes in the timing of high/low flow periods, which the section does not address.

A: We agree with referee 2 that the title is misleading and changed it:

P13026, Sect. 5.2: *"Changes of streamflow under climate change"*

R: P13029, L22: None of the ESMs seem to represent the high flow peak for the period from 1970 in the Limpopo basin well, the Hadley model overestimates while the other models underestimate. This needs to be acknowledged in the uncertainties.

We agree with referee 2 that the model performance for high flows in the Limpopo basin is not adequately taken into account in the discussion of extremes. Therefore we revised that paragraph:

P13029, L22: *"In the Limpopo basin, the extreme discharge resulting from the simulation driven by the Hadley model can be explained by extremely high rainfall. The high sensitivity to weather extremes in the Limpopo Basin most often results in the very high discharge peaks (Table 1; Fig. 9). However, in the*

Limpopo basin not only the uncertainties originating from the climate models are high but also the performance of the hydrological model for peak discharges is rather weak (Fig. 3).”

R: Technical Comments: P13014, L11: “was applied to calibrate” – change applied to used P13014, L16: “they practically cover” – remove practically

A: We thank referee 2 for this advice and changed the first sentence. The second sentence was completely revised see response on comment of referee 1 P13014 L13-17.

P13014, L11: *“Observed river discharge data from the Global Runoff Data Centre was used to calibrate and validate the model (Fekete et al., 1999).”*