

Answers for reviewer #1:

First, we would like to thank István Zsuffa for the very careful reading of our manuscript and useful suggestions.

C1: I am not sure if the word 'Yellow' in itself is sufficient. This river is usually referred to as 'Yellow River'.

Answer: We added description at page 855 line 15:

“The Yellow River (hereafter called Yellow) source region above the gauging....”

C2 (p852 I1): Comment: include “of”

Answer: Included:

“...less studies can be found which intercompare hydrological models and study propagation of uncertainty along the entire model chain of General Circulation Model (GCM) – Regional Climate Model (RCM) – impact models.

C3 (p852 I6-I8): include “the”

Answer: Included:

“A comprehensive intercomparison of hydrological models has been done, e.g., in the Distributed Model Intercomparison Project (Reed et al., 2004; Smith et al., 2004) comparing the performance of 12 hydrological models in three North American river basins.”

C4 (p854 I5): Something is not clear here:

These five 'climate scenarios' are based on ghg projections, aren't they. How are these projections linked to the projections of the four RCPs mentioned in the next sentence? Maybe these sentence should talk only about the five applied GCMs. Later on it can be mentioned that these GCMs have been driven by the ghg projections of the four RCPs in order to generate the climate scenarios (at least this is what Fig. 4 suggests). This means a total number of 20 climate scenarios for each river - so not 5. Do I misunderstand something?

Answer: We used bias-corrected climate scenarios from five GCMs driven by four RCPs (i.e. 20 different time series). We changed the text as follows:

“The bias-corrected climate scenarios from five GCMs (HadGEM2-ES, IPSL-CM5ALR, MIROC-ESM-CHEM, GFDL-ESM2M, NorESM1-M) driven by four Representative Concentration Pathways (RCPs) were provided by the ISI-MIP project (Hempel

et al., 2013) and used as input for impact assessment. The four RCPs are covering a range of emissions and land-use change projections.”

C5 (p856 I10) : Wrong altitude was given.

Answer: Changed:

“The altitude in the drainage area ranges from 4275ma.s.l. in Swiss Alps to 0 ma.s.l. at Rotterdam.”

C6 (p858 I3): Include “the”

Answer: Included:

“The advantage of HBV is that it covers the most important runoff generating processes”

C7 (p860 I11-I12): Delete “the”

Answer: Deleted:

“For the raster based model VIC a grid resolution of 0.125° was used for ~~the~~ all basins.”

C8 (p860 I22): “According to Section 3.1 and Table 3, MSE was not used. NSE was used in all three hydrological models.”

Answer: The text is correct. PEST minimized MSE (mean square error) and not NSE (Nash and Sutcliffe Efficiency). Later we use the more common NSE for the model validation and for model comparison (and not MSE).

C9 (p860 I22): Use plural.

Answer: Changed:

“As ~~an~~ objective functionss the...”

C10 (p860 I27): Wrong citation style.

Answer: Changed:

“As an objective function the Nash and Sutcliffe Efficiency (Nash and Sutcliffe (1970)”

C11 (p861 I23): Wrong citation style.

Answer: Changed:

“The convergent iterative numerical algorithm for the MM-estimates was provided by (Yohai (1987) “

C12(p862 I14): What are N_{Hyd} , N_{Gcm} and N_{Rcp} ?

Answer: N_{Hyd} , N_{Gcm} , N_{Rcp} describe the number of hydrological models, the number GCMs and the number of RCPs, respectively. This explanation is given at page 863 line 13.

C13 (p862 I14): The five GCMs should also be mentioned here.

Answer: Changed .

C14 (p866 I21): "actually only two (see Fig 4) "

Answer: Yes you are right. Now corrected:

"Panel (B) shows for the Upper Niger a decrease in P with increasing T projected by **two** models (IPSL, Nor), and an opposite trend projected by MIROC and GFDL, with the highest increase in P simulated by MIROC. The projected Had P shows the largest differences for RCP 2.6 and the smallest differences for RCP 8.5 in contrast to the four other GCMs."

C15 (p867 I3): Actually I see only one: GFDL (at 8.5)

Answer: Our formulation refers to monthly changes. So it should be correct.

C16 (p867 I5): ?

Answer: Sentence was improved.

Original: "For the Rhine, four climate models show a decrease of P with increasing T (though with some variations), except the projections by Nor, which firstly show a decrease, and then an increase."

Change to: "For the Rhine, all GCMs project an increase in precipitation for RCP 2.6. Except for Nor all the other GCMs show smaller increase in P for all the higher level RCPs. "

C17 (p868 I25): Figure caption lost.

Answer: Added.

C18 (p868 I2): Figure caption lost.

Answer: Added.

C19 (p874 I8): Delete “the”

Answer: Deleted:

“The projected impacts show the best agreement in the Rhine basin,...,”

C20 (p874 I12): Replace “do” by “draw”

Answer: Replaced:

“For the Upper Niger in Africa, having a monsoonal type of climate, scenarios from Climate models are the largest uncertainty source, and therefore clear conclusions on the projections for future are difficult to ~~do~~ draw.”

C21 (p891 Fig.4): Indicate the scenario periods!

Answer: Included.

C22 (p893 Fig.6): Not clear: Why do simulated flows from the reference period (1961-1990) depend on RCP projections? I understand that they do depend on the chosen hydrological and climate models. But projections are for the future! From the past we have to use historical GHG data as boundary conditions for the climate models. Note that unlike flows simulated P and T from the reference period do not depend on RCPs (see Fig. 4). What is the difference between these graphs and the graphs given on Fig. 5? Do I misunderstand something?

Answer: You are absolutely correct. The simulations for the reference period depend only on GCMs. This is the reason why the 4 plots in the first column of the 4x3 plot matrix are identical. We found it easier to compare between the reference period and the scenario period, when plots are arranged in this way. To avoid the misunderstanding, now we moved the Y axes notations to the right.

Also regarding Fig.5 you are right. The 4 plots in the first column of the 4x3 plot matrix are identical to that given in Fig.5. We decided to include Fig.5 separately and with more details because in Fig. 6 the different lines (5x3) are hardly distinguishable. We thought that the details of hydrological model runs in the reference period are important to show. Also the observed discharges are given in Fig.5 (in contrast to Fig.6).

Answers for reviewer #2:

First we like to thank the reviewer for reading our manuscript.

C1 (p.863, line 4) "But in two three cases..."? Should it be "two or three cases"?

Answer: Three. Text is changed.

C2 (p.863, line5) "In case of no ..." should read "In the case of no.."

Answer: Changed.

C3 (p.863, Sentence following eq. (5)) "The token ? indicates". I could not reproduce the "token" shown on my computer and could not discover where it was in the paper.

Answer: The token "°" (small circle) appears in eq. 3 and eq. 4 as a subscription of Y.

C4 (p. 868. last sentence) Figs ??-8

Answer: Corrected.

C5 (p. 869. first sentence) Fig?? for the Upper Niger

Answer: Corrected

C6 (p. 870, first sentence) should read: "summarizes the results of the evaluation of the

Answer: Changed.

C7 (p.897) the number on the figure axes are too small to read.

Answer: Improved

Answers for reviewer #3:

First we like to thank the reviewer for the very careful reading and for pointing out important questions regarding this publication.

Reviewer's comment:

The manuscript may be improved in several places. The authors should also take into account recent research that addresses global multi-model hydrological studies with several GCMs and global hydrology models (GHMs) conducted within the EU project WATCH (see, e.g. Haddeland et al. 2011; Hagemann et al. 2013). In this way sentence as "Nevertheless ..." (abstract line 6-7) and "However, ..." (Introduction, p. 851, line 19-21) are not really true anymore as WATCH (and later ISIMIP) provided a consistent setup for GCMs-GHM studies that allow to create a more complete picture of uncertainties, and first studies already started to deal with this topic (e.g. Hagemann et al. 2013).

Haddeland et al. 2011 Multi-Model Estimate of the Global Terrestrial Water Balance: Setup and First Results. *J. Hydrometeor.* 12, 10.1175/2011JHM1324.1, 869-884.

Hagemann et al. 2013 Climate change impact on available water resources obtained using multiple global climate and hydrology models. *Earth Syst. Dyn.* 4, 129-144, doi:10.5194/esd- 4-129-2013.

Answer:

We agree that the references on the mentioned publications should be included into the manuscript. Therefore we added the references in the introduction. However, the abovementioned publications rely on the global-scale modeling approaches. Our study differs from that, as it is a regional-scale model application. Still, a comparison between the global-scale and regional scale impact assessments makes sense.

Reviewer's question:

What is the motivation to focus in exactly these catchments?

Answer:

From the general set of possible basins we selected three, which belong to different climatic zones. Besides, the chosen basins belong to the ISI-MIP Phase 2 river basins chosen for intercomparison of climate change impacts on water using the regional-scale models.

Reviewer's question:

Why only smaller sub-basins of Niger and Yellow river are considered?

Answer:

This was done for pragmatic reasons:

1. The complete Niger River basin could not be properly calibrated/validated with HBV and VIC since there is no inundation module included in those models. The flow of the Niger is very much altered by the existence of the Inner Niger Delta and related inundation processes, and the applied hydrological models should be able to model all the relevant processes.
2. The Yellow River in its lower reaches is heavily influenced by human water use (irrigation, abstraction for industry and municipal purposes), and therefore

hydrological modelling of the total river basin requires a lot of water management information, which is not easily available. Therefore, the headwater part, which is crucial for water supply to the whole river basin, was chosen in our case.

3. Besides, the sizes of the three chosen basins are comparable, which is also good for consistency of the study.

Reviewer's question:

Why do they authors do not focus on catchments that have been addressed previously in WATCH or ISIMIP to allow an intercomparison of results?

Answer:

Actually, the chosen basins (and models) belong to the set of the ISI-MIP Phase 2 river basins chosen for intercomparison of climate change impacts on water using the regional-scale models.

Also, there is one major difference between the applied regional approach and the global-scale approach, and that is the model calibration to the observed river discharge which is always applied at the regional scale, but very rarely at the global scale. The global models often fail to represent the measured runoff for the regional-scale basins (Hattermann et. Al, 2013).

Hattermann, F.F., Ch. Müller, V. Krysanova, J. Heinke, M. Flörke, S. Eisner, V. Aich, Sh. Huang, T. Vetter, J. Tecklenburg, S. Fournet, S. Liersch, H. Koch and S. Schaphoff. Bridging the global and regional scales in climate impact assessment: an example comparing two river basis located in humid and semi-arid climate zones. . In: Impacts World 2013 Conference Proceedings. Potsdam : Potsdam Institute for Climate Impact Research, pp: XXX. DOI: 10.2312/pik.2013.001

Reviewer's question:

Why does the study only take into account three different hydrology models? This likely leads to an underrepresentation of uncertainty originating from the hydrology models. This question needs to be properly addressed, especially as the data of more hydrology models are available since within ISIMIP, many more hydrology models have been used to conduct simulations using the ISIMIP forcing.

Answer:

This study is a kind of a pilot study of the ISI-MIP Phase 2, where more river basins (11) and more hydrological models (10-11) will be included. We added a sentence to the discussion:

“In the present study only three hydrological models have been used. This might underestimate the overall uncertainty related to hydrological models. A larger number of models (11) and more river basins (11) will be considered in the second phase of the ISI-MIP project allowing for a more detailed intercomparison of climate change impacts. “

Reviewer's comment:

With regard to the representation of results, i.e. impacts: Instead of describing the differences between single GCM driven simulations it may be more useful to focus on the mean changes and the ranges/spreads between the models.

Answer:

In our point it is important to highlight both aspects: the spread/ranges between the models, and the differences caused by the single GCMs and the hydrological models. We also pointed out some results related to single GCMs, because they represent such an important source of uncertainty. The spreads and averages can be seen e.g. in several figures (e.g. Fig. 7 Fig. A.1- A.3), and they are also discussed.

Reviewer's comment:

At several places the figure caption is repeated in the main text (e.g. p. 866 line 7-12). Please avoid this as this is redundant and unnecessary.

Answer:

The descriptions of Figures in the text are checked, and redundant text in several parts was shortened or removed.

Minor commentsC1 (p. 850 – line 14-16)

Comment: "Sentence is difficult to read. Please rewrite."

Answer: Modified:

"The objectives were to analyze and compare climate impacts on future river discharge and to evaluate uncertainties from different sources"

C2 (p. 850 – line 18)

Comment: ... basins. *Robust* results ...

Answer: Changed

C3 (p. 850 – line 24)

Comment: ... in *the* future ...

Answer: Changed

C4 (p. 854 – line 14)

Comment: ... providing *scenarios, hydrological* models

Answer: Sentence was changed.

C5 (p. 858 – line 10-12) & C6 (p. 858 – line 13-14)

Comment 5: Sentence is difficult to read. Please rewrite.

Comment 6: Sentence is a partial repetition of lines 9-10. Please shorten.

Answer: Rewritten and shortened:

In our study a modified semi-distributed version of the HBV model (HBV-D, Krysanova et al., 1999), with a finer spatial disaggregation into subbasins and 10 elevation zones and up to 15 land cover types, was applied. The modification was based on the Nordic HBV version (Saelthun, 1996).

C7 (p. 859 – line 1-4)

Comment: Sentence is difficult to read. Please rewrite.

Answer: Done:

“The hydrotopes are created by overlaying three maps: subbasins, land use and soil. At the hydrotope level all soil water flows, nutrient cycling in soil and vegetation growth are simulated, based on the principle of similarity, i.e. assuming that units within one subbasin that have the same land use and soil types behave similarly.”

C8 (p. 861 – line 4-7)

Comment: ... *the model, eight for SWIM and 19 for HBV*. Four of *the HBV* parameters are ...

Answer: Improved:

“For VIC only five parameters were used to calibrate the model, *eight for SWIM and 19 for HBV*. For the Upper Niger Basin a reduced number of calibration parameters was used, as all snow related parameters were excluded. “

C9 (p. 863 – line 4)

Comment: It is written: “But in two three cases ...” Two or Three?

Answer: Three. The text was changed.

C10 (p. 863 – line 13)

Comment: ... number *of* GCMs ...

Answer: Included

C11 (p. 863 – line 20-21)

Comment: " ... subsampled *in a way*models, *five* GCMs and *four* RCPs ..."

Answer: "in" included. The numbers (underlined) are correct. Always three hydrological models, three GCMs and three RCPs were combined.

New: "Following the general approach from Bosshard et al. (2013) in the present study the five GCMs, four RCPs and three hydrological Models are subsampled **in** a way that all possible combination of **three** hydrological models, three GCMs and three RCPs are fulfilled."

C12 (p. 864 – line 13)

Comment: What are the calibration and validation periods?

Answer: They are given in Tab. 3.

C13 (p. 868 – line 23)

Comment: ..., respectively. *The* trends ...

Answer: Changed

C14 (p. 868 – line 26 and p. 869 - line 3)

Comment: It is written: ... Figs. ??-8 ...

Answer: Corrected.

C15 (p. 872 – line 2)

Comment: For Q90, *CMs* and ...

Answer: Corrected

C16 (p. 872 – line 22)

Comment: ... of *the* hydrological *model* is ...

Answer: Corrected

C16 (p. 873 – line 9)

Comment: ... of *CMs* is highest ...

Answer: Corrected

C17 – (Fig. 2)

Comment: Insufficient quality. I cannot see the grey line. I suggest using thicker lines and replacing the grey line by a black line.

Answer: Improved

C18 – (Figs. 4, 6, 10)

Figs. are too small and the content can hardly be recognized (especially for Fig. 6). In Fig. 10, I can't separate some of the colours/curves given in the tiny panels below each small panel.

Answer: Fig. 4 and 6 are now larger because of different format compared to discussion paper layout. Font sizes in Fig 10 have been improved. Single lines have been removed and replaced by bands.

List of relevant changes:

List of relevant changes:

p. 850 line 9-11: sentence changed:

“Only a few such studies are available until now with the global-scale hydrological models, and our study is intended as a step in this direction applying the regional-scale models.”

p. 850 line 14-16: sentence changed/shortened:

“The objectives were to analyze and compare climate impacts on future river discharge and to evaluate uncertainties from different sources.”

p. 851 line 8-10: sentence changed:

“Very many studies have been undertaken to investigate climate impacts for a number of sectors, globally and at the regional scale, applying different models and emission scenarios, and there are only a few model intercomparison studies using a consistent setup and global-scale models.”

p. 853 line 11: additional sentences including two additional references have been included:

“Another global scale model intercomparison study was done by Haddeland et al. (2011). They compared simulated runoff and evapotranspiration from six land surface models and five global hydrological models for eight large river basins. Their simulations with uncalibrated models (with one exception) showed high spreads in simulated seasonal and annual river runoff and only little agreement to observed runoff. Doing impact assessment with eight of those eleven models and three GCMs Hagemann et al. (2013) found that in some regions uncertainties caused by the impact models are higher than those caused by the climate models.”

p. 854 line 6-14: Paragraph improved:

“The bias-corrected climate scenarios from five GCMs (HadGEM2-ES, IPSL-CM5ALR, MIROC-ESM-CHEM, GFDL-ESM2M, NorESM1-M) driven by four Representative Concentration Pathways (RCPs) were provided by the ISI-MIP project (Hempel et al., 2013) and used as input for impact assessment. The four RCPs are covering a range of emissions and land-use change projections. The objectives were to analyze and compare climate impacts on future river discharge and to evaluate uncertainties from different sources.”

p. 858 line 12-13: sentence deleted.

p. 861 line 3-7: Paragraph improved:

“The numbers of parameters used for calibration differ between the models and the three basins. For VIC only five parameters were used to calibrate the model, eight for SWIM and 19 for HBV. For the Upper Niger Basin a reduced number of calibration parameters was used, as all snow related parameters were excluded. “

Subsection 3.2. “Evaluation of climate scenarios” subsection partly rewritten:

“For the assessment the time period 1961-1990 was chosen as the reference period, and two periods 2021-2050 and 2061-2090 were considered as two scenario periods SP1 and SP2, respectively. The simulated precipitation (P) and temperature (T) were evaluated and compared between the reference and scenario periods for the three basins.

Fig. 4 shows the long-term average monthly simulated P by five GCMs in the reference period 1961–

1990 in comparison with the observed P. As one can see, the simulated P in the reference period is very close to the observed one. Obviously, this is mainly due to the applied bias correction. The differences between the long-term average simulated P in two scenario periods 2021–2050 and 2061–2090 and that in the reference period for four RCPs are also shown in Fig. 4. Additionally, the long-term average annual changes in P related to the long-term average annual changes in T are given in Fig. 4 for the second scenario period.

For the Upper Niger changes in P vary between climate models, especially in months 4-7: two or three climate scenarios show a decrease in P, and other three or two – an increase. The MIROC scenarios show mostly an increase, which is mostly higher than that simulated by other models, reaching 90-140 mm/month in some months of the second scenario period. Panel (B) shows for the Upper Niger a decrease in P with increasing T projected by two models (IPSL, Nor), and an opposite trend projected by MIROC and GFDL, with the highest increase in P simulated by MIROC. The projected Had P shows the largest differences for RCP 2.6 and the smallest differences for RCP 8.5 in contrast to the four other GCMs. Projections for the Upper Yellow show mostly a small or moderate increase in months 5 – 9, which is mostly below 20 mm/month. Only in the second period MIROC and IPSL project higher increases in some months. According to panel (B) the increase in T is accompanied by an increase in P in almost all cases. The projected changes in monthly P for the Rhine are also moderate, mostly within ± 20 mm/month, only the Had scenarios show a stronger decrease in some months, and MIROC – a higher increase in summer months, contradicting to other models. For the Rhine, all GCMs project an increase in precipitation for RCP 2.6. Except for Nor all the other GCMs show smaller increase in P for all the higher level RCPs. After the model calibration and validation, the simulated long-term average discharge driven by the bias-corrected climate model outputs in the period 1961-1990 was compared with the observed one. As is evident from 5, the simulated and observed discharges agree well for all basins and driving climate models. This is mainly due to the bias-correction of climate model outputs. The agreement between 15 simulations for the reference period is very good for the Upper Niger, good for the Rhine, but weaker for the Upper Yellow.”

Subsection 3.4.1. and 3.4.1: Redundancies with figure captions have been removed.

P. 866 line 18-20: sentence deleted.

P.876 line 28: sentence included:

“In the present study only three hydrological models have been used. This might underestimate the overall uncertainty related to hydrological models. A larger number of models (11) and more river basins (11) will be considered in the second phase of the ISI-MIP project allowing for a more detailed intercomparison of climate change impacts.”

Fig. 11: Single lines have been removed and replaced by ranges and interquartile ranges.

Marked-up manuscript version:

Multi-model climate impact assessment and intercomparison for three large-scale river basins on three continents

Tobias Vetter¹, Shaochun Huang¹, Valentin Aich¹, Tao Yang², Xiaoyan Wang², Valentina Krysanova¹, and Fred Hattermann¹

¹Potsdam Institute for Climate Impact Research, Germany

²Hohai University, China

Correspondence to: Tobias Vetter
(vetter@pik-potsdam.de)

Abstract. Climate change impacts on hydrological processes should be simulated for river basins using validated models and multiple climate scenarios in order to provide reliable results for stakeholders. In the last 10-15 years climate impact assessment was performed for many river basins worldwide using different climate scenarios and models. ~~Nevertheless, the~~ However, their results are hardly comparable and do not allow to create a full picture of impacts and uncertainties. Therefore, a systematic intercomparison of impacts is suggested, which should be done for representative regions using state-of-the-art models. ~~Our~~ Only a few such studies are available until now with the global-scale hydrological models, and our study is intended as a step in this direction applying the regional-scale models. The impact assessment presented here was performed for three river basins on three continents: Rhine in Europe, Upper Niger in Africa and Upper Yellow in Asia. For that, climate scenarios from five GCMs and three hydrological models: HBV, SWIM and VIC, were used. Four 'Representative Concentration Pathways' (RCPs) covering a range of emissions and land-use change projections were included. The objectives were to analyze and compare climate impacts on future ~~trends considering three runoff quantiles: Q90, Q50 river discharge and Q10 and on seasonal~~ water discharge, and to evaluate uncertainties from different sources. The results allow drawing some robust conclusions, but uncertainties are large and shared differently between sources in the studied basins. ~~The robust~~ Robust results in terms of trend direction and slope and changes in seasonal dynamics could be found for the Rhine basin regardless which hydrological model or forcing GCM is used. For the Niger River scenarios from climate models are the largest uncertainty source, providing large discrepancies in precipitation, and therefore clear projections are difficult to do. For the Upper Yellow basin, both the hydrological models and climate models contribute to uncertainty in the impacts, though an increase in high flows in the future is a robust outcome assured by all three hydrological models.

1 Introduction

Setting adequate climate stabilization goals and designing appropriate adaptation policies should rely on a sound quantitative understanding of the expected impacts of climate change under different emission scenarios and different levels of global warming. In particular, a comprehensive assessment of climate impacts is urgently needed within the Intergovernmental Panel on Climate Change (IPCC) process. However, the scientific knowledge about the impacts of climate change still remains fragmentary. Very many studies have been undertaken to investigate climate impacts for a number of sectors, globally and at the regional scale, applying different models and emission scenarios, and there are only a few model intercomparison studies using a consistent setup and global-scale models.

Global assessment of climate change impacts is important to inform the global policy makers, especially regarding mitigation efforts. However, climate impacts occur and adaptation policies are implemented at the regional scale, where the projections

from global impact models may be not precise enough. To make sure that climate impact research meets the demand of stakeholders for reliable information at the regional level, projections of climate impacts should be provided at the regional or river basin scale using validated hydrological models and up-to-date scenarios.

40 Of course, numerous studies on climate change impacts are of value in their own right by providing useful knowledge. However, a quantitative synthesis of climate impacts for different regions, including consistent estimation of uncertainties from different sources, is still missing. In order to achieve it, a systematic intercomparison of impacts simulated by several state-of-the-art models performed for a set of representative regions on all continents using an ensemble of climate scenarios is needed. It is now planned in the Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP), and the first results of the
45 fast-track modelling using global-scale models are already available (Warszawski et al., 2014).

While comparison of climate model outputs has a long tradition and is well established at the global and continental scales (IPCC, 2000, 2007; Déqué et al., 2007; Jacob et al., 2013), less studies can be found which intercompare hydrological models and study propagation of uncertainty along the entire model chain of General Circulation Model (GCM) – Regional Climate Model (RCM) – impact models. The ones which can be found use a variety of methods and techniques to assess the contribution
50 of different sources of uncertainty to the total uncertainty.

A comprehensive intercomparison of hydrological models has been done, e.g., in the Distributed Model Intercomparison Project (Reed et al., 2004; Smith et al., 2004) comparing the performance of 12 hydrological models in three North American river basins. A hierarchical validation testing scheme for model application to runoff predictions in gauged and ungauged basins has been proposed by Refsgaard & Knudsen (1996) using three types of hydrological models (lumped, physically-based, semi-
55 distributed) for a catchment in Zimbabwe. An uncertainty intercomparison of different hydrological models in simulating extreme flows for the upper Yellow River has been published by (Chen et al., 2013) comparing the performance of three different hydrological models under current climate conditions. These studies solely investigate the model performance and related uncertainty of hydrological models without looking at simulation of climate change impacts and the related uncertainty propagation (see also Velázquez et al. (2013)).

60 A set of 10 lumped, semi-distributed and fully-distributed hydrological models has been applied in the project LUCHEM aiming at investigating the envelope of predictions on changes in hydrological fluxes due to land use change considering also land use change scenarios (Breuer et al., 2009).

A comprehensive comparison of uncertainty sources for climate change impacts on flood frequency in England has been published by Kay et al. (2009). Six different sources of uncertainty are discussed for two example catchments (future greenhouse
65 gas emissions, Global Climate Model (GCM) structure, downscaling from GCMs, hydrological model structure, hydrological model parameters and the internal variability of the climate system, sampled by applying different GCM initial conditions). The results suggest that uncertainty from GCM structure is by far the largest source of uncertainty. A probabilistic framework for assessing uncertainties in climate change impacts on low-flow scenarios for the River Thames considering a similar set of uncertainty sources is suggested by Wilby & Harris (2006).

70 All these studies compare results for the regional scale, while Schewe et al. (2014) performed a multi-model assessment of water scarcity under climate change comparing results of twelve global hydrological models driven by five GCM projections. The study highlights the large uncertainties associated with both, climate models and hydrological models. The uncertainty introduced by the hydrological models is particularly dominant in many regions affected by water resources scarcity, suggesting a high potential for improved water resource projections through further improvement of hydrological models.

75 Another global scale model intercomparison study was done by Haddeland et al. (2011). They compared simulated runoff and evapotranspiration from six land surface models and five global hydrological models for eight large river basins. Their simulations with uncalibrated models (with one exception) showed high spreads in simulated seasonal and annual river runoff and only little agreement to observed runoff. Doing impact assessment with eight of those eleven models and three GCMs Hagemann et al. (2013) found that in some regions uncertainties caused by the impact models are higher than those caused by the climate models.

80 Recently, a number of studies were published suggesting the use of Analysis of Variance (ANOVA) approach for uncertainty assessment in hydrology and climatology. For example Yip et al. (2011) uses ANOVA to quantify contribution of different uncertainty sources in climate models, distinguishing between model uncertainty, scenario uncertainty and internal variation. Finger et al. (2012) used ANOVA to estimate impacts of climate model uncertainty on water resources projections and
85 hydro-power production in a glaciated catchment in the Swiss during the 21st century. Ott et al. (2013) applied ANOVA and a non-parametric test to address uncertainty sources in an assessment of high-resolution climate change impacts on medium-sized river catchments in Germany, applying an ensemble of RCM climate forcing data to three hydrological models. Using ANOVA approach Bosshard et al. (2013) assessed uncertainties induced by climate models, two different bias correction methods and hydrological models using output of eight RCMs which are fed into two hydrological models of the Upper Rhine. The results
90 indicate that none of the investigated uncertainty sources are negligible, and some of the uncertainties are not attributable to

individual modeling chain components but rather depend upon their interactions, while in total climate model uncertainty has the largest contribution to the entire uncertainty.

Our study is intended as a contribution to the intercomparison of climate change impacts for the water sector at the regional scale. It was done for three large-scale river basins on three continents: the Rhine in Europe, the Upper Niger in Africa and the Upper Yellow River in Asia by applying three hydrological models: SWIM (Krysanova et al., 1998), HBV (Bergström et al., 1995) and VIC (Liang et al., 1994) after their calibration and validation. ~~Five~~ The bias-corrected climate scenarios from five GCMs ~~:(HadGEM2-ES,IPSL-CM5ALR, MIROC-ESM-CHEM, GFDL-ESM2M, and NorESM1-M)~~ driven by four Representative Concentration Pathways (RCPs) were provided by the ISI-MIP project (Hempel et al., 2013) and used as input for impact assessment. ~~Four 'Representative Concentration Pathways' (RCPs)~~ The four RCPs are covering a range of emissions and land-use change projections ~~were included~~. The objectives were (a) to compare climate impacts on seasonal water discharge, (b) to compare future trends considering three runoff quantiles: Q90, Q50 and Q10 in terms of trend direction and slope, and (c) to evaluate uncertainties from different sources, namely: related to climate models (CMs) providing scenarios, ~~related to~~ hydrological models (HMs) and RCPs.

2 Study areas, data and methods

2.1 Study areas: three river basins

2.1.1 Upper Niger

The Niger River basin is the largest basin of West Africa. Its source is located in the Guinean highlands, from which the river flows in the north-eastern direction through the dry Sahelian zone, and then enters the wetter tropical region north of the Gulf of Guinea. In our study only the Upper Niger catchment until the gauge station Koulikoro (Fig. 1) covering an area of about 122,000 km² was taken for impact assessment. In another study Aich et al. (2014) the whole Niger basin was modeled with SWIM, with several intermediate gauges. The same model setup for the Niger as in Aich et al. (2014) was taken in the present study for SWIM simulations.

The studied upper part of the Niger basin spreads over the countries Guinea and Mali, and includes a small part of the Ivory Coast. The topography of the catchment is quite heterogeneous with several steep-sloped tributaries in Upper Guinea that flow into the flat plane of the Niger River. The dominant land cover in the Upper Niger catchment is forest (34 %) followed by savannah (30 %). The climate in the Upper Niger basin is characterized by a dry season (November - May) and a rainy season from June to September (see Table 1). The rainfall which feeds the river comes mainly from the Guinean Highlands during the rainy season. The average annual precipitation of 1495 mm in the Upper Niger is the highest among the three basins in the study.

The catchment area until Koulikoro is not much influenced by human management. There is only one dam, the Selingue (brought into service in 1982), influencing the discharge downstream until Koulikoro. There are no major irrigation schemes in this part of the catchment.

2.1.2 The Upper Yellow

The Yellow River (hereafter called Yellow) source region above the gauging station Tangnaihai belongs to the Qing-Tibetan Plateau of China. With the drainage area of about 122,000 km² it covers approximately 15 % of the entire Yellow River's drainage basin, while supplying 38 % of the River's total runoff (Chen et al., 2013). This area has been designated as a part of the "Three Rivers' Sources" National Nature Reserve, which was created to protect the source region of the Yellow River, the Yangtze and the Mekong (also called China's water tower supplying water to the whole country). The Upper Yellow flows mainly through pastures, swamps, and knolls between mountains, and crystal clear lakes are characteristic for this area. The two major lakes along this part are Lake Zhaling and Lake Eling, with capacities of 4.7 billion m³ and 10.8 billion m³, respectively. The mean altitude of the drainage area is about 4,000 m.

The Yellow River source region belongs to the cold and dry climate zone with the annual average temperature about -4 °C and average annual precipitation lower than 500 mm (see Table 1), with 70 % of precipitation falling from July to October. The headwaters of the Yellow basin are not much influenced by human activities besides some overgrazing and wild herbs digging.

2.1.3 The Rhine

The Rhine river basin covers a drainage area of about 185,000 km² and spreads over nine European countries. The Rhine river starts from the confluence of two small rivers originating in Switzerland, then forms the Swiss-German and Franco-German

borders before flowing through Germany. In its lower part it enters the Netherlands where it forms an extensive delta and finally releases into the North Sea. Its main tributaries are the Main, the Neckar and the Moselle. Approximately two thirds of the Rhine drainage basin are located in Germany. The altitude in the drainage area ranges from 4275 m.a.s.l. in Swiss Alps to 100 m.a.s.l. at Rotterdam. Regarding its hydrological characteristics, the basin can be subdivided into three major sub-areas: the Alpine area, the German Middle Mountain area and the Lowland area. Two major land cover types in the drainage basin are arable land (38 %) and forest (25 %).

The average annual precipitation of 987 mm in the Rhine is lower than in the Upper Niger but higher than in the Upper Yellow (Tab. 1), and seasonality is not very distinct. In the Alpine region, maximum discharge of the Rhine is observed during summer due to snow melt. Downstream of Basel (close to the Swiss-French-German border crossing), a pluvio-nival hydrological regime of the Rhine gradually becomes dominant. The rainfall dominated tributaries (Moselle and Neckar) contribute to the second maximum discharge of Rhine in winter in this part. In the middle and lower Rhine, the winter peak is higher than the summer one, changing the runoff regime into the pluvial type. Compared to the other two rivers in our study, the Rhine is more influenced by human water management. In this study the Rhine basin was modeled until the gauge Rees (160,000 km²).

2.2 Input data and climate scenarios

The Digital Elevation Model (DEM) constructed from the Shuttle Radar Topography Missions with 90 m resolution was used for altitude information (Jarvis et al., 2008). Soil parameters (soil depth, texture, bulk density) were derived from the Digital Soil Map of the World (FAO et al., 2009), and other parameters needed by models (field capacity - by HBV, porosity, field capacity, available water capacity and saturated conductivity – by SWIM and VIC) were derived using pedotransfer functions and tables (hoc AG-Boden, 2006). Land use was parameterized using the Global Land Cover data (GLCF) (Bartholomé & Belward, 2005). Observed river discharge data from the Global Runoff Data Centre was used to calibrate and validate the hydrological models (GRDC, 1998). As climate input for model calibration the WATCH reanalysis data was applied with the grid resolution of 0.5 degrees (Weedon et al., 2011).

For the climate impact assessment the hydrological models were driven with outputs of five bias corrected Earth System Models of the Coupled Model Intercomparison Project Phase 5 (CMIP5): HadGEM2-ES, IPSL-5 CM5A-LR, MIROC-ESM-CHEM, GFDL-ESM2M, and NorESM1-M. Later the following abbreviations are used for scenarios produced by these models: Had, IPSL, MIROC, GFDL and Nor. Climate scenarios were downscaled to a grid resolution of 0.5 degrees and bias-corrected by the ISI-MIP project using a trend-preserving bias-correction method with the WATCH reanalysis data (Hempel et al., 2013). Four RCPs covering a range of emissions and land-use change projections, from RCP 2.6 to RCP 8.5, were included in the study.

2.3 Hydrological models

Two hydrological models: HBV and VIC, and the ecohydrological model SWIM were used in the study.

2.3.1 HBV

The HBV model (Bergström & Forsman, 1973; Bergström et al., 1995) is a conceptual rainfall-runoff model. The model was developed for runoff simulation and hydrological forecasting. The advantage of HBV is that it covers the most important runoff generating processes by quite simple and robust structures where topographic and climate parameters serve as driving forces. Besides, HBV does not require extensive data sets.

In our study a modified semi-distributed version of the HBV model (HBV-D, Krysanova et al. (1999)), with a finer spatial disaggregation into subbasins and 10 elevation zones and up to 15 land cover types, was applied. The modification ~~included the subbasin level into spatial disaggregation and~~ was based on the Nordic HBV version (Saelthun, 1996). ~~The spatial disaggregation scheme includes now subbasins, 10 elevation zones within every subbasin and up to 15 land use classes.~~

HBV has been applied for modeling hydrological processes in countries with different climatic conditions as for example Sweden, China (Zhang & Lindström, 1996), Zimbabwe (Lidén et al., 2001), and Mozambique (Andersson et al., 2011) and the scales ranging from small catchments to the entire Baltic Sea drainage basin (Graham, 1999). The model is also used worldwide in climate impact assessment studies (Menzel et al., 2006; Yu & Wang, 2009).

2.3.2 SWIM

The ecohydrological model SWIM (Soil and Water Integrated Model, Krysanova et al. (1998)) is a continuous-time spatially semi-distributed model of intermediate complexity for river basins. It integrates hydrological processes, vegetation growth, nutrient cycling, erosion and sediment transport at the river basin scale. SWIM was developed based on two models: SWAT

(Arnold et al., 1993) and MATSALU (Krysanova et al., 1989), with the aim to provide a tool for climate and land use change impact assessment in mesoscale and large river basins. The spatial disaggregation scheme in SWIM includes subbasins and hydrotopes. The hydrotopes are created by overlaying three maps: ~~subbasins~~subbasins, land use and soil, ~~and represent the spatial units used to simulate all water flows and~~. At the hydrotope level all soil water flows, nutrient cycling in soil ~~as well as vegetation growth and vegetation growth are simulated~~, based on the principle of similarity, i.e. assuming that units within one subbasin that have the same land use and soil types behave similarly).

The model was validated and applied for impact assessment in many medium and large river basins in Europe, Africa and Asia (Hattermann et al., 2011; Huang et al., 2013; Liersch et al., 2013). An overview of SWIM application for impact studies is given in (Krysanova et al., 2014).

2.3.3 VIC

The Variable Infiltration Capacity (VIC) model (Liang et al., 1994, 1996) is a semi-distributed hydrological model for large-scale applications. The land surface processes in VIC are modeled as a grid of large (usually >1 km), flat, uniform cells, and the sub-grid heterogeneity (e.g. in elevation, land cover) is handled using statistical distribution functions. The water and energy balances at the land surface can be simulated at a daily or sub-daily time step. The runoff processes are represented through the variable infiltration curve, parameterization of the effects of sub-grid variability in soil moisture holding capacity and a representation of the non-linear baseflow. The modeller can subdivide each grid cell into arbitrary number of tiles, each corresponding to the fraction of the cell covered by a particular land cover (e.g. grassland, coniferous forest, etc.). VIC takes into account snow in several forms: ground snow pack, snow in the vegetation canopy, and snow on top of lake ice. The processes in grid cells are simulated independently, and the routing of water flow is performed separately from the land surface simulation.

VIC has been extensively applied in climate impact studies for a number of large river basins over the continental US and the globe (Hamlet & Lettenmaier, 1999; Su & Xie, 2003; Christensen & Lettenmaier, 2007).

2.3.4 Setup and calibration of three models

The three models differ in their levels of complexity, mathematical process formulation and spatial resolution. For example, vegetation growth is simulated only in SWIM, whereas HBV and VIC use fixed monthly plant characteristics. Spatial resolution of SWIM and HBV is finer than that in VIC, though statistical distribution functions allow to account for vegetation and soil processes in VIC as well. On the other hand, VIC describes land – atmosphere processes with more detail than the other two models. Table 2 describes some major differences between three models. Among them, the differences in soil and vegetation representation are important.

In this study SWIM and HBV were set up with different spatial representation compared to VIC. For the raster based model VIC a grid resolution of 0.125 degrees was used for ~~the~~-all basins. For example, for the Rhine basin 1433 grid cells were simulated by VIC. For SWIM and HBV applications the basins were subdivided into subbasins with an average area of 100-200 km² using SRTM digital elevation model. So, for the Rhine 1668 subbasins were created. All the subbasins and grid cells were further disaggregated considering land use and elevation zones. In addition, SWIM used soil information along with land use for disaggregation of subbasins into hydrotopes.

In total, for the Rhine basin 26.961 units were simulated by VIC, 41.976 hydrotopes were modeled by SWIM, and 69.589 units were simulated by HBV, with an average areas of 5.9 km², 3.8 km², and 2.3 km², correspondingly.

All hydrological models were calibrated using the observed discharge at the basin outlet. For SWIM and VIC automatic calibration was performed with the software package PEST (Doherty, 2005) using the mean square error (MSE) between the observed and simulated discharges as an objective function.

A multi-objective calibration was applied for all basins modeled with HBV using the NSGA II algorithm (Deb et al., 2002). As ~~an objective function~~objective functions the Nash and Sutcliffe Efficiency ~~Nash & Sutcliffe (1970)~~(Nash & Sutcliffe, 1970) of untransformed (NSE) and log-transformed (logNSE) observed and simulated discharges was taken. After optimization the one simulation from the Pareto front was selected, which was closest to the points NSE = 1; logNSE = 1, reflecting the theoretically best possible value for NSE and logNSE.

The numbers of parameters used for calibration differ between the models and the three basins. For VIC only five parameters were used to calibrate the model, ~~and eight parameters were used for the SWIM calibration. For HBV eight for SWIM and 19 parameters were used for model calibration. Four of those parameters are snow related, and they were not applied for the Upper Niger Basin.~~

for HBV. For the Upper Niger Basin a reduced number of calibration parameters was used, as all snow related parameters were excluded.

2.4 Method of trend analysis

Trends in projected runoff were calculated for three runoff quantiles, which reflect the annual high flow conditions (Q10), the annual low flow conditions (Q90) and the medium flow conditions (Q50). To avoid that the same event will be separated between two adjacent years, the hydrological years have been taken instead of the calendar years for calculation of annual Q50 and Q10. As Q90 may occur at the end or at the beginning of a new hydrological year, we used a different approach for this quantile. For calculating Q90, a year starts from a month when, on average, the high flow is reached (different months in our three basins). The quantiles were calculated for the scenario period 2010-2099, and then analyzed for trends.

The trends in the projected runoff quantiles were calculated using the robust linear MM-estimator (Yohai, 1987; Koller & Stahel, 2011). The MM-estimates are calculated by a three-step procedure: 1) a regression estimate with a high breakdown-point, 2) an M-estimate of the errors scale using residuals from step 1, and 3) an M-estimate of the regression parameters based on a proper re-descending psi-function. The M-estimates with a monotone psi-function were introduced by Huber et al. (1973). The convergent iterative numerical algorithm for the MM-estimates was provided by (Yohai, 1987), and are included in the statistical software R. For the trend analysis in our study the function “lmrob” from the package “robust” in R was used.

Compared to the ordinary least squares prediction (OLS), the robust trend estimator is less sensitive to outliers or extreme values and to deviations from the Gaussian distribution. For example, one extremely wet year at the end of the considered time period would have a significant influence on the predicted trend when using the OLS method (bad leverage point), but it would not influence much the trend prediction using the MM estimator. On the other hand, the MM-estimator is usually less efficient (i.e. provides the higher p-values) compared to the OLS prediction. In the present study the standard setup of “lmrob” with an asymptotic relative efficiency of 95 % was applied. The statistically significant trends correspond to the p-values lower than 0.05.

2.5 Method of uncertainty evaluation

For evaluation of different sources of uncertainty an analysis of variance (ANOVA) is performed. ANOVA is a tool for partitioning observed variances into different sources of contributing variation. In the present study three factors are used for variance decomposition (three-way-ANOVA). The total sum of squares (SST) is defined as:

$$SST = \sum_{i=1}^{N_{Hyd}} \sum_{j=1}^{N_{Gcm}} \sum_{k=1}^{N_{Rcp}} (Y_{ijk} - \bar{Y}_{ooo})^2 \quad (1)$$

where Y_{ijk} is the particular value corresponding to hydrological model i , climate model j and RCP k , respectively, and \bar{Y}_{ooo} is the overall mean. According to ANOVA theory SST can be split into seven fractions:

$$SST = \underbrace{SS_{Hyd} + SS_{Gcm} + SS_{RCP}}_{main\ effects} + \underbrace{SS_{Hyd*Gcm} + SS_{Hyd*Rcp} + SS_{Gcm*Rcp} + SS_{Hyd*Gcm*Rcp}}_{interaction\ terms} \quad (2)$$

The total sum of squares is partitioned into three main effects SS_{Hyd} , SS_{Gcm} , SS_{Rcp} corresponding to three hydrological models, five GCMs and four RCPs, respectively. In addition, there are four interaction terms $SS_{Hyd*Gcm}$, $SS_{Hyd*Rcp}$, $SS_{Gcm*Rcp}$, $SS_{Hyd*Gcm*Rcp}$, describing the situation where effects are non-additive or nonlinear. For example the precipitation trends in the Upper Niger (Fig. 2) show noticeable interaction effects. There are strong dependencies of precipitation on the RCPs for each single GCM. But in two-three cases GCMs show negative and in two cases GCMs show a positive trends. In the case of no interactions all the lines would run in parallel.

Exemplarily the calculations of one main effect, one first order interaction term and the second order interactions are given below:

$$SS_{Hyd} = N_{Gcm} N_{Rcp} \sum_{i=1}^{N_{Hyd}} (\bar{Y}_{ioo} - \bar{Y}_{ooo})^2 \quad (3)$$

$$SS_{Hyd*Gcm} = N_{Rcp} \sum_{i=1}^{N_{Hyd}} \sum_{j=1}^{N_{Gcm}} (\bar{Y}_{ijo} - \bar{Y}_{ioo} - \bar{Y}_{ojo} + \bar{Y}_{ooo})^2 \quad (4)$$

$$SS_{Hyd*Gcm*Rcp} = SST - SS_{Hyd} - SS_{Gcm} - SS_{Rcp} - SS_{Hyd*RCP} - SS_{Hyd*Gcm} - SS_{Gcm*Rcp} \quad (5)$$

where N_{Hyd} , N_{Gcm} , N_{Rcp} describe the number of hydrological models, the number of GCMs and the number of RCPs, respectively. The token ° indicates averaging over a particular index.

Bosshard et al. (2013) showed that different sample sizes of the uncertainty sources result in a biased variance estimation. To avoid such a bias Bosshard et al. (2013) complemented the ANOVA with a subsampling scheme. Following the general approach from Bosshard et al. (2013) in the present study the five GCMs, four RCPs and three hydrological Models are subsampled in a way that all possible combination of three hydrological models, three GCMs and three RCPs are fulfilled. So, for the four RCPs we have four combinations (c_1 - c_4) of selecting a subsample with $N_{Rcp} = 3$: , namely: $c_1 = 2.6; 4.5; 6.5$, $c_2 = 2.6; 4.5; 8.5$, $c_3 = 2.6; 6.5; 8.5$ and $c_4 = 4.5; 6.5; 8.5$. For the five GCMs there are 10 possible combinations. These four possible RCP combinations and 10 possible GCM combinations are finally combined to a total number of 40 subsamples. Each of the 40 subsamples contains 27 simulations (using 3 GCMs, 3 hydrological models and 3 RCPs. For each of these 40 subsamples the ANOVA calculation using the formulas above is fulfilled.

After calculation of all the partial sums of squares for all 40 subsamples, the unbiased variance fractions η^2 related to different components can be calculated. For example, the partial variance related to hydrological models can be calculated as:

$$\eta_{Hyd}^2 = \frac{1}{40} \sum_{m=1}^{40} \frac{SS_{Hyd}(m)}{SST(m)} \quad (6)$$

The calculation of the contributing variance fraction for all the other six components is analogous to Formula 6.

3 Results

3.1 Calibration and validation of hydrological models

The results of calibration and validation in terms of fit NSE and percent bias (PBIAS) are presented in Table 3 and Fig. 2 for three models and three basins. In general, the validation results are good, with NSE ranging between 0.81 and 0.93 for the daily time step. The lowest NSE is 0.81 for the VIC application in the Rhine basin. The percent bias values are between -3.6 % and +3.8 % for the simulations in the validation period.

Figure 2 shows a comparison of a 30 day moving average of daily river discharge: simulated by three hydrological models and observed for five years in both the calibration and validation periods, and a comparison of the long-term average seasonal discharges in the calibration and validation periods. As is evident, dynamics are simulated adequately by all three models in the three basins. The river discharges simulated by HBV and SWIM are very similar, and the VIC outputs show some moderate differences in the Upper Yellow with an underestimation of discharge in the first part of the year and an overestimation of discharge in September and October.

In addition, the annual simulated values of Q90, Q50 and Q10 in the calibration and validation periods were compared to those estimated from the observed time series, and results are presented in Fig. 3 and Table 4. All hydrological models show a good performance for high flow conditions (Q10) in all three basins. The coefficient of correlation ranges between 0.78 and 0.97 for the validation period, and the bias is between -13.6 % and +14.6 %, and in 7 cases from 9 it is between -7 % and +8.4 %, for the same period.

In the Rhine basin there is also a good agreement between the observed and simulated Q50 and Q90 values. The lowest coefficient of correlation for the two quantiles is 0.88 for Q90 simulated by VIC. The highest bias for Q90 in the validation period is -17.2 % (VIC). For Q50 the highest bias is +6.3 % for the simulation run with HBV.

The simulation results for Q90 and Q50 in the Upper Yellow and the Upper Niger are not as good as for the Rhine basin. For Q50 in the Upper Niger the coefficient of correlation ranges between 0.78 and 0.92 in the validation period, but a large bias of +40.5 % was found for SWIM. In the Upper Yellow the maximum bias is lower (20.1 % for VIC) but here the correlation between the observations and simulations is lower (ranges between 0.33 and 0.7).

For the low flow simulations (Q90) the biases in the validation period are even higher than for Q50. The maximum bias of +108.4 % was found for the HBV simulations in the Upper Niger. In the Upper Yellow VIC showed a maximum bias of -49.5 %. The coefficient of correlation ranges between 0.18 and 0.91 in the Upper Yellow and between 0.52 and 0.72 in the Upper Yellow.

Summarizing the hydrological model validation, we can conclude that in general the results are good in terms of NSE and PBIAS for river discharge, in terms of correlation and PBIAS for high flows Q10 in all three basins, and for Q50 and Q90 in

the Rhine. However, the results for Q50 and Q90 in the Upper Yellow and the Upper Niger are weaker. In our view, this could be due to high seasonality in runoff in these basins. For example, in the Niger the average Q10 is nearly 100 times higher than Q90. Therefore, low flow gets small weights in the calibration process leading to bias in simulations.

In total, there is no hydrological model which outperforms the others.

3.2 Evaluation of climate scenarios

For the assessment the time period 1961-1990 was chosen as the reference period, and two periods 2021-2050 and 2061-2090 were considered as two scenario periods SP1 and SP2, respectively. The simulated precipitation (P) and temperature (T) were evaluated and compared between the reference and scenario periods for the three basins. ~~The upper panel (A) of~~

Fig. 4 shows the long-term average monthly simulated P ~~by five GCMs in the reference period 1961-1990~~ in comparison with the observed P ~~(left), and the~~. ~~As one can see, the simulated P in the reference period is very close to the observed one. Obviously, this is mainly due to the applied bias correction.~~ The differences between the long-term average monthly-simulated P in two scenario periods ~~2021-2050 and 2061-2090~~ ~~2021-2050 and 2061-2090~~ and that in the reference period for ~~4 RCPs 2.6, 4.5, 6.0, 8.5 (right).~~ ~~The lower panel (B) of four RCPs are also shown in Fig. 4 shows.~~ Additionally, the long-term average annual changes in P related to the long-term average annual changes in T ~~in three basins in are given in Fig. 4 for the second scenario period for 4 RCPs.~~ ~~As one can see, the simulated P in the reference period is very close to the observed one. Obviously, this is mainly due to the applied bias correction.~~

For the Upper Niger changes in P vary between climate models, especially in months 4-7: two or three climate scenarios show a decrease in P, and other three or two – an increase. The MIROC scenarios show mostly an increase, which is mostly higher than that simulated by other models, reaching 90-140 mm/mon in some months of the second scenario period. Panel (B) shows for the Upper Niger a decrease in P with increasing T projected by ~~three models, two models (IPSL, Nor),~~ and an opposite trend projected by MIROC and GFDL, with the highest increase in P simulated by MIROC. ~~Here, RCP8.5-related points are clearly located at the right hands side of the graph, and those related to RCP2.6 – on the left one~~ ~~The projected Had P shows the largest differences for RCP 2.6 and the smallest differences for RCP 8.5 in contrast to the four other GCMs.~~

Projections for the Upper Yellow show mostly a small or moderate increase in months 5 – 9, which is mostly below 20 mm/mon. Only in the second period MIROC and IPSL project higher increases in some months. According to panel (B) the increase in T is accompanied by an increase in P in almost all cases.

The projected changes in monthly P for the Rhine are also moderate, mostly within ± 20 mm/mon, only the Had scenarios show a stronger decrease in some months, and MIROC – a higher increase in summer months, contradicting to other models. For the Rhine, ~~four climate models show a decrease of P with increasing T (though with some variations), except the projections by Nor, which firstly show a decrease, and then an increase~~ ~~all GCMs project an increase in precipitation for RCP 2.6. Except for Nor all the other GCMs show smaller increase in P for all the higher level RCPs.~~

After the model calibration and validation, the simulated long-term average discharge driven by the bias-corrected climate model outputs in the period 1961-1990 was compared with the observed one. As is evident from 5, the simulated and observed discharges agree well for all basins and driving climate models. This is mainly due to the bias-correction of climate model outputs. ~~The agreement between 15 simulations for the reference period is very good for the Upper Niger, good for the Rhine, but weaker for the Upper Yellow.~~

3.3 Impacts on seasonal dynamics

After the calibration and validation of the three hydrological models they were run for the period 1971-2099 using five GCM scenarios for four RCPs providing 60 time series, which were analysed for long-term average seasonal dynamics and trends.

Figure 6 presents the long-term average seasonal discharge for the reference period 1961-1990 ~~(left), for the second scenario period 2061-2090 (middle) and the difference between the second scenario and the reference periods (right).~~ ~~Fifteen trajectories in every box show simulations of three hydrological models (HMs) driven by five climate models (CMs), the colours correspond to the driving CMs. The agreement between 15 simulations for the reference period is very good for the Upper Niger, good for the Rhine, but weaker for the Upper Yellow.~~

For the Upper Niger a high discrepancy between simulations driven by different CMs is visible, and agreement between HMs is higher: three curves of the same colour are close to one another. The increase of discharge projected by simulations driven by MIROC is the highest in the rainy season from July to December (and increases from RCP 2.6 to RCP 8.5), followed by those driven by GFDL, which also show an increase. On the opposite, the IPSL and Nor driven projections show a decrease in discharge in this season, and the simulations using Had scenarios are in between.

For the Upper Yellow River changes in average daily discharge are smaller compared to the Upper Niger. The simulations driven by two CMs: IPSL and MIROC project an increase in second half of the year (months 7-11), whereas the results driven

by three other CMs show rather moderate changes. The magnitude of changes is clearly and steadily increasing from RCP 2.6 to RCP 8.5.

375 For the Rhine a decrease in summer period (results driven by four CMs of five), and a moderate increase in winter time are projected, which corresponds well to the previous impact assessments for this basin (see e.g. Huang et al. (2010)). The MIROC-driven outputs show mostly an increase for all months and RCPs. Uncertainty related to CM is visually higher compared to that related to HM (clustering of curves with the same colour). The projections by SWIM and VIC agree very well (though not visible in Fig. 6).

380 The results driven by the climate model MIROC for all three basins show the highest discharge in the scenario period compared to the outputs driven by four other CMs in almost all cases. This correlates well with higher precipitation projected by MIROC for the studied basins. In general, notably lower uncertainty related to HMs compared to CMs is visible in most cases, especially for the Upper Niger and Rhine.

3.4 Impacts on trends: magnitude of change and direction

385 After that linear trends were calculated using a robust statistical method for three variables: annual median runoff Q50, and low and high annual percentiles Q90 and Q10 representing the low and high flow conditions, respectively, ~~the~~ trends were calculated for the period 2010 — 2099, and the significance of trends was evaluated at the 5 % level. The results of trend analysis for the three basins in terms of the slope of trend and trend direction (and significance) are presented in Fig. 7 – 9, and Fig. A1 - A3.

3.4.1 Slope of trend

390 Two examples of trends for the median flow Q50 simulated by the three hydrological models driven by the Had climate scenario and RCP 8.5 are presented in Fig. 7 for the Upper Niger (~~left~~) and Upper Yellow (~~right~~). As is evident, in the first case all three models agree on a downward trend (all – statistically significant). For the Upper Yellow the models disagree: SWIM shows an increasing trend, and VIC and HBV – a decreasing one.

395 Figure A1 - A3 shows the slopes of trends in Q90, Q50 and Q10. The results are grouped by climate models and by hydrological models. In addition, Fig. 8 shows aggregated results for the slopes of trends as boxplots for four RCPs, ~~showing median lines, degree of dispersion (50 % of all outputs as boxes = interquartil range), and outliers (exceeding median \pm 2.5 x interquartil range, as circles).~~

400 For the Rhine slopes of the trends in two variables Q90 and Q50 are mainly positive under RCP 2.6. The trends are negative under three other RCPs, almost steadily decreasing from RCP 2.6 to RCP 8.5. The slopes of Q10 are positive under RCP 2.6, negative under RCP 8.5, and uncertain under RCP 4.5 and RCP 6.0.

The results in terms of slope of the trend for the Upper Niger are highly uncertain for all variables and RCPs. There is no case where at least 75 % of the model runs show the same trend direction. Only for RCP 2.6 and Q50 more than 75 % model runs agree on declining trends.

405 For the Upper Yellow all three variables, Q90, Q50 and Q10 show an increase under RCP 2.6 and RCP 4.5, and the slopes of trend in Q10 representing high flows are positive for all four RCPs.

The graphs Fig. A1 - Fig. A3 allow to compare visually uncertainties related to CMs and HMs and to conclude that the agreement between HMs is higher than that between CMs, especially for the Upper Niger and Rhine. The outputs driven by MIROC show the highest slopes for the Upper Niger and Upper Yellow.

3.4.2 Trend significance

410 Figure 9 summarizes the results of ~~the~~ evaluation the direction of trends for all 180 time series. The direction of trends in Q90, Q50 and Q10 are grouped by basins (~~horizontal bands~~), four RCPs (~~vertical bands~~), driving climate models (~~horizontal lines~~), and three hydrological models: ~~first arrow corresponds to HBV, second one to SWIM, and third one to VIC. The arrows with statistically significant trends are thicker, and their direction up or down shows an increase or decrease, correspondingly.~~ Only statistically significant trends are analyzed below.

415 For the Niger much more significant trends were found for RCP 8.5 than for the other three. In this case, all simulations (except one) driven by Had and IPSL show a significant downward trend, and almost all simulations driven by MIROC show a significant upward trend. For Q50 the Nor-driven projections corresponding to RCP 8.5 agree with those of Had and IPSL, and two of three GFDL projections agree with those of MIROC. The direction of trends in Q50 for three other RCPs is positive in all simulation driven by MIROC, and in each of these three cases there is a simulation driven by one of other CMs which shows downward trend simulated by all three HMs: GFDL for RCP 2.6, Nor for RCP 4.5, and IPSL for RCP 6.0. The results

420

driven by MIROC show the upward trends in Q50 and Q10 for all RCPs. In total, a high discrepancy between climate models is obvious for this basin.

For the Upper Yellow a downward trend in Q90 and Q50 simulated by HBV and SWIM driven by 3-4 climate models was found for RCP 8.5, whereas the VIC results either do not show a significant trend, or even disagree and show an upward trend (in three cases). The results for Q10 show a moderate agreement in increasing trends, with a growing level of significance from RCP 2.6 to RCP 8.5. Only for the Yellow several cases could be found, where hydrological models disagree and show significant opposite trends.

Evaluating results for the Rhine, we can conclude that with a moderate certainty all three runoff quantiles are showing positive trends for RCP 2.6. The direction of change is different for three other RCPs. A significant downward trend in Q50 and Q90 was found in simulations driven by 3-4 climate models for RCP 4.5 and RCP 6.0. Q90 and Q50 show a stronger downward trend in nearly all simulations for RCP 8.5. Regarding Q10, there is only a small number of significant trends and the trend direction is changing from positive trend in RCP 2.6 scenario to a negative one in RCP 8.5 scenario. In general, a good agreement between CM-driven simulations and HM outputs can be stated for the Rhine.

3.5 Evaluation of uncertainty

The sources of uncertainties were analysed using the ANOVA method described in Section 2.5 for the slopes of the linear trend and changes in long-term average seasonal dynamics (Fig. 10 and Fig. 11). The variance decomposition of the projected slopes for three runoff quantiles is presented in Fig. 10.

For the projected high flow trends (Q10) in the Upper Niger mainly CMs are important uncertainty contributors, followed by the GCM/RCP interactions and followed by the main effects of RCPs. Together, they contribute to about 95 % to the total uncertainty. For Q50 a similar pattern can be observed. For trends in low flow in the Upper Niger, the hydrological models become more important contributors to uncertainty. Their contribution to the overall uncertainty is about as big as the contribution of CMs and the GCM/RCP interactions.

In the Upper Yellow the contribution of the hydrological models to the overall uncertainty of projected trends is more pronounced than in the Upper Niger. For Q90 and Q50 their effect is comparable to the effect of the CMs. Together, they explain 58 % of the variance for Q10 and 49 % of the total variance for Q10, respectively. For the high flows the CMs are dominating contributors, followed by the hydrological models and the GCM/RCP interaction term.

A different pattern can be seen for the Rhine. Here, the RCPs have the highest influence on the projected trends for all three runoff quantiles. For the trends in low flows, the RCPs contribute 70 % to the overall uncertainty. For Q50 and Q10, the CMs become more important contributors of uncertainty, but still lower than RCPs. For Q90, Ms-CMs and RCPs together explain about 70 % and for Q50 85 % of the overall uncertainty. In the Rhine basin, there is not much contribution of the hydrological models on the uncertainty of the projected trends.

The sources of uncertainty related to the changes in the long-term average seasonal dynamics of runoff are shown in Fig. 11a-c. Figure 11a shows the unscaled results for three basins in two selected scenario periods. To highlight the periods of the year where the changes among the different runoff projections are large (high uncertainty) the graphs presented in Fig. 11a were further scaled. The scaling was performed in accordance to the variability in absolute changes (Fig. 11b), and also in accordance to the variability in relative changes (Fig. 11c). As variance estimator the interquartile range of absolute changes (Fig. 11b) and the interquartile range of relative changes (Fig. 11c) was taken for the scaling. When difference between low flow and high flow river runoff are very large, like it is the case in the the Upper Niger as well as in the Upper Yellow, the scaling of the variance contributions in only absolute terms would lead to an under representation of uncertainties with small absolute but high relative changes. This is the reason for using two scaling approaches.

As is evident from Fig. 11a-c, in the Upper Niger the CMs are the main source of uncertainty, especially in the high flow period. In the first four to five months of the year, hydrological models have a higher contribution, but the overall uncertainty is low in this period (see. Fig. 11b). The low flow periods become more pronounced when looking at the relative scaling (Fig. 11c). During this periods the importance of hydrological models is increasing. In general the differences in the variance decomposition between the two selected scenario periods are relatively small.

In the Upper Yellow the hydrological models are about as important as the climate models. The contribution of hydrological models to the overall uncertainty is highest during the second runoff peak in autumn. In this period the uncertainty related to hydrological models is higher than that related to climate models. This pattern is present in both scenario periods. The overall uncertainty in projected seasonal changes is highest during the rainy season, when looking at scaling with absolute differences (Fig. 11b). Looking at relative scaling (Fig. 11c) there is a notable peak at beginning of the year caused by hydrological models. The reason for this peak is not yet clear. The same as in the Upper Niger, the differences between the two scenario periods are small, but influence of hydrological models is getting higher in the second scenario period.

In contrast to these two basins, for the Rhine also the RCPs become more important contributors of uncertainty, but only for the second scenario period (2061-2090), and mainly in summer and autumn. As in the other two basins, looking at the whole year, the variance contribution of CMs is highest, followed by that of hydrological models. For the Rhine, a clear difference between the first and the second scenario periods regarding sources of uncertainty is visible. In the first scenario period, the highest variability in projected absolute changes is in the late autumn and early winter. For 2061-2090, the period of high variability is expanded to the late summer period.

Comparing 10 and Fig. 11 we can conclude that the fractions of uncertainties for the slopes of trends and the long-term average seasonal dynamics are mostly consistent. Only for the Rhine basin, the main sources of uncertainties differ. Namely, whereas for changes in seasonal dynamics the uncertainty related to GCMs is the highest, the RCPs are the main contributors to overall uncertainty for the projected trends in runoff quantiles.

4 Summary and discussion

The study intercompared the climate impacts on runoff generation and river discharge across three river basins on three continents using three hydrological models driven by climate scenarios from five bias-corrected global climate models for four RCPs.

Evaluation and validation of models. The validation of the hydrological models provided good results for river discharge and high flows (Q10) in all three basins, and for the median flow (Q50) and low flow (Q90) in the Rhine. However, the validation results for Q50 and Q90 in the Upper Yellow and the Upper Niger were weaker, probably due to high seasonality of runoff in these basins. No one of the three hydrological models outperformed the other two. The evaluation of climate model outputs in the historical period by comparing simulations driven by climate models with the observed discharge has shown a good agreement.

Robust impacts and uncertainty sources. Regarding the sources of uncertainty in the overall results, we found that the GCM structure is in most cases the largest source of uncertainty for simulated river flows under climate change conditions. The projected impacts show the best agreement in the Rhine basin, despite given differences in precipitation projections. Therefore, the robust results in terms of trend direction and slope and changes in seasonal dynamics could be found for the Rhine River basin regardless which hydrological model or forcing GCM is used.

For the Upper Niger in Africa, having a monsoonal type of climate, scenarios from climate models are the largest uncertainty source, and therefore clear conclusions on the projections for future are difficult to deduce. It is evident that during the rainy season, with high and intensive precipitation, the driving GCM simulations dominate river runoff and contribute most to the total uncertainty. However, during the dry season, when evapotranspiration dominates the hydrological processes, hydrological models clearly contribute much more to the total uncertainty.

For the Upper Yellow in Asia, both the hydrological models and climate models contribute to uncertainty in the impacts, though an increase in high flows in future is a robust result assured by all three hydrological models. In this basin, also having a monsoonal type of climate but lower temperatures than in the Niger basin, the snow melt processes contribute to runoff, and the highest contribution of the hydrological models to the total uncertainty occurs at the end of the rainy season and in winter.

Uncertainty related to RCPs The uncertainty related to RCPs (also called scenario uncertainty) arises due to incomplete knowledge about future emissions. For temperature changes other studies found an increased contribution of scenario uncertainty to the overall uncertainty for the second half of the 21st century (Yip et al., 2011). Regarding precipitation projections, Hawkins & Sutton (2011) found that uncertainties are in general outperformed by climate models, also at the end of the century. The large uncertainty contribution of climate models for precipitation projections is probably the reason for the small contribution of RCPs to the overall uncertainty in the present study. Only in the Rhine RCPs represent an important "driving" factor where the more certain projected trends in temperature are probably more relevant for projected discharges than the precipitation projections.

Uncertainty related to GCMs. The dominance of GCM related uncertainty in impact studies is reported also in other studies, e.g. (Kay et al., 2009). This is not surprising as GCMs are not able yet to reproduce some variables (like precipitation) due to their coarse resolution and current model structure describing related processes. Besides, they cannot be calibrated and validated in the same way against observed data as it is usually done for hydrological models (see Blöschl & Montanari (2010) and the related discussion). In addition, GCMs have more degrees of freedom as they have to model atmosphere and hydrosphere and all feedbacks within one model system at the global scale with only greenhouse gas emissions as driver, whereas the hydrological models are specialized to simulate hydrological processes and are usually calibrated and validated for the region of interest. Also, the hydrological models are very sensitive to climate variability and change making the climate boundary conditions as given by climate scenarios even more important. Summarizing all the results it can be concluded that providing more robust climate scenarios is a pre-condition of obtaining more robust hydrological impacts.

525 **Uncertainty related to hydrological models** It is likely that uncertainty of hydrological models increases with the increase of complexity of hydrological processes in the studied basins. As a result, the largest uncertainty related to hydrological models was found for the Upper Yellow river, where both snow melt and precipitation are important for the runoff generation. Only a minor uncertainty related to hydrological models was found for the Upper Niger, where a simple rainfall-runoff process is prevailing.

530 Looking at the projected long-term average seasonal dynamics, we can see that in certain periods of the year hydrological models contribute almost as much to the total uncertainty as the driving GCMs. The analysis shows that this is normally the case when certain processes simulated by the hydrological models dominate the generation of river runoff, and these processes are simulated differently by three models. During the summer season in the Rhine basin, for example, the water balance is negative (viz. monthly evapotranspiration is higher than precipitation). The evapotranspiration is a process modelled by the hydrological models and the different approaches implemented in the three models contribute more to the overall uncertainty whenever evapotranspiration dominates the water balance. The second period when the contribution of the hydrological models to the overall uncertainty is relatively high in the Rhine basin is during the late winter when snow melt processes, also simulated by the hydrological models, become relevant.

540 Important, when discussing the contribution of hydrological models to the entire scenario uncertainty, is to compare the processes considered in the hydrological models, and complexity of their description. It is often argued that the complex physically-based models ought to be the better choice when performing model projections (Bergström, 1991; Abbott & Refsgaard, 1996). However, the more complex models normally need more parameter input, and problems may occur whenever the additional parameters and processes are sensitive to changes in the boundary conditions. For example, vegetation processes are usually considered in the more complex hydrological models (like SWIM), and parameterized in the more simple hydrological models (like HBV). On the one hand, this may be advantageous for a more realistic description of evapotranspiration, but, on the other hand, under climate change and especially under high end scenarios vegetation cover may change (e.g. summer to winter crops, coniferous to mixed or to deciduous forest, etc.), and the corresponding adjustment of the vegetation module would be needed. Another example is water management: it can be implemented during the reference period, but may become obsolete under climate change conditions, so that for hydrological models which are considering water management also relevant management scenarios should be defined.

545 In the present study only three hydrological models have been used. This might underestimate the overall uncertainty related to hydrological models. A larger number of models (11) and more river basins (11) will be considered in the second phase of the ISI-MIP project allowing for a more detailed intercomparison of climate change impacts.

555 **Uncertainties from different sources: what are the ways to reduce them?** The ideal case would be to improve description of processes in climate and hydrological models so that the climate models agree better in the climate trends for one specific RCP scenario and region, and the hydrological models agree better in impact projections. However, climate processes are very complex due to different feedbacks within the climate system (IPCC, 2007), and some uncertainties will always remain. For example, looking at the changes in precipitation as outlined in the latest IPCC report (IPCC, 2013), one can identify regions where most climate models agree in the trend direction (for example in the tropics with an increase in precipitation and the sup-tropics with mostly a decrease in precipitation), but the transition zones will always be subject to uncertainty as one cannot expect that all climate models will exactly agree on the borderline between changes in precipitation and other climate variables. Regarding hydrological models, more efforts are needed to improve simulation of different processes and their performance in complex basins, e.g. those in higher altitudes where snow and glacier processes are important, and those with human water management playing an important role in overall water balance.

565 When accepting the fact that it is not really possible to decrease substantially the range of uncertainty, one should invest more in analyzing the distribution of uncertainty by gathering more information (e.g. integrating ensembles of climate and hydrological models in the impact study). Doing so, conclusions will get a higher explanatory power as one will learn more about e.g. focal points of change and possible outliers, and thus increase the robustness of the overall results.

570 Besides, some scientists (Greene et al., 2006; Zhu et al., 2013) recommend a model weighting scheme as a feasible approach to reduce uncertainties in climate impact studies. This method gives large weights to the skillful models and minor weights to the models which do not match the observed dynamics. The Impact model intercomparison is still a relatively new field of research (Schewe et al., 2014), and most studies now are focusing on robust results and the sources of uncertainty in terms of model types and data processing (climate models, impact models, bias correction). Less attention is paid to the specific and fundamental processes implemented, and how to improve their description in the models. Regarding the hydrological part, this concerns mainly the runoff generation and related processes, including evapotranspiration, vegetation dynamics, snow melt, etc. The differences in the description of these processes contribute to the total impact uncertainty differently over the season and in various regions. In general, our study shows that the intercomparison of impacts is very important for producing more reliable results of climate impact assessment for the regions and reducing fragmentarity of impacts at the global scale. Besides,

the model intercomparison and analysis of results allow finding the ways to improve climate and impact models and reduce uncertainty for more reliable impact studies in future.

References

- Abbott, M. B. & Refsgaard, J. C. (1996). *Distributed hydrological modeling*. Distributed hydrological modeling. Kluwer Academic Publishers, Dordrecht, Boston, London.
- Aich, V., Liersch, S., Vetter, T., Huang, S., Tecklenburg, J., Hoffmann, P., Koch, H., Fournet, S., Krysanova, V., Müller, E. N., & Hattermann, F. F. (2014). Comparing impacts of climate change on streamflow in four large african river basins. *Hydrology and Earth System Sciences*, 18(4), 1305–1321.
- Andersson, L., Samuelsson, P., & Kjellström, E. (2011). Assessment of climate change impact on water resources in the pungwe river basin. *Tellus A*, 63(1), 138–157.
- Arnold, J. G., Allen, P. M., & Bernhardt, G. (1993). A comprehensive surface-groundwater flow model. *Journal of hydrology*, 142(1), 47–69.
- Bartholomé, E. & Belward, A. (2005). Glc2000: a new approach to global land cover mapping from earth observation data. *International Journal of Remote Sensing*, 26(9), 1959–1977.
- Bergström, S. (1991). Principles and confidence in hydrological modelling. *Nord. Hydrol.*, 22(2), 123–136.
- Bergström, S. & Forsman, A. (1973). Development of a conceptual deterministic rainfall-runoff model. *Nordic hydrology*, 4(3), 147–170.
- Bergström, S. m, S., Singh, V. P., et al. (1995). The hbv model. *Computer models of watershed hydrology.*, (pp. 443–476).
- Blöschl, G. & Montanari, A. (2010). Climate change impacts—throwing the dice? *Hydrological processes*, 24(3), 374–381.
- Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., & Schaer, C. (2013). Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. *WATER RESOURCES RESEARCH*, 49(3), 1523–1536.
- Breuer, L., Huisman, J. A., Willems, P., Bormann, H., Bronstert, A., Croke, B. F. W., Frede, H. G., Graeff, T., Hubrechts, L., Jakeman, A. J., Kite, G., Lanini, J., Leavesley, G., Lettenmaier, D. P., Lindstroem, G., Seibert, J., Sivapalan, M., & Viney, N. R. (2009). Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use. *ADVANCES IN WATER RESOURCES*, 32(2), 129–146.
- Chen, X., Yang, T., Wang, X., Xu, C.-Y., & Yu, Z. (2013). Uncertainty intercomparison of different hydrological models in simulating extreme flows. *Water resources management*, 27(5), 1393–1409.
- Christensen, N. S. & Lettenmaier, D. P. (2007). A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the colorado river basin. *Hydrology & Earth System Sciences*, 11(4), 1417–1434.
- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *Ieee Transactions On Evolutionary Computation*, 6(2), PII S 1089–778X(02)04101–2.
- Déqué, M., Rowell, D., Lüthi, D., Giorgi, F., Christensen, J., Rockel, B., Jacob, D., Kjellström, E., De Castro, M., & van den Hurk, B. (2007). An intercomparison of regional climate simulations for europe: assessing uncertainties in model projections. *Climatic Change*, 81(1), 53–70.
- Doherty, J. (2005). *PEST: Model Independent Parameter Estimation. fifth edition of user manual*. Brisbane, Australia: Watermark Numerical Computing.
- FAO, IIASA, ISRIC, & ISSCAS (2009). Jrc (2009) harmonized world soil database (version 1.1). *FAO, Rome and IIASA, Laxenburg*.
- Finger, D., Heinrich, G., Gobiet, A., & Bauder, A. (2012). Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century. *WATER RESOURCES RESEARCH*, 48.
- Graham, P. (1999). Modelling runoff to the baltic basin. *Ambio*, 28, 328–334.
- GRDC (1998). Global runoff data. *The Global Runoff Data Centre, D-56002 Koblenz, Germany*.
- Greene, A. M., Goddard, L., & Lall, U. (2006). Probabilistic multimodel regional temperature change projections. *JOURNAL OF CLIMATE*, 19(17), 4326–4343.
- Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voss, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P., & Yeh, P. (2011). Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results. *JOURNAL OF HYDROMETEOROLOGY*, 12(5), 869–884.
- Hagemann, S., Chen, C., Clark, D. B., Folwell, S., Gosling, S. N., Haddeland, I., Hanasaki, N., Heinke, J., Ludwig, F., Voss, F., & Wiltshire, A. J. (2013). Climate change impact on available water resources obtained using multiple global climate and hydrology models. *Earth System Dynamics*, 4(1), 129–144.
- Hamlet, A. F. & Lettenmaier, D. P. (1999). Effects of climate change on hydrology and water resources in the columbia river basin. *JAWRA Journal of the American Water Resources Association*, 35(6), 1597–1623.
- Hattermann, F. F., Weiland, M., Huang, S., Krysanova, V., & Kundzewicz, Z. W. (2011). Model-supported impact assessment for the water sector in central germany under climate change—a case study. *Water resources management*, 25(13), 3113–3134.
- Hawkins, E. & Sutton, R. (2011). The potential to narrow uncertainty in projections of regional precipitation change. *CLIMATE DYNAMICS*, 37(1-2), 407–418.
- Hempel, S., Frieler, K., Warszawski, L., Schewe, J., & Piontek, F. (2013). A trend-preserving bias correction - the ISI-MIP approach. *EARTH SYSTEM DYNAMICS*, 4(2), 219–236.

- hoc AG-Boden, A. (2006). *Bodenkundliche Kartieranleitung. 5. verbesserte und erweiterte Auflage*. Stuttgart: Bundesanstalt für Geowissenschaften und Rohstoffe.
- Huang, S., Hattermann, F. F., Krysanova, V., & Bronstert, A. (2013). Projections of climate change impacts on river flood conditions in germany by combining three different rcms with a regional eco-hydrological model. *Climatic change*, 116(3-4), 631–663.
- 640 Huang, S., Krysanova, V., Österle, H., & Hattermann, F. F. (2010). Simulation of spatiotemporal dynamics of water fluxes in germany under climate change. *Hydrological Processes*, 24(23), 3289–3306.
- Huber, P. J. et al. (1973). Robust regression: asymptotics, conjectures and monte carlo. *The Annals of Statistics*, 1(5), 799–821.
- IPCC (2000). Ipcc special report emission scenarios - summary for policymakers. a special report of ipcc working group iii. intergovernmental panel on climate change.
- 645 IPCC (2007). Climate change 2007: Impacts, adaptation and vulnerability - summary for policymakers. working group ii contribution to the fourth assessment report of the intergovernmental panel on climate change. ipcc secretariat.
- IPCC (2013). Climate change 2013: The physical science basis.
- Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., et al. (2013). Euro-cordex: new high-resolution climate change projections for european impact research. *Regional Environmental Change*, (pp. 1–16).
- 650 Jarvis, A., Reuter, H., Nelson, A., & Guevara, E. (2008). Hole-filled srtm for the globe version 4. available from the CGIAR-CSI SRTM 90m Database (<http://srtm.csi.cgiar.org>).
- Kay, A., Davies, H., Bell, V., & Jones, R. (2009). Comparison of uncertainty sources for climate change impacts: flood frequency in england. *Climatic Change*, 92(1-2), 41–63.
- 655 Koller, M. & Stahel, W. A. (2011). Sharpening wald-type inference in robust regression for small samples. *Computational Statistics & Data Analysis*, 55(8), 2504–2515.
- Krysanova, V., Bronstert, A., & Muller-Wohlfeil, D. (1999). Modelling river discharge for large drainage basins: from lumped to distributed approach. *HYDROLOGICAL SCIENCES JOURNAL-JOURNAL DES SCIENCES HYDROLOGIQUES*, 44(2), 313–331.
- Krysanova, V., Hattermann, F., Huang, S., Hesse, C., Vetter, T., Liersch, S., Koch, H., & Kundzewicz, Z. (2014). Modelling climate and land use change impacts with swim: lessons learnt from multiple applications. *Hydrological Sciences Journal*.
- 660 Krysanova, V., Meiner, A., Roosaare, J., & Vasilyev, A. (1989). Simulation modelling of the coastal waters pollution from agricultural watershed. *Ecological modelling*, 49(1), 7–29.
- Krysanova, V., Müller-Wohlfeil, D.-I., & Becker, A. (1998). Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. *Ecological modelling*, 106(2), 261–289.
- 665 Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 99(D7), 14415–14428.
- Liang, X., Wood, E. F., & Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the vic-2l model: Evaluation and modification. *Global and Planetary Change*, 13(1), 195–206.
- Lidén, R., Harlin, J., Karisson, M., & Rahmberg, M. (2001). Hydrological modelling of fine sediments in the odzi river, zimbabwe. *Water SA*, 27(3), 303–314.
- 670 Liersch, S., Cools, J., Kone, B., Koch, H., Diallo, M., Reinhardt, J., Fournet, S., Aich, V., & Hattermann, F. F. (2013). Vulnerability of rice production in the inner niger delta to water resources management under climate variability and change. *Environmental Science & Policy*, 34, 18–33.
- Menzel, L., Thielen, A. H., Schwandt, D., & Bürger, G. (2006). Impact of climate change on the regional hydrology–scenario-based modelling studies in the german rhine catchment. *Natural Hazards*, 38(1-2), 45–61.
- Nash, J. & Sutcliffe, J. (1970). River flow forecasting through conceptual models part i—a discussion of principles. *Journal of hydrology*, 10(3), 282–290.
- Ott, I., Duethmann, D., Liebert, J., Berg, P., Feldmann, H., Ihringer, J., Kunstmann, H., Merz, B., Schaedler, G., & Wagner, S. (2013). High-Resolution Climate Change Impact Analysis on Medium-Sized River Catchments in Germany: An Ensemble Assessment. *JOURNAL OF HYDROMETEOROLOGY*, 14(4), 1175–1193.
- 680 Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., , & Participants, D. (2004). Overall distributed model intercomparison project results. *Journal of Hydrology*, 298, 27 – 60. The Distributed Model Intercomparison Project (DMIP).
- Refsgaard, J. C. & Knudsen, J. (1996). Operational validation and intercomparison of different types of hydrological models. *Water Resources Research*, 32(7), 2189–2202.
- 685 Saelthun, N. R. (1996). The nordic hbv model. *Norwegian Water Resources and Energy Administration Publication*, 7, 1–26.
- Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., Dankers, R., Eisner, S., Fekete, B. M., Colon-Gonzalez, F. J., Gosling, S. N., Kim, H., Liu, X., Masaki, Y., Portmann, F. T., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L., & Kabat, P. (2014). Multimodel assessment of water scarcity under climate change. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 111(9), 3245–3250.
- 690 Smith, M. B., Seo, D.-J., Koren, V. I., Reed, S. M., Zhang, Z., Duan, Q., Moreda, F., & Cong, S. (2004). The distributed model intercomparison project (dmip): motivation and experiment design. *Journal of Hydrology*, 298(1–4), 4 – 26. The Distributed Model Intercomparison Project (DMIP).
- Su, F. & Xie, Z. (2003). A model for assessing effects of climate change on runoff in china. *Progress in Natural Science*, 13(9), 701–707.

- 695 Velázquez, J., Schmid, J., Ricard, S., Muerth, M., St-Denis, B. G., Minville, M., Chaumont, D., Caya, D., Ludwig, R., & Turcotte, R. (2013).
An ensemble approach to assess hydrological models contribution to uncertainties in the analysis of climate change impact on water
resources. *Hydrol. Earth Syst. Sci.*, 17, 565–578.
- Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., & Schewe, J. (2014). The Inter-Sectoral Impact Model Intercomparison
Project (ISI-MIP): Project framework. *PNAS*, 111(9), 3228–3232.
- 700 Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., & Best, M. (2011).
Creation of the watch forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth
century. *Journal of Hydrometeorology*, 12(5), 823–848.
- Wilby, R. L. & Harris, I. (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the river thames,
uk. *Water Resources Research*, 42(2).
- 705 Yip, S., Ferro, C. A. T., Stephenson, D. B., & Hawkins, E. (2011). A Simple, Coherent Framework for Partitioning Uncertainty in Climate
Predictions. *JOURNAL OF CLIMATE*, 24(17), 4634–4643.
- Yohai, V. J. (1987). High breakdown-point and high-efficiency robust estimates for regression. *Annals of Statistics*, 15(2), 642–656.
- Yu, P.-S. & Wang, Y.-C. (2009). Impact of climate change on hydrological processes over a basin scale in northern taiwan. *Hydrological
processes*, 23(25), 3556–3568.
- 710 Zhang, X. & Lindström, G. (1996). A comparative study of a swedish and a chinese hydrological model1. *JAWRA Journal of the American
Water Resources Association*, 32(5), 985–994.
- Zhu, J., Forsee, W., Schumer, R., & Gautam, M. (2013). Future projections and uncertainty assessment of extreme rainfall intensity in the
United States from an ensemble of climate models. *CLIMATIC CHANGE*, 118(2), 469–485.

Table 1. Characteristics of the three river basins

	Rhine (Rees)	Upper Yellow (Tanghaihai)	Upper Niger (Koulikoro)
Area [km ²]	160.000	110.000	122.000
Altitude range (min/mean/max) [m]	10/495/4275	2673/4256/6248	289/463/1407
Average temperature (1971-2000) [°C]	8.6	-2.0	26.5
Temp. of coldest/warmest month [°C]	0.3/17.4	-14.2/8.23	23.8/28.6
Annual precipitation (1971-2000) [mm]	987	520	1495
Prec. of driest/wettest month [mm]	69/97	0/113	3/323
Dominant land cover [%]	cropland 38 forest 25 grassland 9	grassland 90 bare soil 4 heather 3	forest 34 savanna 30 cropland 24

Table 2. Differences between spatial disaggregation, climate input and representation of main components in three hydrological models used in the study

Features	HBV	SWIM	VIC
Spatial disaggregation	Subbasins, ten elevation zones and land use classes within them	Subbasins and hydrotops (based on land use and soil types within subbasins)	Grid cells, sub-grid heterogeneity (elevation, land cover) is handled via statistical distributions
Climate data input	2 parameters: T_{mean}^b , precipitation	6 parameters: T_{min}^a , T_{mean}^b , T_{max}^c , precipitation, air humidity, radiation	5 parameters: T_{min}^a , T_{max}^c , precipitation, air humidity, wind speed
Representation of soils	One soil layers, 2 soil parameters	Up to 10 soil layers, 11 soil parameters	typically 3 soil layers, 19 parameters
Representation of vegetation	Fixed monthly plant characteristics	Simulation of plant growth using EPIC approach	Fixed monthly plant characteristics
Calculation of potential evapotranspiration	Blaney-Criddle	Priestley-Taylor	Penman-Monteith
Method to calculate snow melt	Degree-day	Extended degree-day method	Two-layer energy-balance model at the snow surface
Runoff routing method	Simple time lag method	Muskingum method	Linearized St. Venant's equations

^a minimum temperature^b mean temperature^c maximum temperature

Table 3. Calibration and validation results with the daily time step. NSE = Nash & Sutcliffe efficiency; pbias = percent bias

	Calibration			Validation		
Rhine	HBV	VIC	SWIM	HBV	VIC	SWIM
Period	1981-1990	1981-1990	1981-1990	1991-2000	1991-2000	1991-2000
NSE	0.75	0.81	0.87	0.83	0.81	0.9
pbias [%]	0.0	-3.6	0.6	3.8	-9.3	-3.6
Upper Yellow	HBV	VIC	SWIM	HBV	VIC	SWIM
Period	1961-1970	1961-1970	1961-1970	1971-1980	1971-1980	1971-1980
NSE	0.89	0.79	0.82	0.88	0.75	0.75
pbias [%]	1.5	-6.9	-0.1	2.6	0.4	4.6
Upper Niger	HBV	VIC	SWIM	HBV	VIC	SWIM
Period	1961-1970	1961-1970	1961-1970	1971-1980	1971-1980	1971-1980
NSE	0.91	0.88	0.92	0.93	0.87	0.91
pbias [%]	6.8	3.3	2.8	3.4	-5.7	2.0

Table 4. Coefficient of correlation (cor) and percent bias (pbias) for the three runoff quantiles (Q90, Q50, Q10) in the calibration and the validation period

	Calibration			Validation		
Rhine	HBV	SWIM	VIC	HBV	SWIM	VIC
Period	1981-1990	1981-1990	1981-1990	1991-2000	1991-2000	1991-2000
Q90 cor	0.96	0.94	0.9	0.93	0.93	0.88
Q90 pbias	4.4	0.0	-13.0	5.8	-8.0	-17.2
Q50 cor	0.93	0.94	0.94	0.95	0.98	0.98
Q50 pbias	3.8	8.0	3.6	6.3	1.1	-4.7
Q10 cor	0.95	0.94	0.93	0.96	0.96	0.95
Q10 pbias	-3.9	-3.4	-5.8	0.4	-1.9	-7.5
Upper Yellow	HBV	SWIM	VIC	HBV	SWIM	VIC
Period	1961-1970	1961-1970	1961-1970	1971-1980	1971-1980	1971-1980
Q90 cor	-0.16	0.66	0.02	0.18	0.91	0.25
Q90 pbias[%]	39.6	1.9	-56.2	25.4	14.4	-49.5
Q50 cor	0.72	0.82	0.87	0.33	0.70	0.57
Q50 pbias[%]	4.2	-3.4	-35.0	-3.5	5.7	-20.1
Q10 cor	0.84	0.93	0.88	0.97	0.89	0.79
Q10 pbias[%]	3.1	-1.5	11.3	6.0	8.4	14.6
Upper Niger	HBV	SWIM	VIC	HBV	SWIM	VIC
Period	1961-1970	1961-1970	1961-1970	1971-1980	1971-1980	1971-1980
Q90 cor	0.74	0.79	0.26	0.52	0.34	0.72
Q90 pbias	61.6	4.2	37.6	108.4	27.0	63.2
Q50 cor	0.69	0.40	0.08	0.92	0.83	0.78
Q50 pbias	30.1	38.1	41.1	9.6	40.5	-4.0
Q10 cor	0.71	0.84	0.71	0.78	0.81	0.79
Q10 pbias	2.1	-1.7	-2.7	-3.8	-10.9	13.6

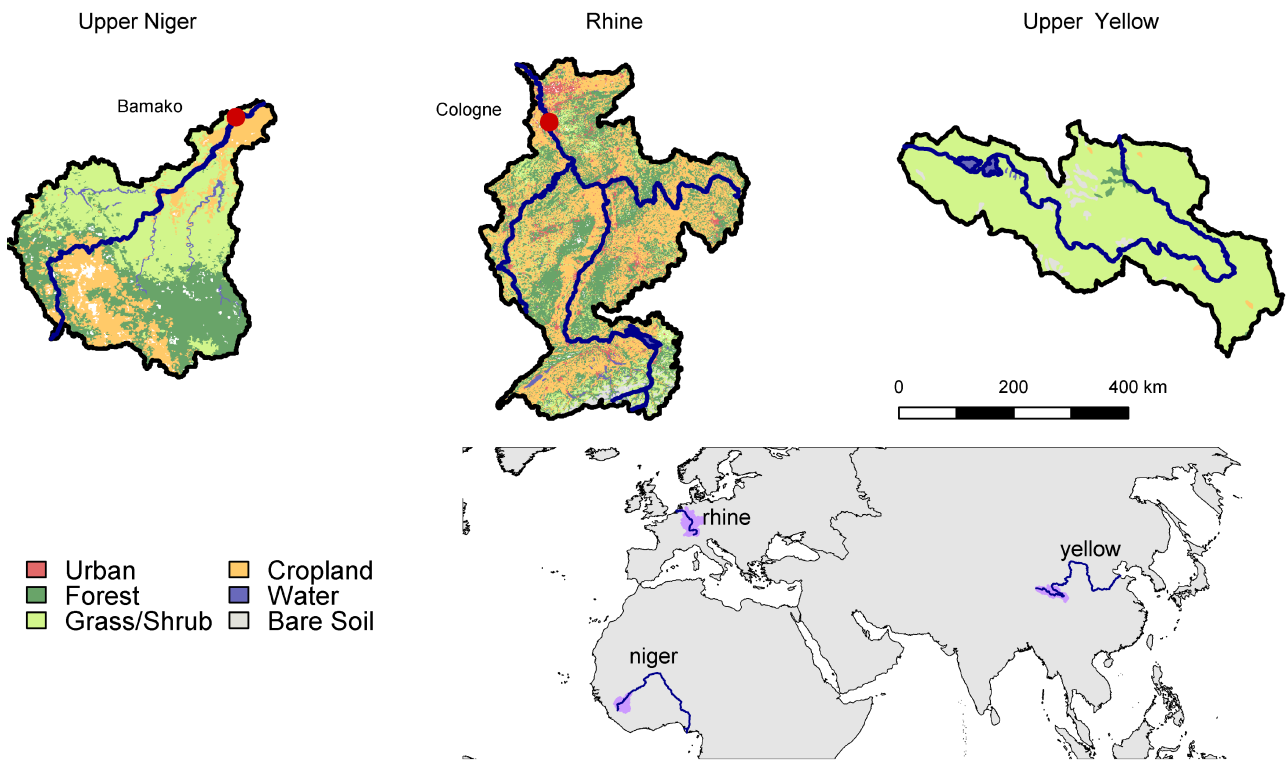


Figure 1. Land use maps of three basins under study: the Upper Niger, the Rhine and the Upper Yellow

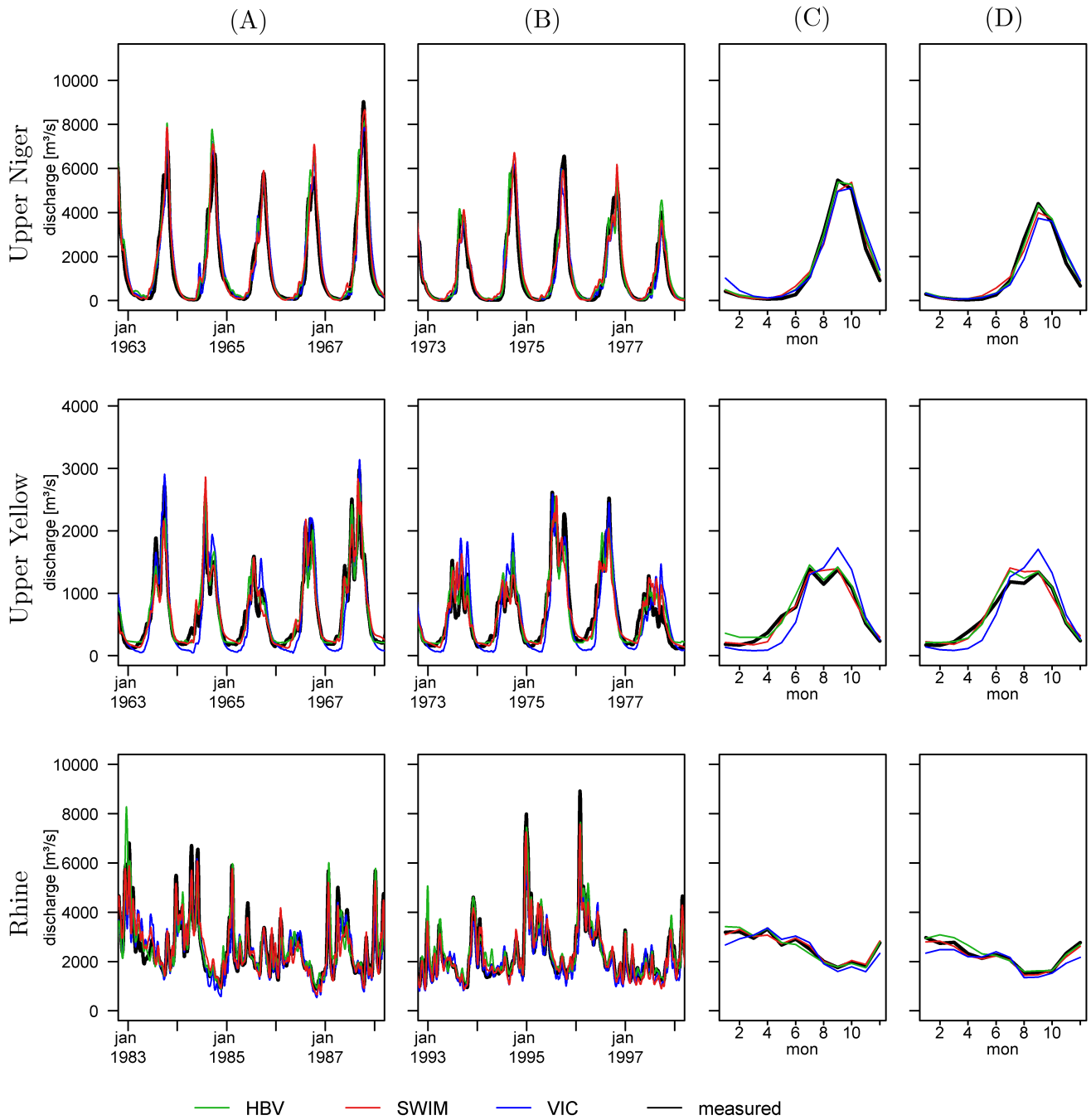


Figure 2. Results of calibration and validation of three models in three basins: comparison of the simulated and observed 30-day-moving-average discharges for five years in the calibration (A) and validation (B) periods, and comparison of the long-term average seasonal discharges in the calibration (C) and validation (D) periods.

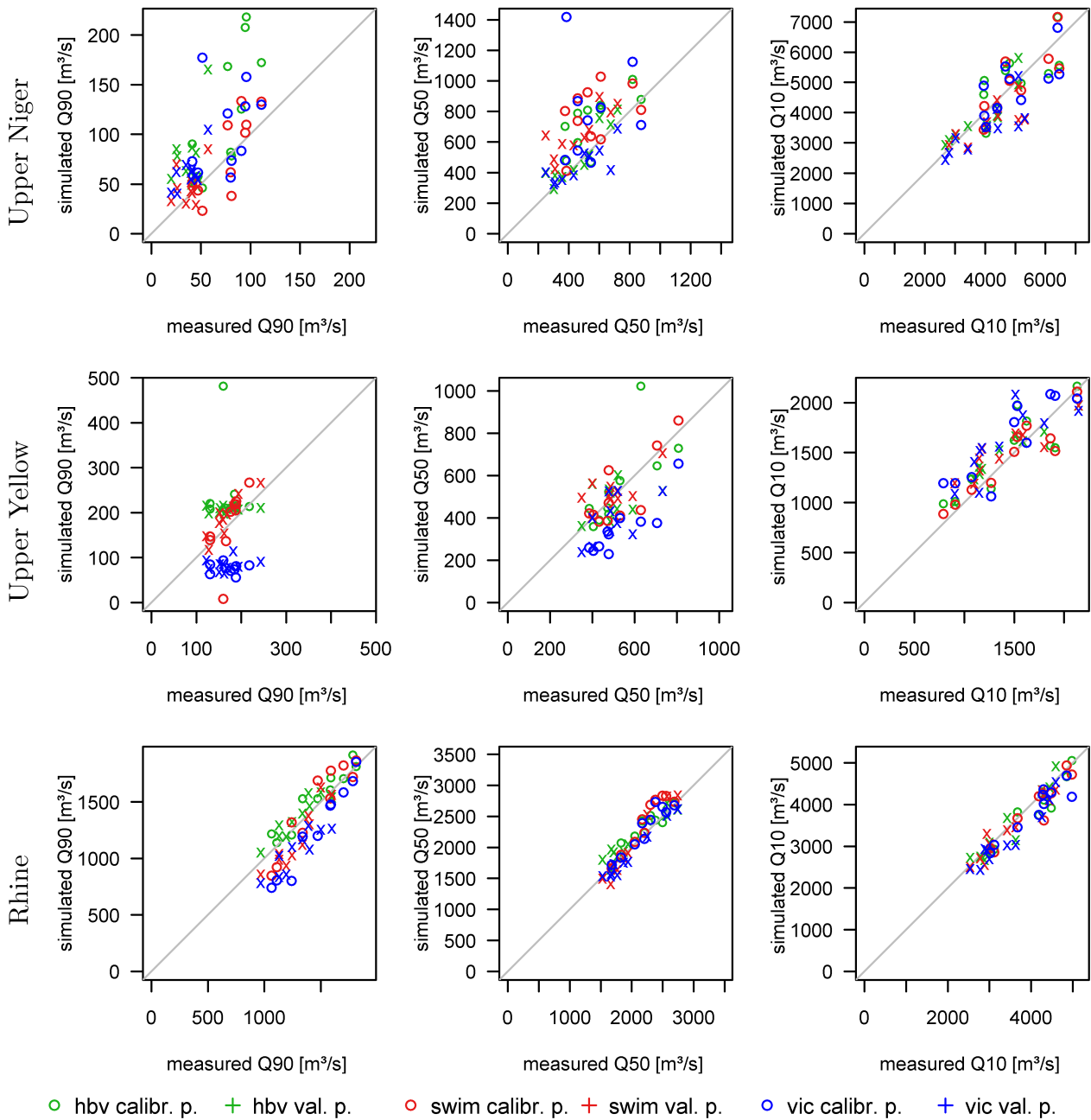


Figure 3. Comparison of simulated and observed annual values of Q90, Q50 and Q10 in the calibration and validation periods for three basins: the Upper Niger, Upper Yellow and Rhine.

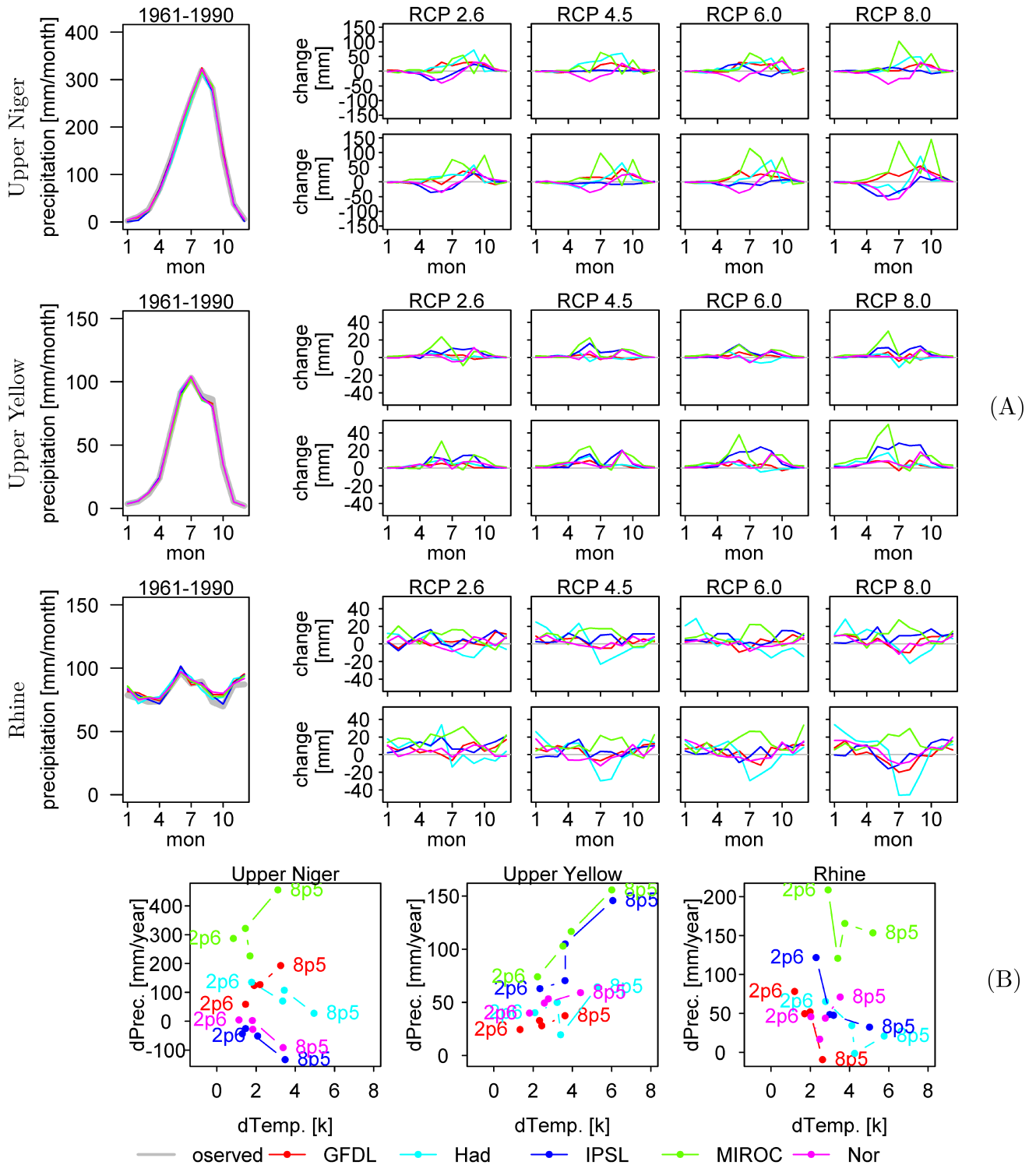


Figure 4. Evaluation of precipitation (P) and temperature (T) simulated by five climate models in three basins: (A) left: comparison of the observed and simulated long-term average monthly P in the reference period 1961-1990; right: differences between simulated P in two scenario periods 2021-2050 and 2061-2090 and that in the reference period for 4 RCPs 2.6, 4.5, 6.0, 8.5; and (B) the long-term average annual changes in P (dPrec) in relation to the long-term average annual changes in T (dTemp) in the second scenario period related to those in the reference period for four RCPs.

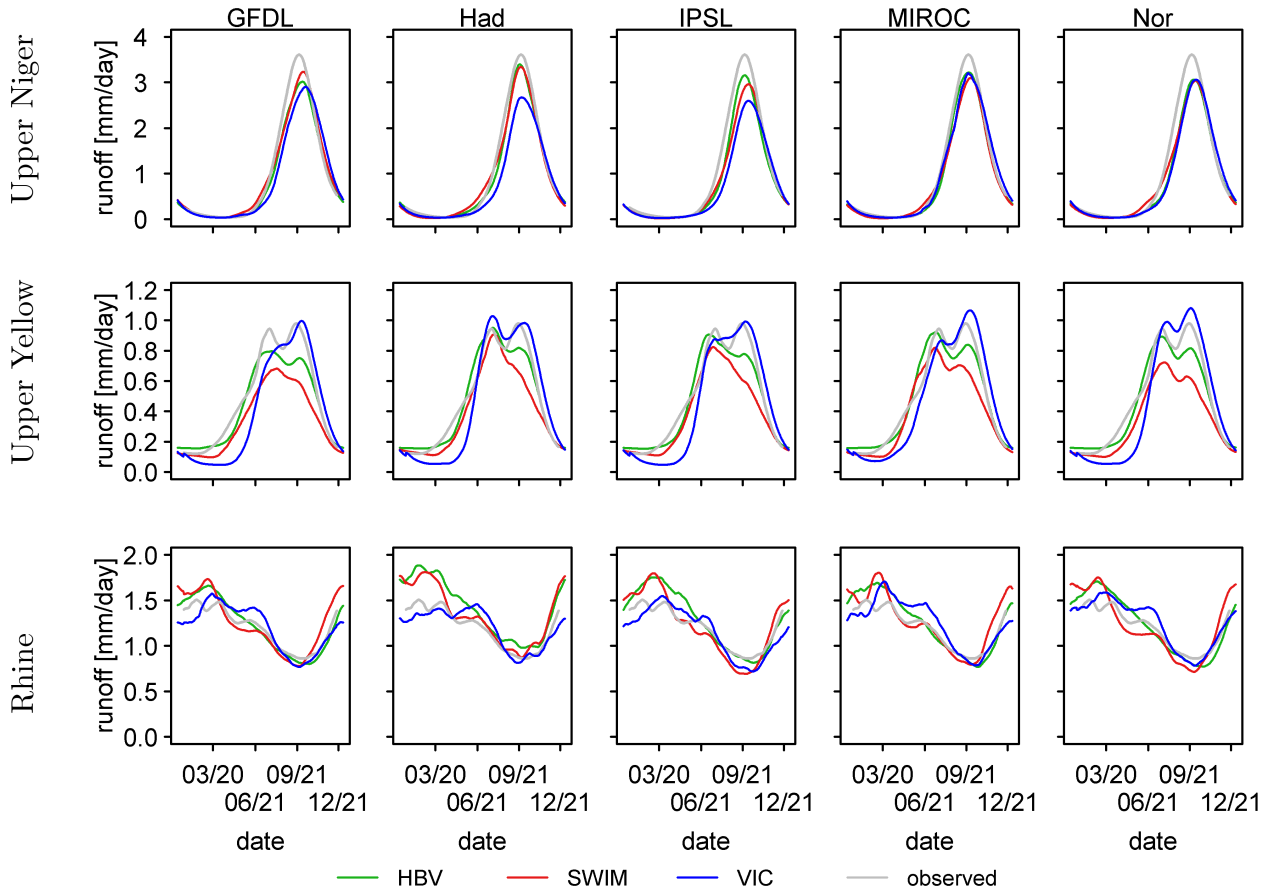


Figure 5. Comparison of the long-term average seasonal observed discharge in 1961-1990 with discharge driven by five climate models and three hydrological models for the same period

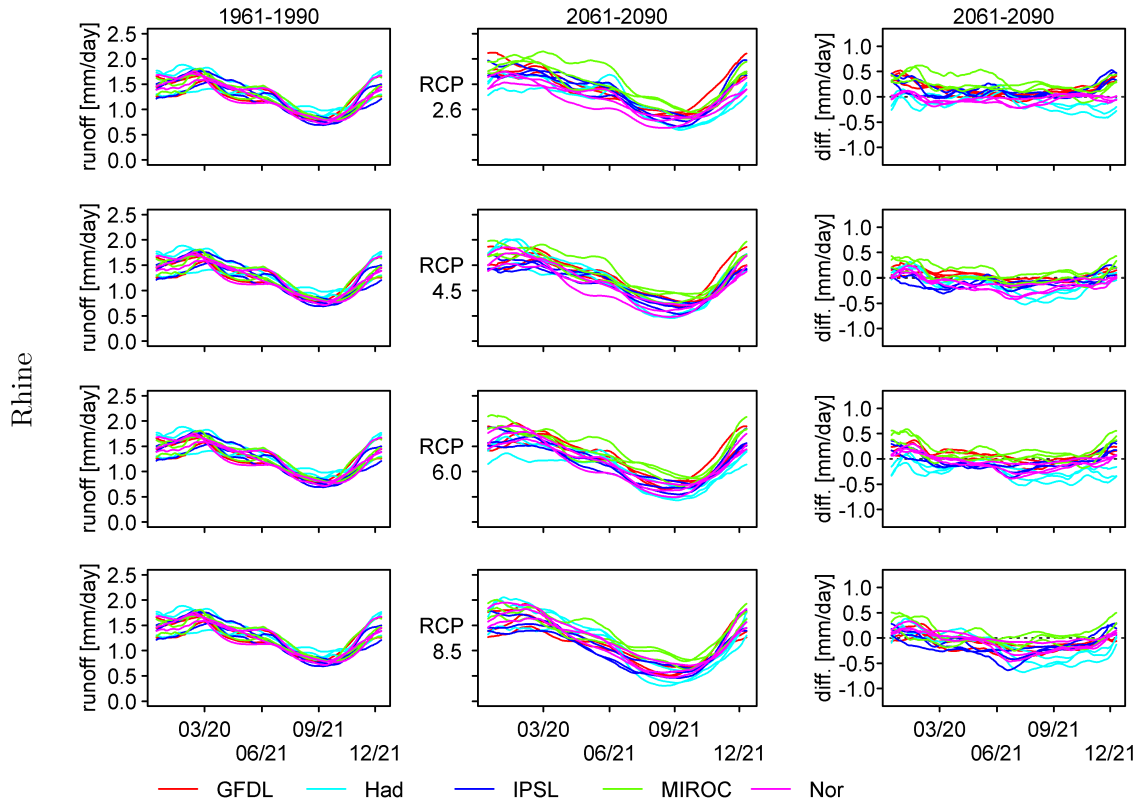


Figure 6. Simulated long-term average seasonal dynamics of river discharge in the reference period 1961-1990 and scenario period 2061-2090 for four RCPs, and the differences in discharge between these two periods; simulations are differentiated by colors corresponding to climate models, whereby three simulations produced by three hydrological models have the same color.

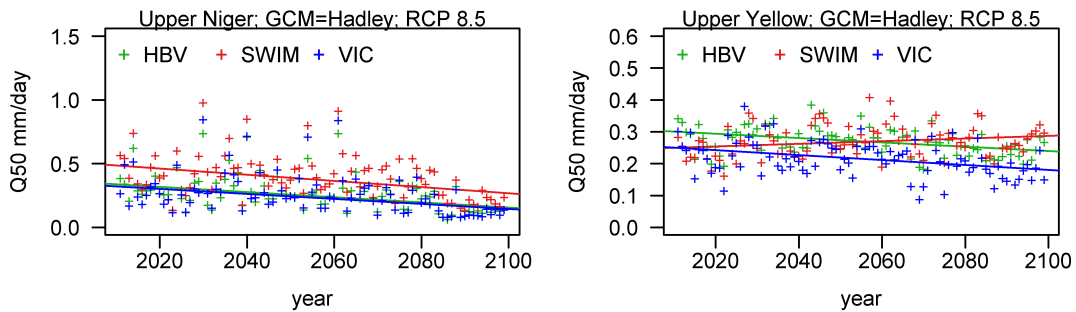


Figure 7. Two examples showing annual median flows Q50 estimated from simulations by three models driven by the Had climate scenario data corresponding to RCP8.5 and the corresponding trend lines (all presented trends are significant) for the Upper Niger (left) and Upper Yellow (right) in the period 2010 - 2099.

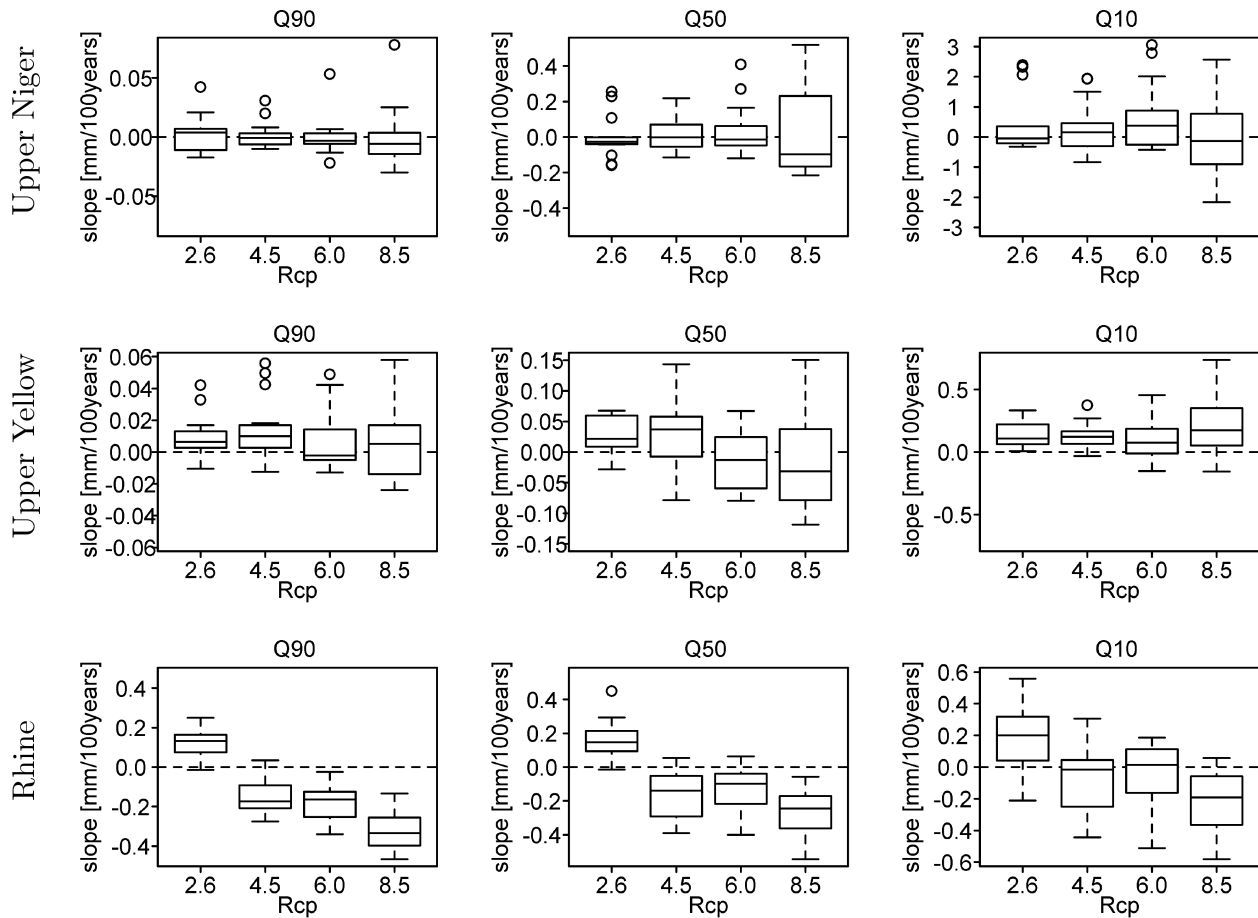


Figure 8. Boxplots for the slopes of the linear trend for Upper Niger, Upper Yellow and Rhine in three runoff quantiles grouped by RCPs. Outliers as circles. outlier= distance to median (center line of box) exceeding 2.5 times the interquartile range (length of the box)

Rcp 2.6				Rcp 4.5				Rcp 6.0				Rcp 8.0			
Niger	Q90	Q50	Q10	Niger	Q90	Q50	Q10	Niger	Q90	Q50	Q10	Niger	Q90	Q50	Q10
GFDL	↓↓↓	↓↓↓	↑↓↓	GFDL	↓↑↓	↓↑↑	↑↑↑	GFDL	↓↑↓	↓↓↓	↓↓↓	GFDL	↓↑↑	↑↑↑	↑↑↑
Had	↓↓↓	↓↓↓	↓↓↓	Had	↑↑↑	↓↑↑	↑↑↑	Had	↓↑↓	↓↑↑	↑↑↑	Had	↓↓↓	↓↓↓	↓↓↓
IPSL	↑↑↑	↓↓↓	↓↓↓	IPSL	↓↓↓	↓↓↓	↓↓↓	IPSL	↓↓↓	↓↓↓	↓↓↓	IPSL	↓↓↓	↓↓↓	↓↓↓
MIROC	↑↑↑	↑↑↑	↑↑↑	MIROC	↓↑↓	↑↑↑	↑↑↑	MIROC	↓↑↑	↑↑↑	↑↑↑	MIROC	↑↑↑	↑↑↑	↑↑↑
Nor	↑↑↑	↓↓↓	↑↑↑	Nor	↓↓↓	↓↓↓	↓↓↓	Nor	↓↑↓	↓↑↑	↑↑↑	Nor	↓↑↑	↓↓↓	↓↓↓
Yellow	Q90	Q50	Q10	Yellow	Q90	Q50	Q10	Yellow	Q90	Q50	Q10	Yellow	Q90	Q50	Q10
GFDL	↑↑↓	↑↑↑	↑↑↑	GFDL	↓↓↑	↓↓↑	↓↑↑	GFDL	↓↓↓	↓↓↓	↓↓↓	GFDL	↓↓↑	↓↓↓	↓↓↑
Had	↑↑↑	↑↑↑	↑↑↑	Had	↑↑↑	↑↑↑	↑↑↑	Had	↓↓↓	↓↓↓	↓↓↑	Had	↑↑↑	↓↓↑	↑↑↑
IPSL	↓↓↑	↓↓↑	↑↑↑	IPSL	↓↓↑	↓↓↑	↑↑↑	IPSL	↓↓↑	↓↓↑	↑↑↑	IPSL	↓↓↑	↓↓↑	↑↑↑
MIROC	↑↑↑	↑↑↑	↑↑↑	MIROC	↑↑↑	↑↑↑	↑↑↑	MIROC	↑↑↑	↓↑↑	↑↑↑	MIROC	↑↑↑	↑↑↑	↑↑↑
Nor	↑↑↑	↑↑↑	↑↑↑	Nor	↑↑↑	↑↑↑	↑↑↑	Nor	↓↑↑	↑↑↑	↑↑↑	Nor	↓↓↑	↓↓↓	↓↓↑
Rhine	Q90	Q50	Q10	Rhine	Q90	Q50	Q10	Rhine	Q90	Q50	Q10	Rhine	Q90	Q50	Q10
GFDL	↑↑↑	↑↑↑	↑↑↑	GFDL	↓↓↓	↓↓↓	↑↑↑	GFDL	↓↓↓	↑↓↑	↑↑↑	GFDL	↓↓↓	↓↓↓	↓↓↑
Had	↑↓↓	↓↓↓	↓↓↓	Had	↑↓↓	↓↓↓	↓↓↓	Had	↓↓↓	↓↓↓	↓↓↓	Had	↓↓↓	↓↓↓	↓↓↓
IPSL	↑↑↑	↑↑↑	↑↑↑	IPSL	↓↓↓	↓↓↓	↓↓↓	IPSL	↓↓↓	↓↓↓	↓↓↓	IPSL	↓↓↓	↓↓↓	↓↓↑
MIROC	↑↑↑	↑↑↑	↑↑↑	MIROC	↑↑↑	↓↑↓	↑↑↑	MIROC	↓↓↓	↓↑↓	↑↑↑	MIROC	↓↓↓	↓↓↓	↓↓↑
Nor	↑↑↑	↑↑↑	↑↓↓	Nor	↓↓↓	↓↓↓	↓↓↑	Nor	↓↓↓	↓↓↓	↑↓↑	Nor	↓↓↓	↓↓↓	↓↓↑

Figure 9. Direction of trends in Q90, Q50 and Q10 for for the Upper Niger, Upper Yellow and Rhine grouped by basins (three horizontal bands), four RCPs (four vertical bands), driving climate models (horizontal lines), and three hydrological models. First arrows in the small boxes corresponds to HBV, second one to SWIM , and third one to VIC. The arrows with statistically significant trends are thicker, and their direction up or down shows an increase or decrease, correspondingly.

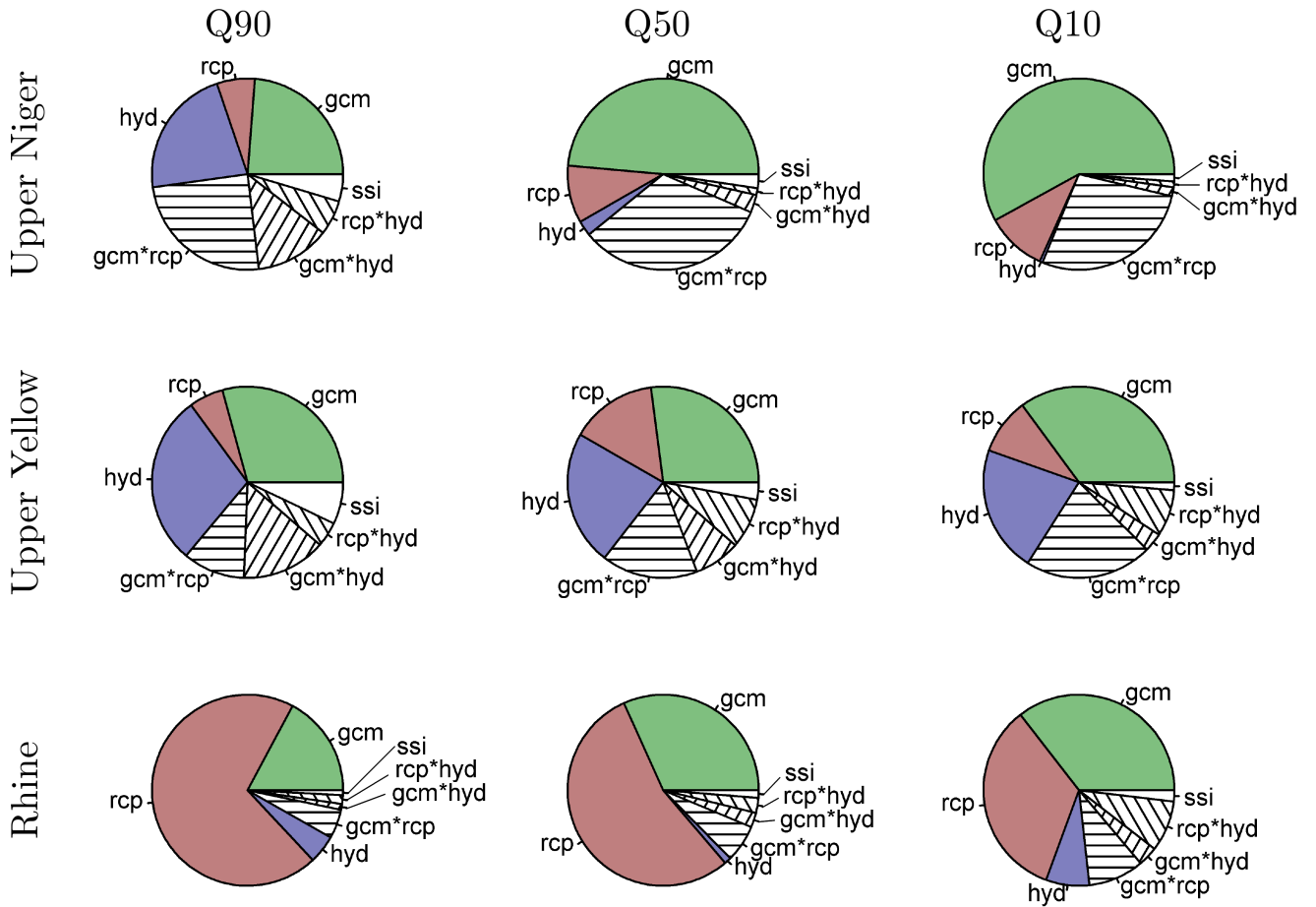


Figure 10. Contribution of different sources of uncertainties to overall uncertainty in the projected slopes of trends for three runoff quantiles Q90, Q50 and Q10 for the Upper Niger, Upper Yellow and Rhine

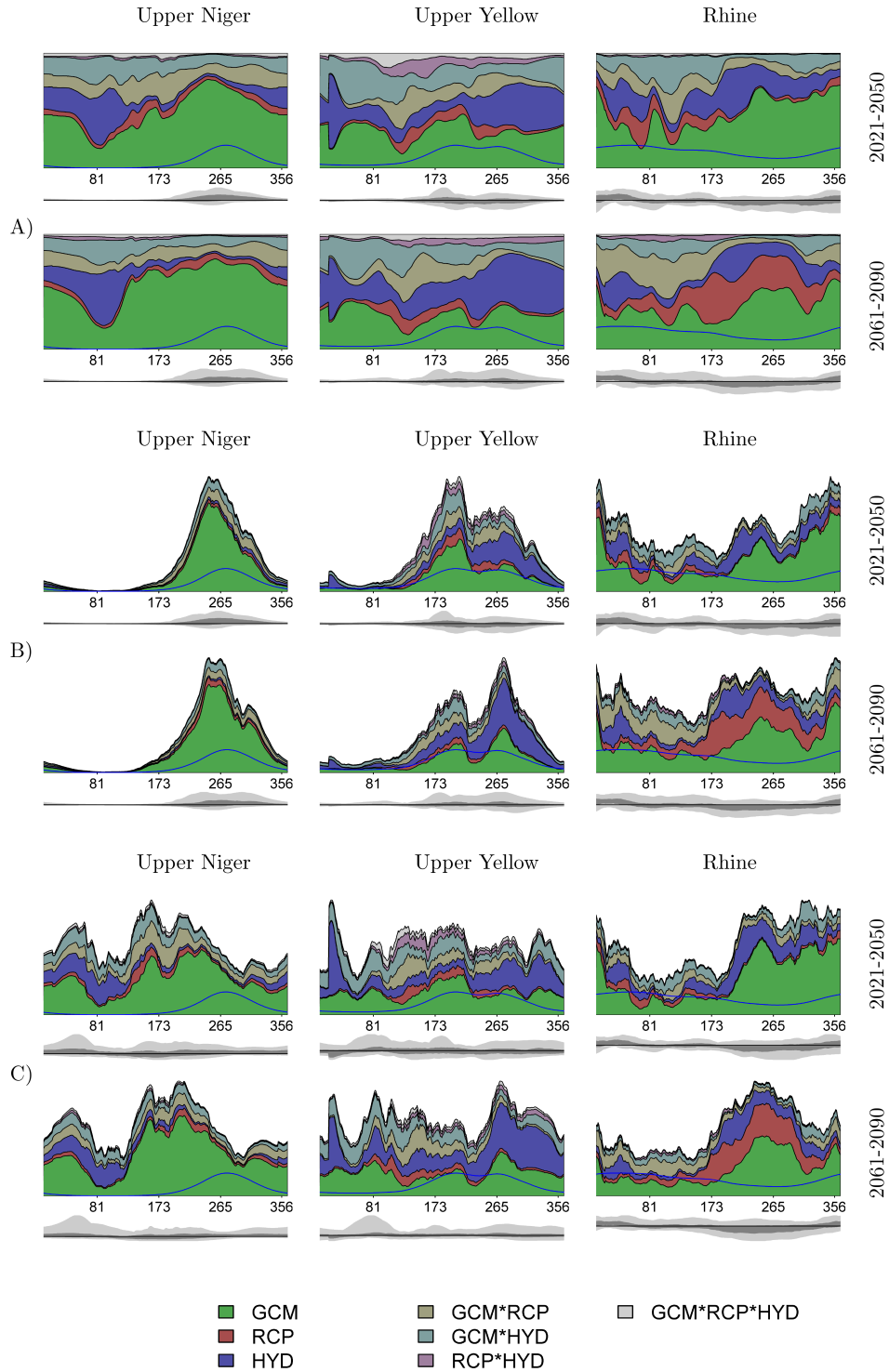


Figure 11. Contribution of different sources of uncertainties to the overall uncertainty in the projected long-term average seasonal dynamics for the Upper Niger, Upper Yellow and Rhine basins for two scenario periods. Blue line denotes the long-term average runoff in the reference period. **Grey lines below the colored graph show the projected absolute changes.** A: unscaled results; B: results scaled by variability (interquartile range) of absolute changes in river discharge; C: results scaled by variability (interquartile range) of relative changes in river discharge. **Ranges (light gray) and interquartile ranges (dark gray) of projected absolute changes (A,B) and projected relative changes (C) are given below each colored graph.**

Appendix A

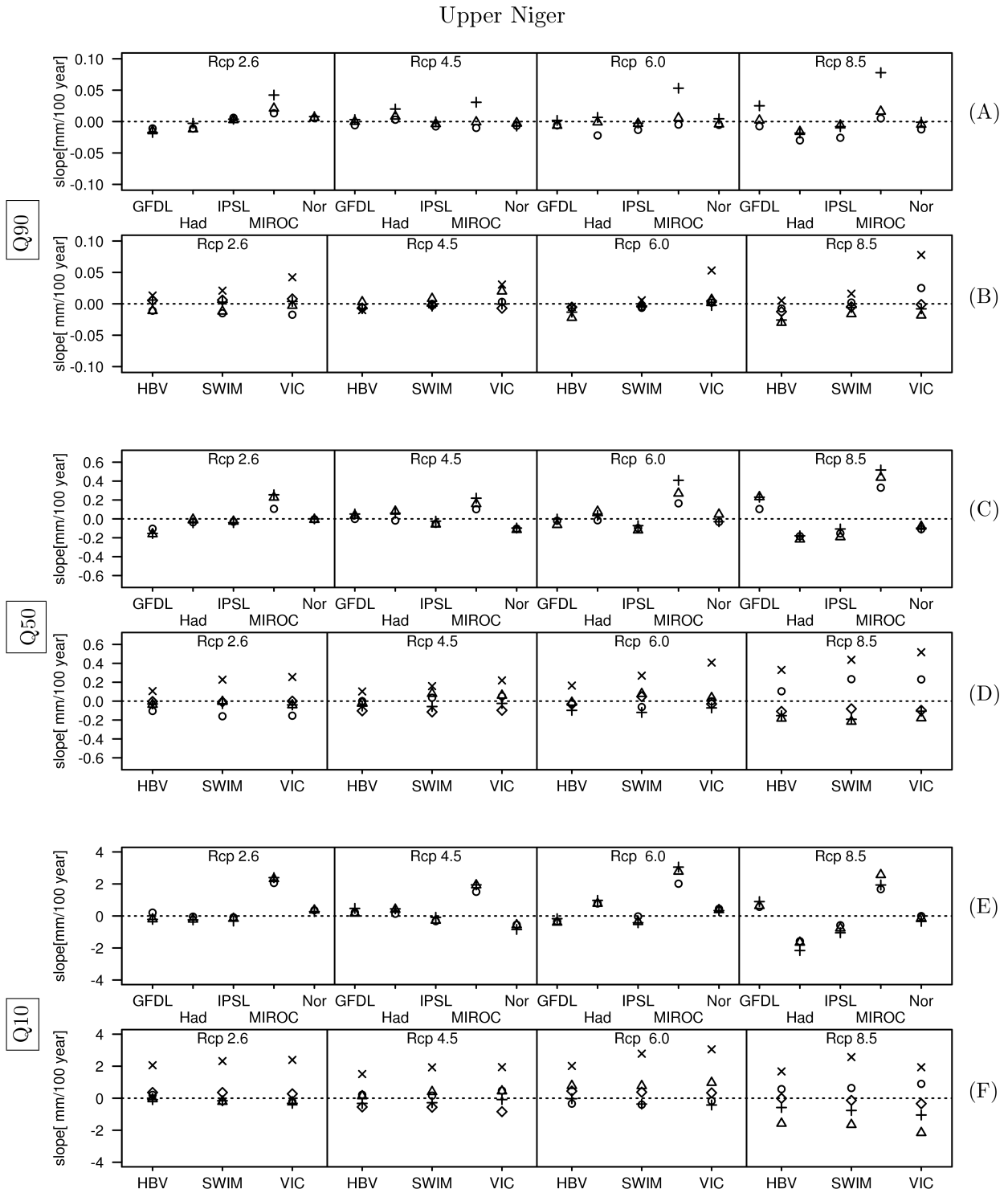


Figure A1. Slopes of trends in (A, B) low flow percentile Q90, (C, D) medium discharge Q50 and (E, F) high flow percentile Q10 grouped by climate models (A, C, E) and by hydrological models (B, D, F) for the Upper Niger.

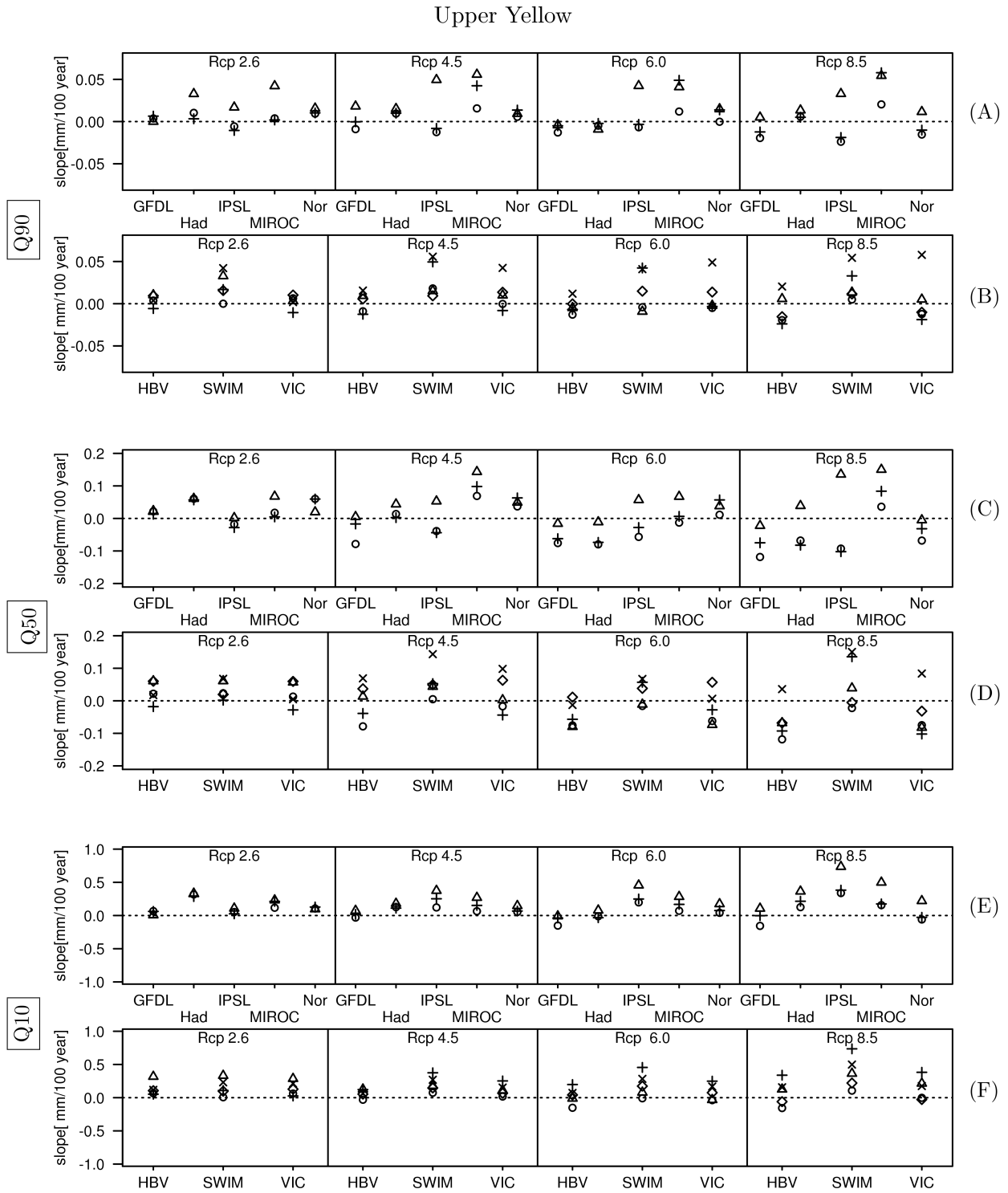


Figure A2. Slopes of trends in (A, B) low flow percentile Q90, (C, D) medium discharge Q50 and (E,F) high flow percentile Q10 grouped by climate models (A, C, E) and by hydrological models (B, D, F) for the Upper Yellow.

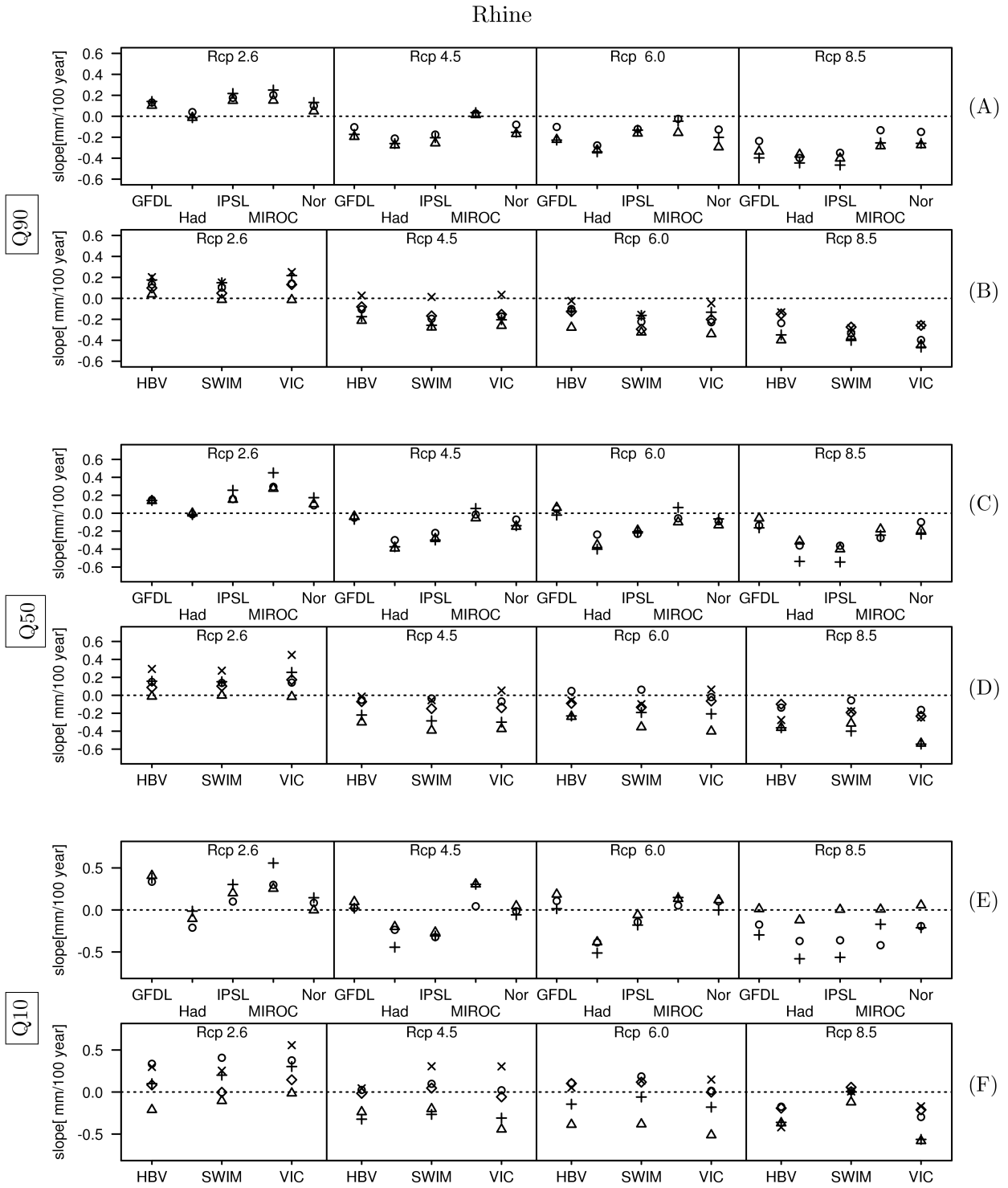


Figure A3. Slopes of trends in (A, B) low flow percentile Q90, (C, D) medium discharge Q50 and (E, F) high flow percentile Q10 grouped by climate models (A, C, E) and by hydrological models (B, D, F) for the Rhine.