



POTSDAM-INSTITUT FÜR  
KLIMAFOLGENFORSCHUNG

**Originally published as:**

**Vetter, T., Huang, S., Aich, V., Yang, T., Wang, X., Krysanova, V., Hattermann, F. (2015):** Multi-model climate impact assessment and intercomparison for three large-scale river basins on three continents. - *Earth System Dynamics*, 6, 1, 17-43

DOI: [10.5194/esd-6-17-2015](https://doi.org/10.5194/esd-6-17-2015)



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

T. Vetter<sup>1</sup>, S. Huang<sup>1</sup>, V. Aich<sup>1</sup>, T. Yang<sup>2</sup>, X. Wang<sup>2</sup>, V. Krysanova<sup>1</sup>, and F. Hattermann<sup>1</sup>

<sup>1</sup>Potsdam Institute for Climate Impact Research, Potsdam, Germany

<sup>2</sup>Hohai University, Nanjing, China

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

Received: 4 June 2014 – Published in Earth Syst. Dynam. Discuss.: 4 July 2014

Revised: 11 December 2014 – Accepted: 13 December 2014 – Published: 22 January 2015

**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 has been performed for many river basins worldwide using different climate scenarios and models. However, their results are hardly comparable, and do not allow one 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. Only a few such studies have been available until now with the global-scale hydrological models, and our study is intended as a step in this direction by applying the regional-scale models. The impact assessment presented here was performed for three river basins on three continents: the Rhine in Europe, the Upper Niger in Africa and the Upper Yellow in Asia. For that, climate scenarios from five general circulation models (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 river discharge and to evaluate uncertainties from different sources. The results allow one to draw some robust conclusions, but uncertainties are large and shared differently between sources in the studied basins. Robust results in terms of trend direction and slope and changes in seasonal dynamics could be found for the Rhine basin regardless of 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 ensured 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 cli-

mate change still remains fragmentary. Very many studies have been undertaken to investigate climate impacts for a number of sectors, globally and on 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 for informing the global policy makers, especially regarding mitigation efforts. However, climate impacts occur and adaptation policies are implemented on the regional scale, where

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

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 fast-track modeling 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 on the global and continental scales (IPCC, 2000, 2007; Déqué et al., 2007; Jacob et al., 2013), fewer studies can be found that 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 that can be found use a variety of methods and techniques to assess the contribution 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 and Knudsen (1996) using three types of hydrological models (lumped, physically based, semi-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).

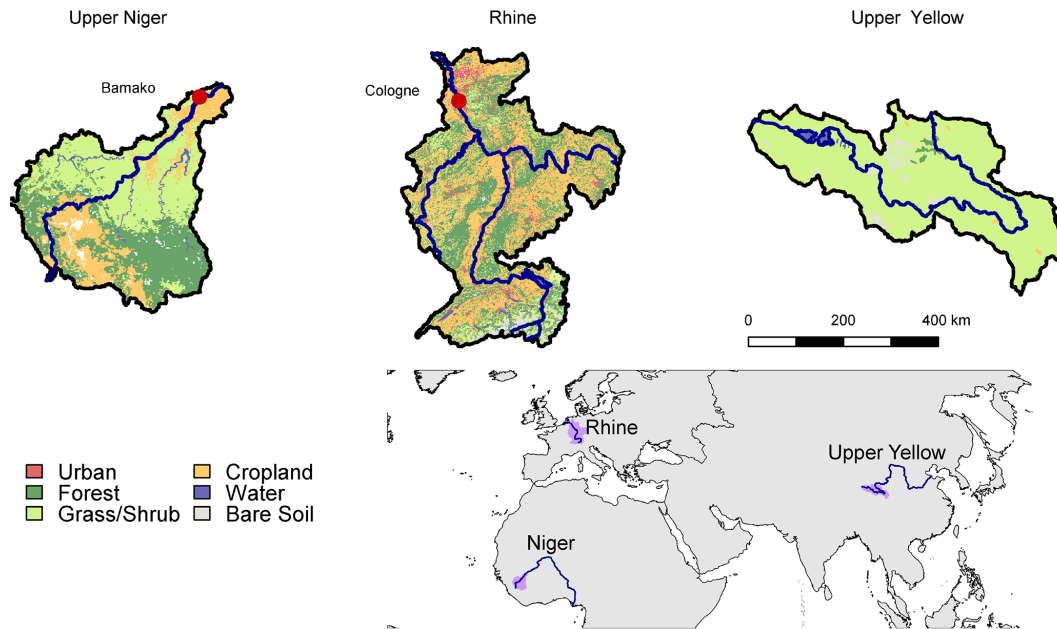
A set of 10 lumped, semi-distributed and fully distributed hydrological models has been applied in the LUCHEM project aiming at investigating the envelope of predictions of changes in hydrological fluxes due to land-use change, also considering land-use change scenarios (Breuer et al., 2009).

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 (fu-

ture greenhouse gas emissions, 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 and Harris (2006).

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 by 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 resource scarcity, suggesting a high potential for improved water resource projections through further improvement of hydrological models. 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 with 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.

Recently, a number of studies were published suggesting the use of the analysis of variance (ANOVA) approach for uncertainty assessment in hydrology and climatology. For example, Yip et al. (2011) use ANOVA to quantify the 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 resource projections and hydro-power production in a glaciated catchment in Switzerland during the twenty-first 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 the ANOVA approach, Bosshard et al. (2013) assessed uncertainties induced by climate models, two different bias correction methods and hydrological models using the output of eight RCMs that are fed into two hydrological models of the Upper Rhine. The results indicate that none of the investigated uncertainty sources is negligible, and some of the uncertainties are not attributable to individual modeling chain components, but rather depend upon their interac-



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

tions, 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 on 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 and Singh, 1995) and VIC (Liang et al., 1994), after their calibration and validation. 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 cover a range of emissions and land-use change projections. The objectives were (a) to compare climate impacts on seasonal water discharge, (b) to compare future trends considering three runoff quantiles,  $Q_{90}$ ,  $Q_{50}$  and  $Q_{10}$ , in terms of trend direction and slope, and (c) to evaluate uncertainties from different sources, namely related to climate models (CMs) providing scenarios, hydrological models (HMs) and RCPs.

## 2 Study areas, data and methods

### 2.1 Study areas: three river basins

From the general set of possible basins, we selected three basins on three continents, which belong to different climatic zones. Besides, the three chosen basins belong to the ISI-MIP Phase 2 (Warszawski et al., 2014) river basins chosen

for intercomparison of climate change impacts on water using regional- and global-scale models.

#### 2.1.1 Upper Niger

The Niger River basin is the largest basin of western Africa. Its source is located in the Guinean highlands, from which the river flows in the northeasterly 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 Koulikoro gauge station (Fig. 1) covering an area of about 122 000 km<sup>2</sup> was taken for impact assessment. The complete Niger River basin could not be properly modeled with HBV and VIC, since there is no inundation module included in those models, and the flow of the Niger is very much altered by the existence of the inner Niger Delta and related inundation processes. 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 of 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 that feeds the river comes mainly from the Guinean Highlands during the rainy season. The average



**Table 1.** Characteristics of the three river basins.

	Rhine (Rees)	Upper Yellow (Tanghaihai)	Upper Niger (Koulikoro)
Area (km <sup>2</sup> )	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

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) in its lower reaches is heavily influenced by human water use like irrigation, and abstraction for industry and municipal purposes. Hydrological modeling 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. The Yellow River source region above the Tangnaihahi gauging station belongs to the Qing–Tibetan Plateau of China. This Upper Yellow basin with the drainage area of about 122 000 km<sup>2</sup> covering approximately 15 % of the entire Yellow River’s drainage basin, while supplying 38 % of the river’s total runoff (Chen et al., 2013), was chosen for our study. 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 of this area. The two major lakes along this part are Lake Zhaling and Lake Eling, with capacities of 4.7 and 10.8 billion m<sup>3</sup>, respectively. The mean altitude of the drainage area is about 4000 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 River basin are not much influenced by human activities, besides some overgrazing and wild herb digging.

### 2.1.3 The Rhine

The Rhine River basin covers a drainage area of about 185 000 km<sup>2</sup> 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 French–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 is located in Germany. The altitude in the drainage area ranges from 4275 m a.s.l. in the Swiss Alps to 0 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 (Table 1), and seasonality is not very distinct. In the Alpine region, the maximum discharge of the Rhine is observed during summer due to snowmelt. 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 the 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 Rees gauge (160 000 km<sup>2</sup>).

## 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 and Belward, 2005). Observed river discharge data from the Global Runoff Data Centre were used to calibrate and validate the hydrological models (GRDC, 1998). As climate input for model calibration, the WATCH reanalysis data were applied with a grid resolution of 0.5° (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° 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 RCP2.6 to RCP8.5, were included in the study.

## 2.3 Hydrological models

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

### 2.3.1 HBV

The HBV model (Bergström and Forsman, 1973; Bergström and Singh, 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 was based on the Nordic HBV version (Saelthun, 1996).

HBV has been applied for modeling hydrological processes in countries with different climatic conditions, as, for example, Sweden, China (Zhang and 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 and Wang, 2009).

### 2.3.2 SWIM

The SWIM (Soil and Water Integrated Model, Krysanova et al., 1998) ecohydrological model 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 on 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 of providing 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, 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. 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. (2015).

### 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, or 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 modeler can subdivide each grid cell into an arbitrary number of tiles, each corresponding to the fraction of the cell covered by a particular land cover (i.e., 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 applied extensively in climate impact studies for a number of large river basins over the continental US and the globe (Hamlet and Lettenmaier, 1999; Su and Xie, 2003; Christensen and Lettenmaier, 2007).

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

Features	HBV	SWIM	VIC
Spatial disaggregation	Subbasins, 10 elevation zones and land-use classes within them	Subbasins and hydrotopes (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	Two parameters: $T_{\text{mean}}^b$ , precipitation	Six parameters: $T_{\text{min}}^a$ , $T_{\text{mean}}^b$ , $T_{\text{max}}^c$ , precipitation, air humidity, radiation	Five parameters: $T_{\text{min}}^a$ , $T_{\text{max}}^c$ , precipitation, air humidity, wind speed
Representation of soils	One soil layer, two 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 the EPIC approach	Fixed monthly plant characteristics
Calculation of potential evapotranspiration	Blaney–Criddle	Priestley–Taylor	Penman–Monteith
Method to calculate snowmelt	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 equations

<sup>a</sup> Minimum temperature, <sup>b</sup> mean temperature, <sup>c</sup> maximum temperature.

### 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 one 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 the 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 representations compared to VIC. For the VIC raster-based model, a grid resolution of  $0.125^\circ$  was used for all basins. For example, for the Rhine basin, 1433 grid cells were simulated by VIC. For the SWIM and HBV applications, the basins were subdivided into subbasins with an average area of 100–200 km<sup>2</sup> using the 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 average areas of 5.9, 3.8, and 2.3 km<sup>2</sup>, respectively.

All hydrological models were calibrated using the observed discharge at the basin outlet. For SWIM and VIC, automatic calibration was performed with the PEST software package (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, the Nash and Sutcliffe efficiency (Nash and 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  $\text{NSE} = 1$  and  $\text{logNSE} = 1$ , reflecting the theoretically best possible value for NSE and logNSE.

The number 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.

## 2.4 Method of trend analysis

Trends in projected runoff were calculated for three runoff quantiles, which reflect the annual high-flow conditions ( $Q_{10}$ ), the annual low-flow conditions ( $Q_{90}$ ) and the medium-flow conditions ( $Q_{50}$ ). To avoid the same event being separated between two adjacent years, the hydrological years have been taken instead of the calendar years for calculation of annual  $Q_{50}$  and  $Q_{10}$ . As  $Q_{90}$  may occur at the end or at the beginning of a new hydrological year, we used a different approach for this quantile. To calculate  $Q_{90}$ , 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 and 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 error 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 (1973). The convergent iterative numerical algorithm for the MM estimates was provided by Yohai (1987), and is included in the R statistical software. For the trend analysis in our study, the “lmrob” function from the “robust” package 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

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

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:

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

The total sum of squares is partitioned into three main effects  $SS_{\text{Hyd}}$ ,  $SS_{\text{Gcm}}$ , and  $SS_{\text{Rcp}}$ , corresponding to three hydrological models, five GCMs and four RCPs, respectively. In addition, there are four interaction terms  $SS_{\text{Hyd}*\text{Gcm}}$ ,  $SS_{\text{Hyd}*\text{Rcp}}$ ,  $SS_{\text{Gcm}*\text{Rcp}}$ , and  $SS_{\text{Hyd}*\text{Gcm}*\text{Rcp}}$ , 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 three cases, GCMs show negative trends and, in two cases, GCMs show 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:

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

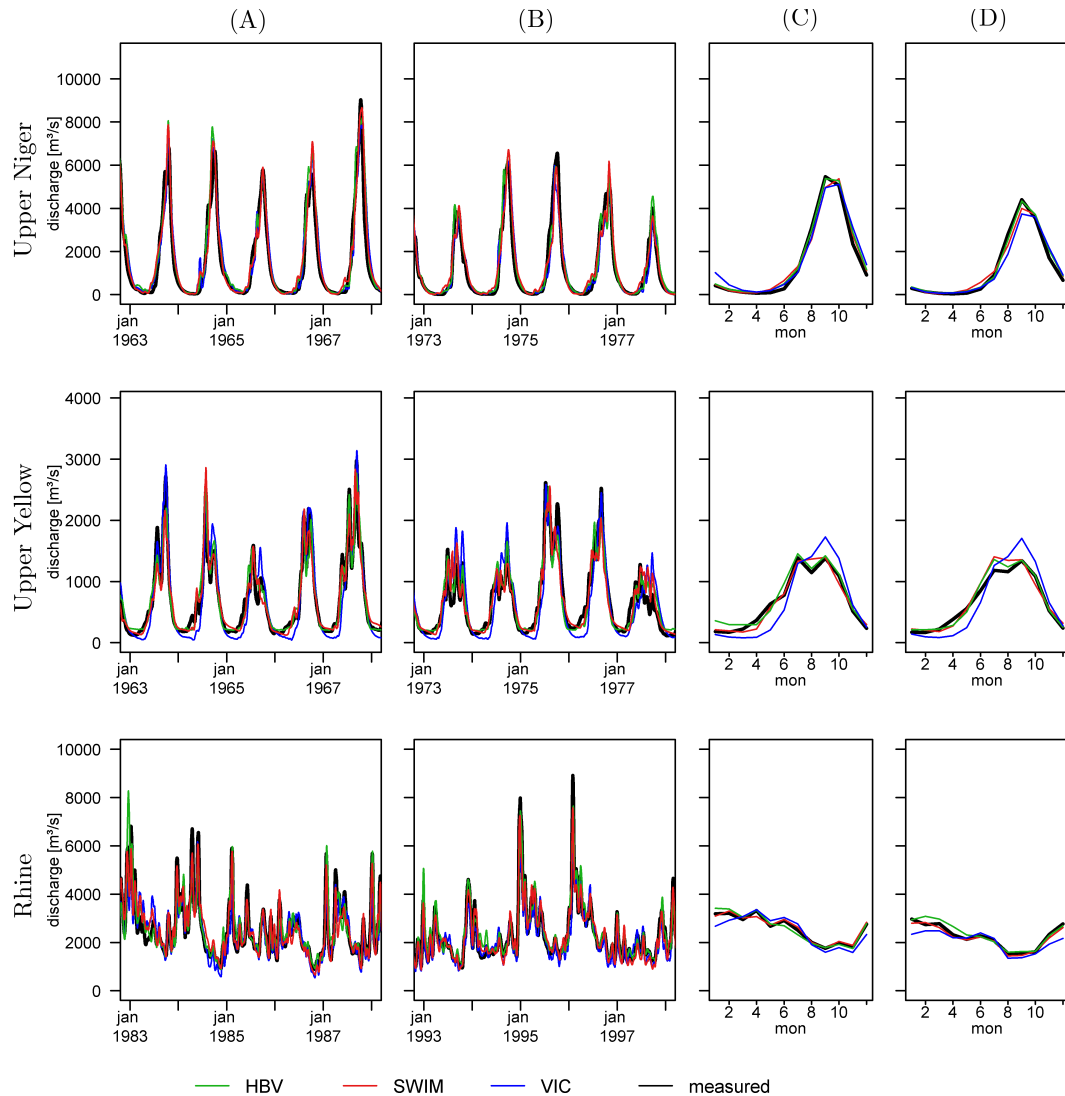
$$SS_{\text{Hyd}*\text{Gcm}} = N_{\text{Rcp}} \sum_{i=1}^{N_{\text{Hyd}}} \sum_{j=1}^{N_{\text{Gcm}}} (\bar{Y}_{ijo} - \bar{Y}_{ioo} - \bar{Y}_{ojo} + \bar{Y}_{ooo})^2 \quad (4)$$

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

where  $N_{\text{Hyd}}$ ,  $N_{\text{Gcm}}$ , and  $N_{\text{Rcp}}$  describe the number of hydrological models, the number of GCMs and the number of RCPs, respectively. The token  $\bar{\phantom{x}}$  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 combinations of three hydrological models, three GCMs and three RCPs are fulfilled; so, for the four RCPs, we have four combinations ( $c_1$ – $c_4$ ) for selecting a subsample with  $N_{\text{Rcp}} = 3$ , namely  $c_1 = 2.6, 4.5,$  and  $6.5$ ,  $c_2 = 2.6, 4.5,$  and  $8.5$ ,  $c_3 = 2.6, 6.5,$  and  $8.5$ , and  $c_4 = 4.5, 6.5,$  and  $8.5$ . For the five GCMs, there are 10 possible combinations. These four possible RCP combinations and 10 possible GCM combinations are finally combined into 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.





**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 5 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.

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

$$\eta_{\text{Hyd}}^2 = \frac{1}{40} \sum_{m=1}^{40} \frac{\text{SS}_{\text{Hyd}}(m)}{\text{SST}(m)}. \quad (6)$$

The calculation of the contributing variance fraction for all the other six components is analogous to Eq. (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 from 0.81 to 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 5 years in both the calibration and validation periods, and a comparison of the long-term aver-

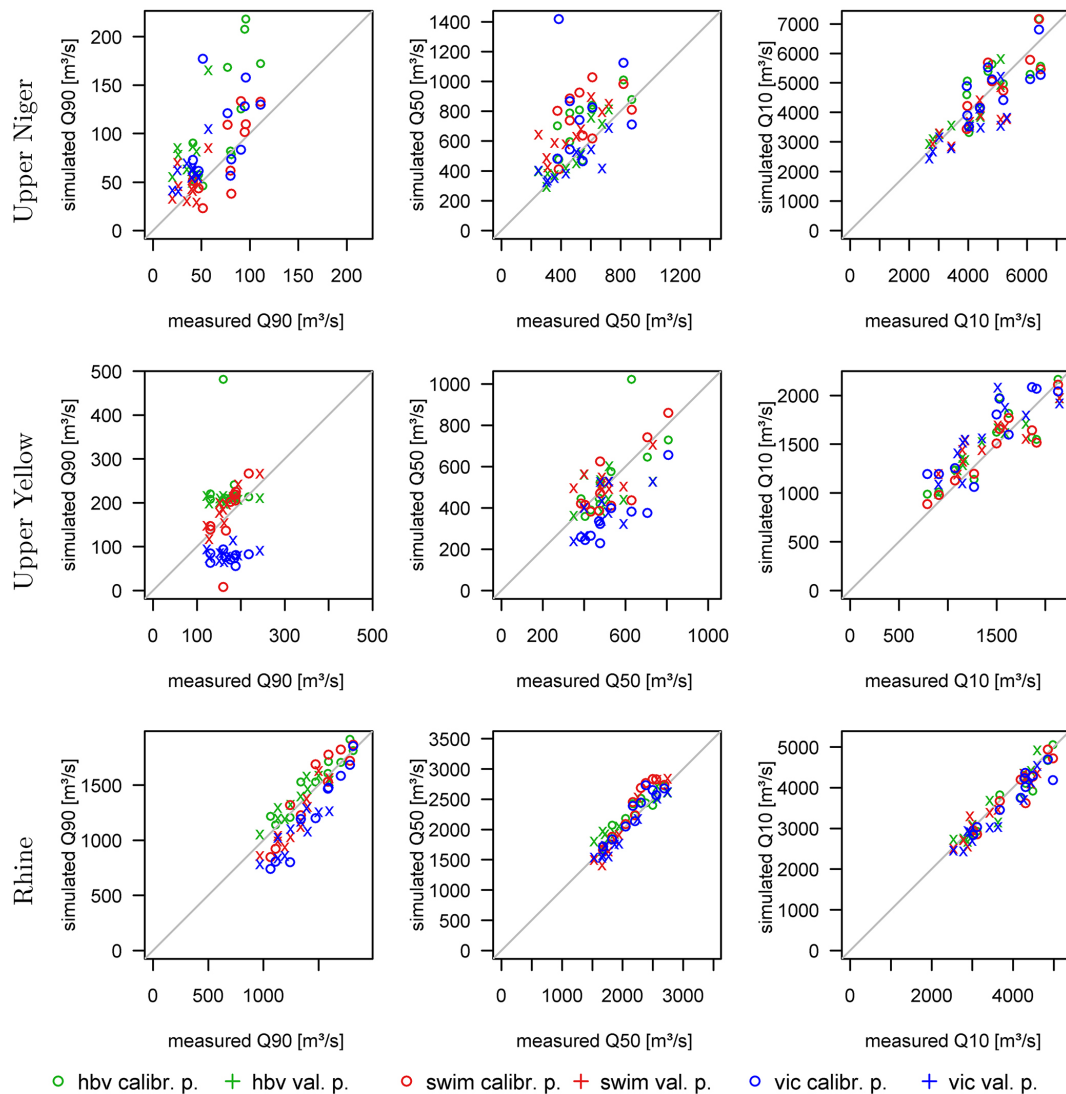
**Table 3.** Calibration and validation results with the daily time step. NSE = Nash and Sutcliffe efficiency; pbias = percent bias.

<i>Rhine</i>	Calibration			Validation		
	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
<i>Upper Yellow</i>	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
<i>Upper Niger</i>	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 ( $Q_{90}$ ,  $Q_{50}$ ,  $Q_{10}$ ) in the calibration and the validation period.

<i>Rhine</i>	Calibration			Validation		
	HBV	SWIM	VIC	HBV	SWIM	VIC
Period	1981–1990	1981–1990	1981–1990	1991–2000	1991–2000	1991–2000
$Q_{90}$ cor	0.96	0.94	0.9	0.93	0.93	0.88
$Q_{90}$ pbias	4.4	0.0	−13.0	5.8	−8.0	−17.2
$Q_{50}$ cor	0.93	0.94	0.94	0.95	0.98	0.98
$Q_{50}$ pbias	3.8	8.0	3.6	6.3	1.1	−4.7
$Q_{10}$ cor	0.95	0.94	0.93	0.96	0.96	0.95
$Q_{10}$ pbias	−3.9	−3.4	−5.8	0.4	−1.9	−7.5
<i>Upper Yellow</i>	HBV	SWIM	VIC	HBV	SWIM	VIC
Period	1961–1970	1961–1970	1961–1970	1971–1980	1971–1980	1971–1980
$Q_{90}$ cor	−0.16	0.66	0.02	0.18	0.91	0.25
$Q_{90}$ pbias (%)	39.6	1.9	−56.2	25.4	14.4	−49.5
$Q_{50}$ cor	0.72	0.82	0.87	0.33	0.70	0.57
$Q_{50}$ pbias (%)	4.2	−3.4	−35.0	−3.5	5.7	−20.1
$Q_{10}$ cor	0.84	0.93	0.88	0.97	0.89	0.79
$Q_{10}$ pbias (%)	3.1	−1.5	11.3	6.0	8.4	14.6
<i>Upper Niger</i>	HBV	SWIM	VIC	HBV	SWIM	VIC
Period	1961–1970	1961–1970	1961–1970	1971–1980	1971–1980	1971–1980
$Q_{90}$ cor	0.74	0.79	0.26	0.52	0.34	0.72
$Q_{90}$ pbias	61.6	4.2	37.6	108.4	27.0	63.2
$Q_{50}$ cor	0.69	0.40	0.08	0.92	0.83	0.78
$Q_{50}$ pbias	30.1	38.1	41.1	9.6	40.5	−4.0
$Q_{10}$ cor	0.71	0.84	0.71	0.78	0.81	0.79
$Q_{10}$ pbias	2.1	−1.7	−2.7	−3.8	−10.9	13.6





**Figure 3.** Comparison of simulated and observed annual values of  $Q_{90}$ ,  $Q_{50}$  and  $Q_{10}$  in the calibration and validation periods for three basins: the Upper Niger, Upper Yellow and Rhine.

age 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  $Q_{90}$ ,  $Q_{50}$  and  $Q_{10}$  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 ( $Q_{10}$ ) in all three basins. The coefficient of correlation ranges from 0.78 to 0.97 for the validation period, and the bias is

between  $-13.6$  and  $+14.6$  %, and in seven cases of nine 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  $Q_{50}$  and  $Q_{90}$  values. The lowest coefficient of correlation for the two quantiles is 0.88 for  $Q_{90}$  simulated by VIC. The highest bias for  $Q_{90}$  in the validation period is  $-17.2$  % (VIC). For  $Q_{50}$  the highest bias is  $+6.3$  % for the simulation run with HBV.

The simulation results for  $Q_{90}$  and  $Q_{50}$  in the Upper Yellow and the Upper Niger are not as good as for the Rhine basin. For  $Q_{50}$  in the Upper Niger, the coefficient of correlation ranges from 0.78 to 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 from 0.33 to 0.7).

For the low-flow simulations ( $Q_{90}$ ), the biases in the validation period are even higher than for  $Q_{50}$ . 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 from 0.18 to 0.91 in the Upper Yellow and from 0.52 to 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  $Q_{10}$  in all three basins, and for  $Q_{50}$  and  $Q_{90}$  in the Rhine. However, the results for  $Q_{50}$  and  $Q_{90}$  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  $Q_{10}$  is nearly 100 times higher than  $Q_{90}$ . Therefore, low flow gets small weights in the calibration process, leading to bias in simulations.

In total, there is no hydrological model that 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.

Figure 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 the 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<sup>-1</sup> 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 RCP2.6 and the smallest differences for RCP8.5, 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<sup>-1</sup>. Only in the second period do 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  $\pm 20$  mm month<sup>-1</sup>; only the Had scenarios show a stronger decrease in some months, and MIROC a higher increase in summer months, contradicting other models. For the Rhine, all GCMs project an increase in precipitation for RCP2.6. Except for Nor, all the other GCMs show a 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 Fig. 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

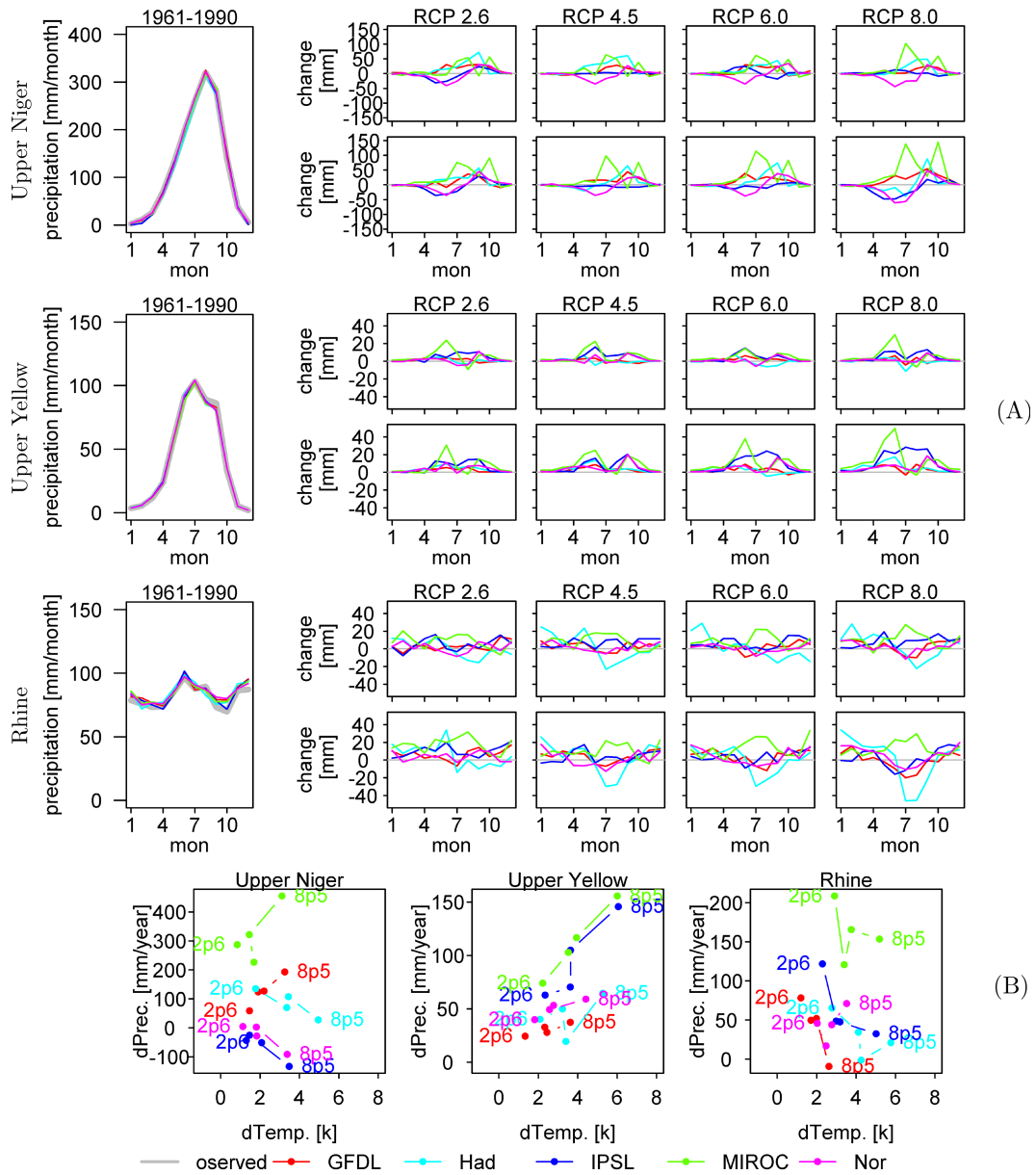
After the calibration and validation of the three hydrological models, they were run for the period 1971–2099 using five GCM scenarios for four RCPs providing 60 time series, which were analyzed for long-term average seasonal dynamics and trends.

Figure 6 presents the long-term average seasonal discharge for the reference period 1961–1990, for the second scenario period 2061–2090 and the difference between the second scenario and the reference periods.

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 color are close to one another. The increase in discharge projected by simulations driven by MIROC is highest in the rainy season from July to December (and increases from RCP2.6 to RCP8.5), followed by those driven by GFDL, which also show an increase. On the opposite end, 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 the second half of the year (months 7–11), whereas the results driven by three other CMs show rather moderate changes. The magnitude of the changes clearly and steadily increases from RCP2.6 to RCP8.5.

For the Rhine, a decrease in the 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



**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 four RCPs: 2.6, 4.5, 6.0, and 8.5. **(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.

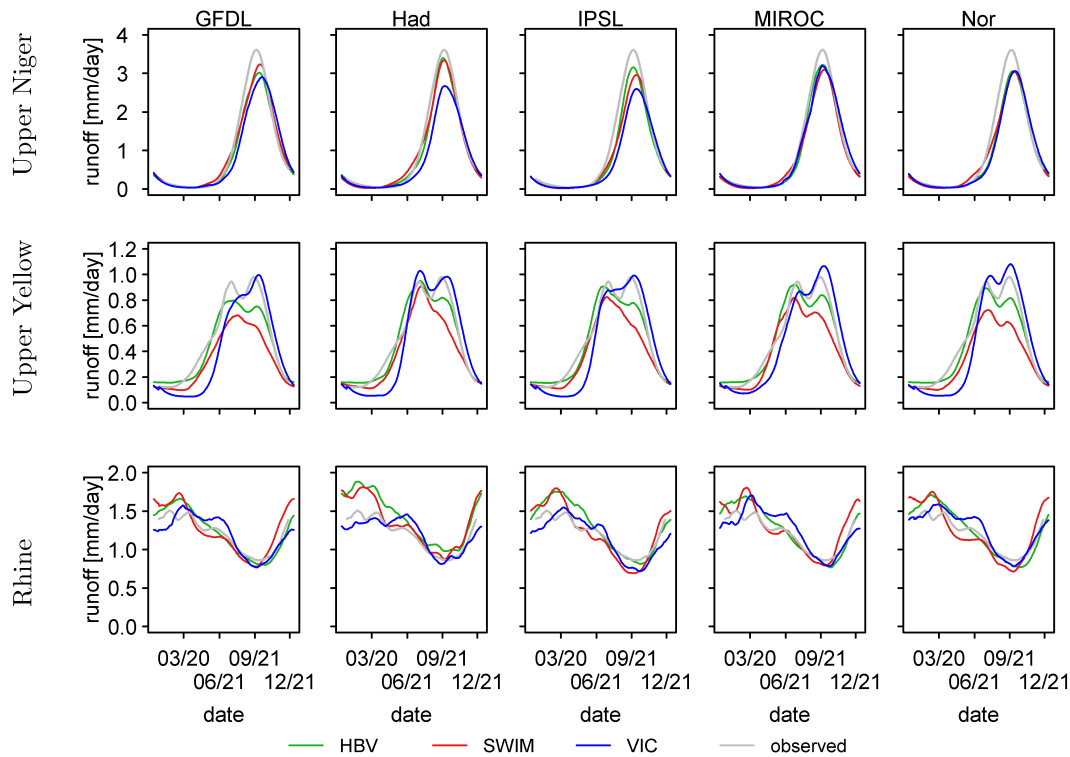
for all months and RCPs. Uncertainty related to CM is visually higher compared to that related to HM (clustering of curves with the same color). The projections by SWIM and VIC agree very well (though are not shown in Fig. 6).

The results driven by the MIROC climate model 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 the 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

After that, linear trends were calculated using a robust statistical method for three variables: annual median runoff  $Q_{50}$ ,



**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.

and low and high annual percentiles  $Q_{90}$  and  $Q_{10}$  representing the low- and high-flow conditions, respectively. The trends were calculated for the period 2010–2099, and the significance of the trends was evaluated at the 5 % level. The results of trend analysis for the three basins in terms of the slopes of trend and trend direction (and significance) are presented in Figs. 7–9 and Figs. A1–A3.

### 3.4.1 Slope of trend

Two examples of trends for the median flow  $Q_{50}$  simulated by the three hydrological models driven by the Had climate scenario and RCP8.5 are presented in Fig. 7 for the Upper Niger and Upper Yellow. 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.

Figures A1–A3 show the slopes of trends in  $Q_{90}$ ,  $Q_{50}$  and  $Q_{10}$ . 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.

For the Rhine, slopes of the trends in the two variables  $Q_{90}$  and  $Q_{50}$  are mainly positive under RCP2.6. The trends are negative under three other RCPs, almost steadily decreasing from RCP2.6 to RCP8.5. The slopes of  $Q_{10}$  are positive

under RCP2.6, negative under RCP8.5, and uncertain under RCP4.5 and RCP6.0.

The results in terms of slopes 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 RCP2.6 and  $Q_{50}$  do more than 75 % of the model runs agree on declining trends.

For the Upper Yellow, all three variables,  $Q_{90}$ ,  $Q_{50}$ , and  $Q_{10}$ , show an increase under RCP2.6 and RCP4.5, and the slopes of the trend in  $Q_{10}$  representing high flows are positive for all four RCPs.

The graph Figs. A1–A3 allow one 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

Figure 9 summarizes the results of the evaluation of the direction of trends for all 180 time series. The direction of trends in  $Q_{90}$ ,  $Q_{50}$  and  $Q_{10}$  are grouped by basins, four RCPs, driving climate models, and three hydrological models. Only statistically significant trends are analyzed below.

For the Niger, much more significant trends were found for RCP8.5 than for the other three. In this case, all sim-

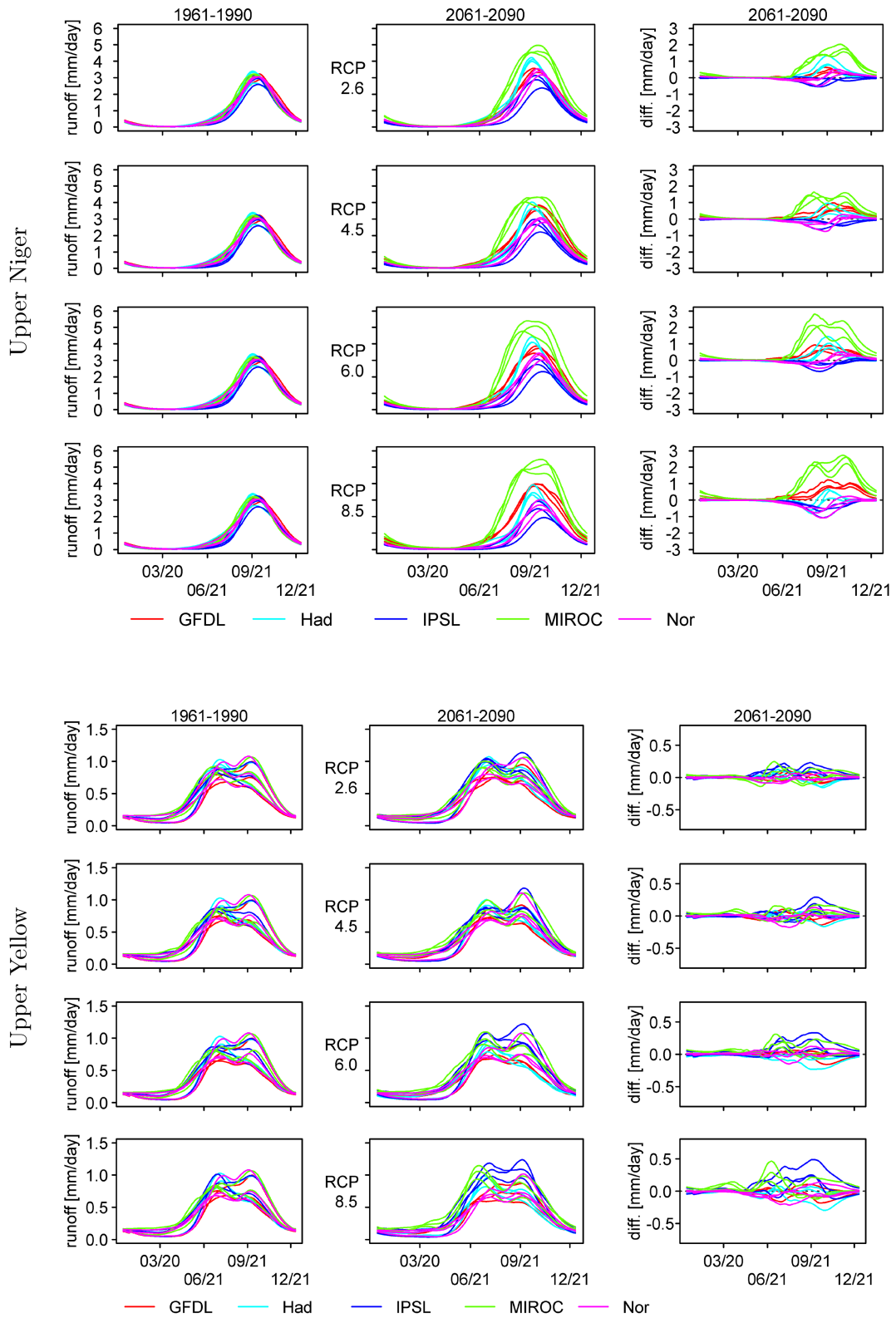
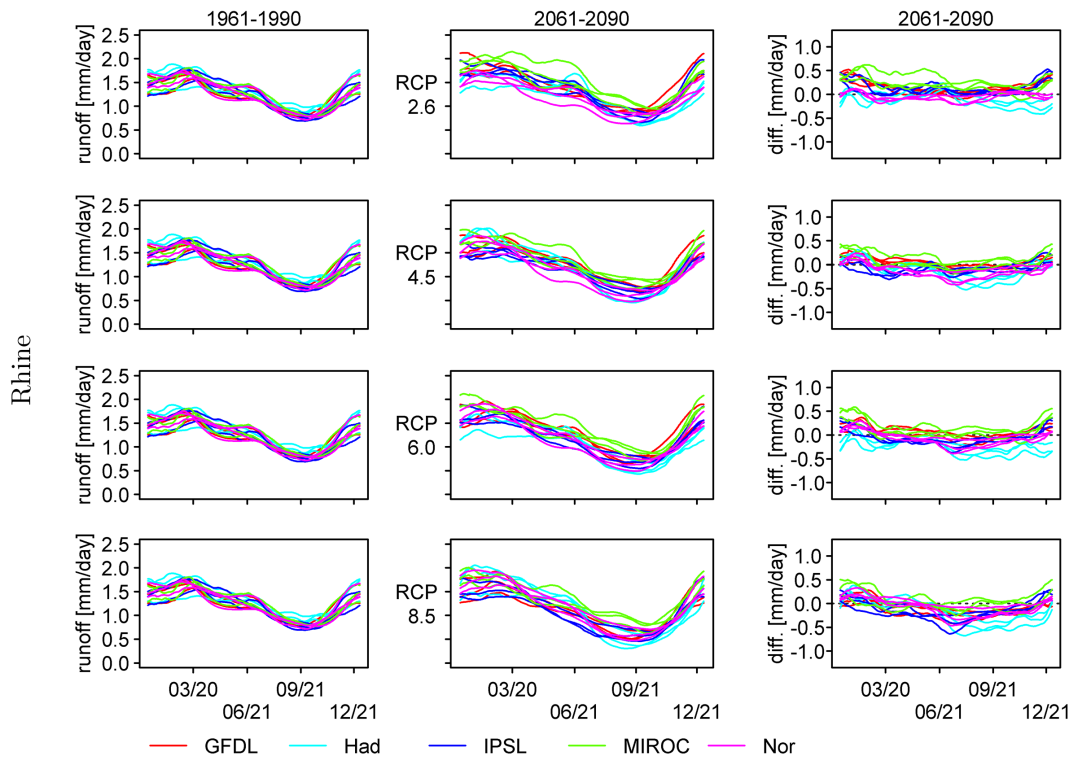
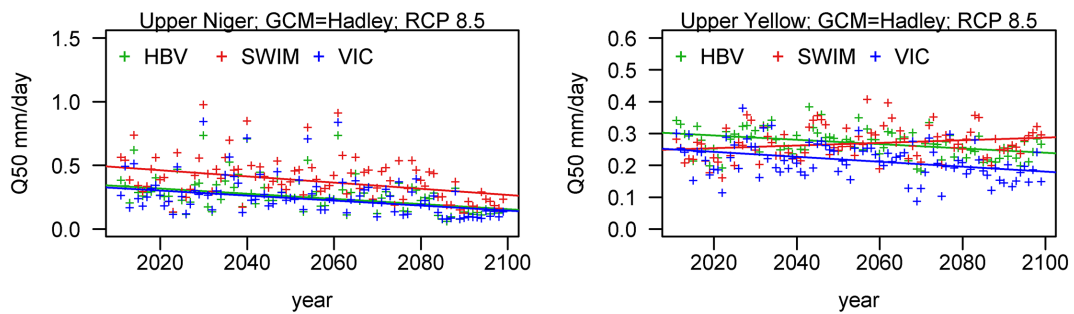


Figure 6.





**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.



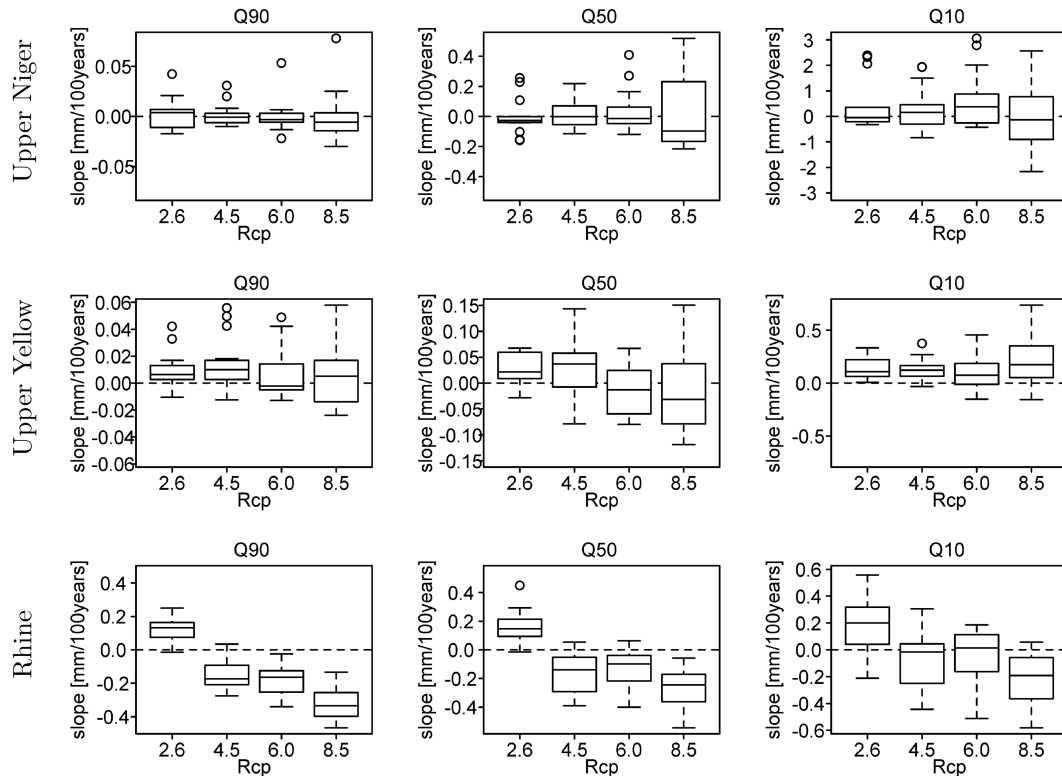
**Figure 7.** Two examples showing annual median flows  $Q_{50}$  estimated from simulations by three models driven by the Had climate scenario data corresponding to RCP8.5 and the corresponding trend lines (all presented trends are significant) for the Upper Niger (left panel) and Upper Yellow (right panel) in the period 2010–2099.

ulations (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  $Q_{50}$ , the Nor-driven projections corresponding to RCP8.5 agree with those of Had and IPSL, and two of three GFDL projections agree with those of MIROC. The direction of trends in  $Q_{50}$  for three other RCPs is positive in all simulations driven by MIROC, and in each of these three cases there is a simulation driven by one of the other CMs that shows a downward trend simulated by all three HMs: GFDL for RCP2.6, Nor for RCP4.5, and IPSL for RCP6.0. The results driven by MIROC

show the upward trends in  $Q_{50}$  and  $Q_{10}$  for all RCPs. In total, a high discrepancy between climate models is obvious for this basin.

For the Upper Yellow, a downward trend in  $Q_{90}$  and  $Q_{50}$  simulated by HBV and SWIM driven by three to four climate models was found for RCP8.5, 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  $Q_{10}$  show a moderate agreement in increasing trends, with a growing level of significance from RCP2.6 to RCP8.5. Only





**Figure 8.** Boxplots for the slopes of the linear trend for the 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).

for the Yellow could several cases be found, where hydrological models disagree and show significant opposite trends.

Evaluating results for the Rhine, we can conclude with a moderate certainty that all three runoff quantiles show positive trends for RCP2.6. The direction of change is different for three other RCPs. A significant downward trend in  $Q_{50}$  and  $Q_{90}$  was found in simulations driven by three to four climate models for RCP4.5 and RCP6.0.  $Q_{90}$  and  $Q_{50}$  show a stronger downward trend in nearly all simulations for RCP8.5. Regarding  $Q_{10}$ , there is only a small number of significant trends, and the trend direction changes from a positive trend in the RCP2.6 scenario to a negative one in the RCP8.5 scenario. In general, a good agreement between CM-driven simulations and HM outputs can be stated for the Rhine.

### 3.5 Evaluation of uncertainty

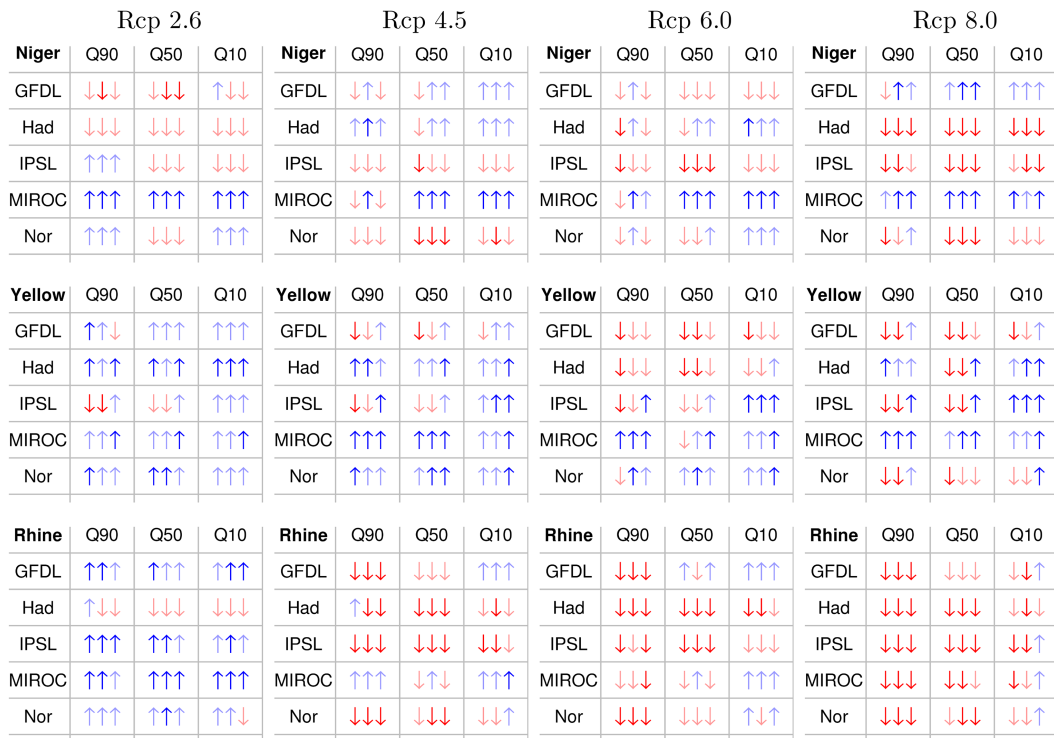
The sources of uncertainties were analyzed using the ANOVA method described in Sect. 2.5 for the slopes of the linear trend and changes in long-term average seasonal dynamics (Figs. 10 and 11). The variance decomposition of the projected slopes for three runoff quantiles is presented in Fig. 10.

For the projected high-flow trends ( $Q_{10}$ ) 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 about 95 % to the total uncertainty. For  $Q_{50}$ , 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 in projected trends is more pronounced than in the Upper Niger. For  $Q_{90}$  and  $Q_{50}$ , their effect is comparable to the effect of the CMs. Together, they explain 58 % of the variance for  $Q_{10}$  and 49 % of the total variance for  $Q_{10}$ , respectively. For the high flows, the CMs are the dominant 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  $Q_{50}$  and  $Q_{10}$ , the CMs become more important contributors of uncertainty, but are still lower than RCPs. For  $Q_{90}$ , CMs and RCPs together explain about 70 % and, for  $Q_{50}$ , 85 % of the overall uncertainty. In the Rhine basin, there is not much contribution of the hydrological models in the uncertainty of the projected trends.



**Figure 9.** Direction of trends in  $Q_{90}$ ,  $Q_{50}$  and  $Q_{10}$  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. The first arrow in the small boxes corresponds to HBV, the second one to SWIM, and the third one to VIC. The arrows with statistically significant trends are thicker, and their direction up or down shows an increase or decrease, respectively.

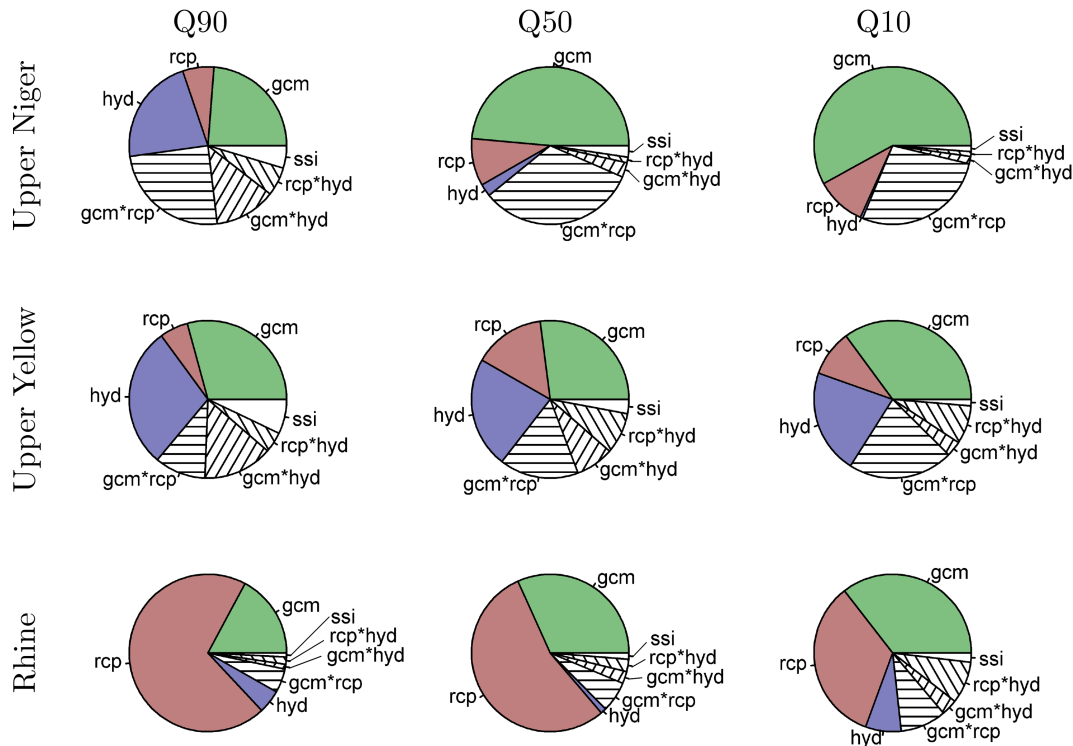
The sources of uncertainty related to the changes in the long-term average seasonal dynamics of runoff are shown in Fig. 11a–c. Figure 11a shows the unscaled results for three basins in two selected scenario periods. To highlight the periods of the year where the changes among the different runoff projections are large (high uncertainty), the graphs presented in Fig. 11a were scaled further. The scaling was performed in accordance with the variability in absolute changes (Fig. 11b), and also in accordance with the variability in relative changes (Fig. 11c). As variance estimators, the interquartile range of absolute changes (Fig. 11b) and the interquartile range of relative changes (Fig. 11c) were taken for the scaling. When the differences between low-flow and high-flow river runoff are very large, as is the case in the Upper Niger as well as in the Upper Yellow, the scaling of the variance contributions in absolute terms only 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 4 to 5 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 these periods, the importance of hydrological models increases. 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 the influence of hydrological models becomes higher in the second scenario period.

In contrast to these two basins, for the Rhine, the RCPs also 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,



**Figure 10.** Contribution of different sources of uncertainties to overall uncertainty in the projected slopes of trends for the three runoff quantiles  $Q_{90}$ ,  $Q_{50}$  and  $Q_{10}$  for the Upper Niger, Upper Yellow and Rhine.

a clear difference between the first and 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 Figs. 10 and 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 do the main sources of uncertainties differ; that is, 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.

### 4.1 Evaluation and validation of models

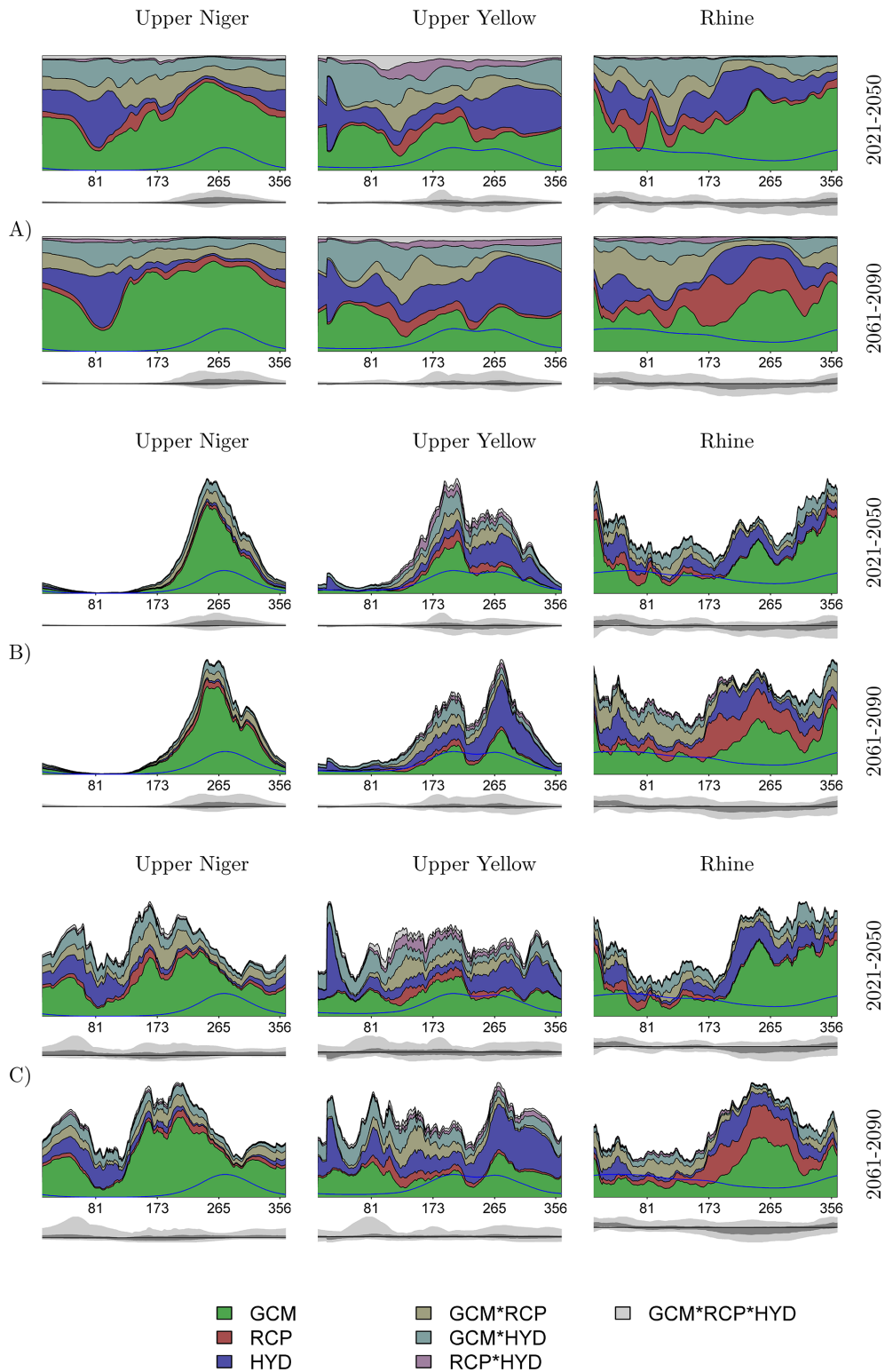
The validation of the hydrological models provided good results for river discharge and high flows ( $Q_{10}$ ) in all three

basins, and for the median flow ( $Q_{50}$ ) and low flow ( $Q_{90}$ ) in the Rhine. However, the validation results for  $Q_{50}$  and  $Q_{90}$  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 good agreement.

### 4.2 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 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 of 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 the future are difficult to draw. 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



**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. The blue line denotes the long-term average runoff in the reference period. **(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.

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 ensured by all three hydrological models. In this basin, also having a monsoonal type of climate but lower temperatures than in the Niger basin, the snowmelt 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.

#### 4.3 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 twenty-first century (Yip et al., 2011). Regarding precipitation projections, Hawkins and 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 do 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.

#### 4.4 Uncertainty related to GCMs

The dominance of GCM-related uncertainty in impact studies is also reported 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 is usually done for hydrological models (see Blöschl and 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 on the global scale with only greenhouse gas emissions as a 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 precondition for obtaining more robust hydrological impacts.

#### 4.5 Uncertainty related to hydrological models

It is likely that uncertainty in hydrological models increases with the increase in 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 snowmelt 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 prevails.

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 modeled 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 snowmelt processes, also simulated by the hydrological models, become relevant.

It is important, when discussing the contribution of hydrological models to the entire scenario uncertainty, to compare the processes considered in the hydrological models, and the 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 and 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 simpler 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 (i.e., summer to winter crops, coniferous to mixed or 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 that are considering water management, relevant management scenarios should also be defined.

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 mod-



els (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.

#### 4.6 Uncertainties from different sources: what are the ways to reduce them?

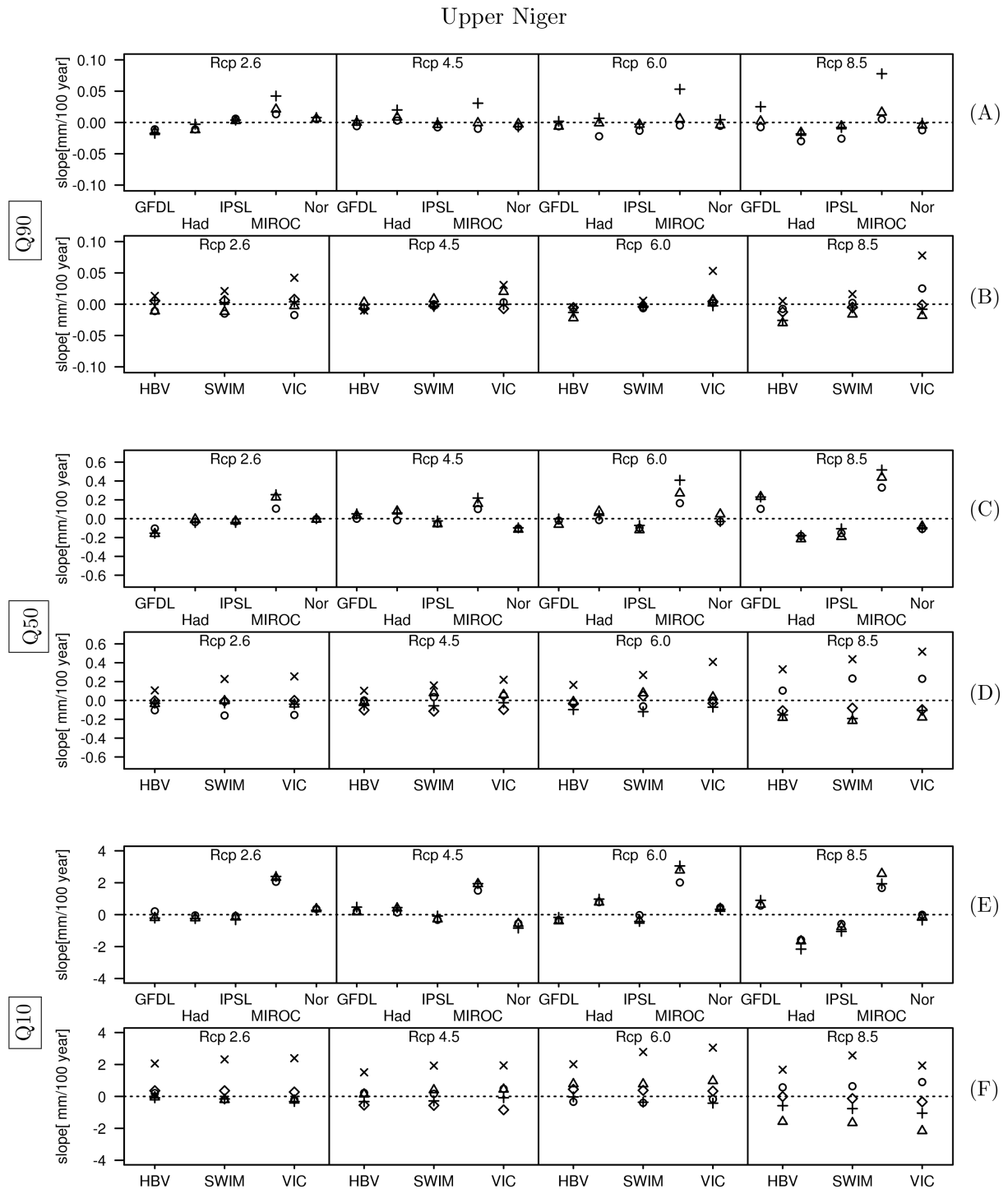
The ideal case would be to improve the description of processes in climate and hydrological models so that the climate models agree better in the climate trends for one specific RCP scenario and region, and the hydrological models agree better in impact projections. However, climate processes are very 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 sub-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 the simulation of different processes and their performance in complex basins, e.g., those at higher altitudes where snow and glacier processes are important, and those with human water management playing an important role in overall water balance.

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 into 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.

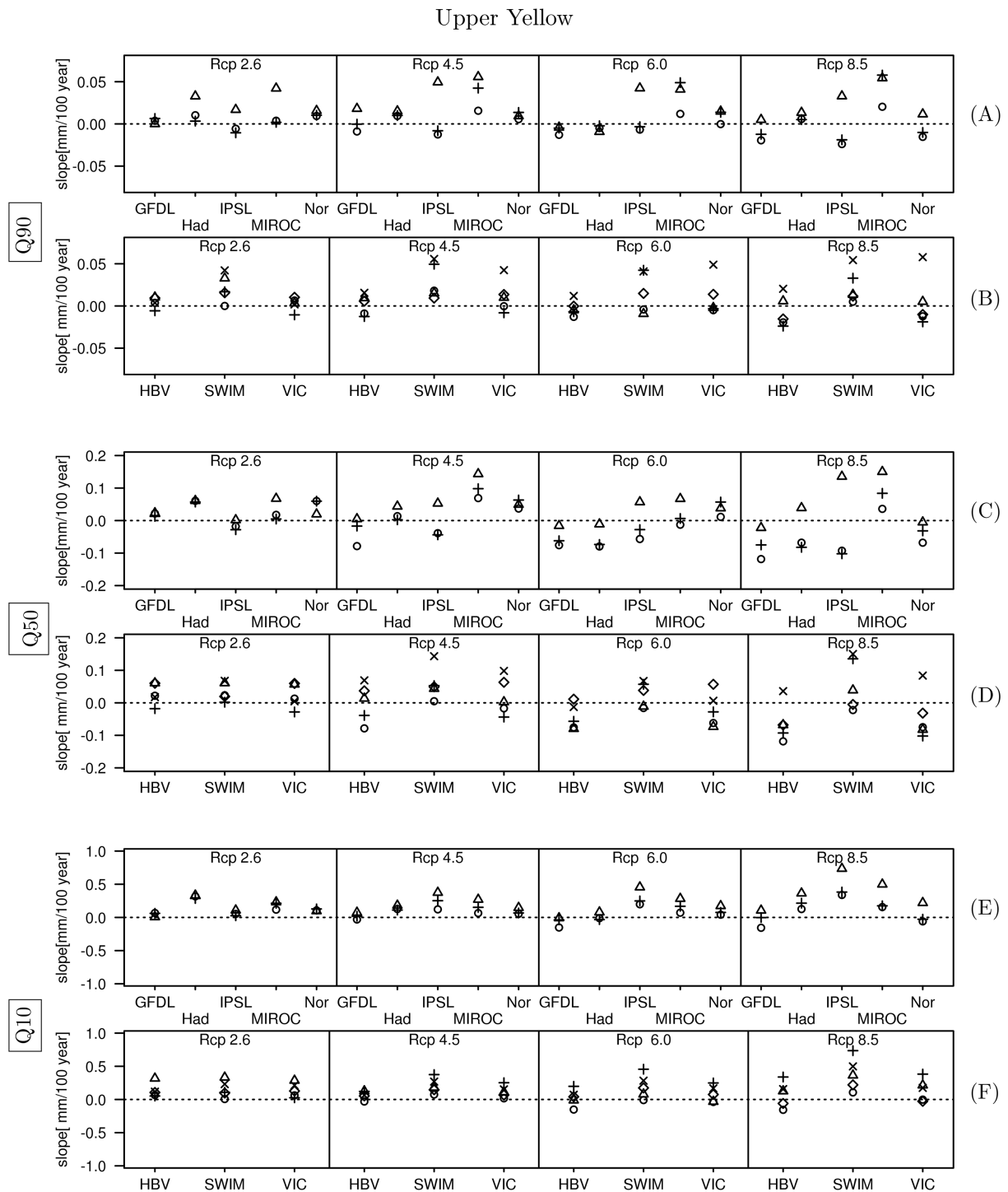
Besides, some scientists (Greene et al., 2006; Zhu et al., 2013) recommend a model-weighting scheme as a feasible approach to reducing uncertainties in climate impact studies. This method gives large weights to the skillful models and minor weights to the models that do not match the observed dynamics. 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, snowmelt, 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 for reducing fragmentarity of impacts on 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.



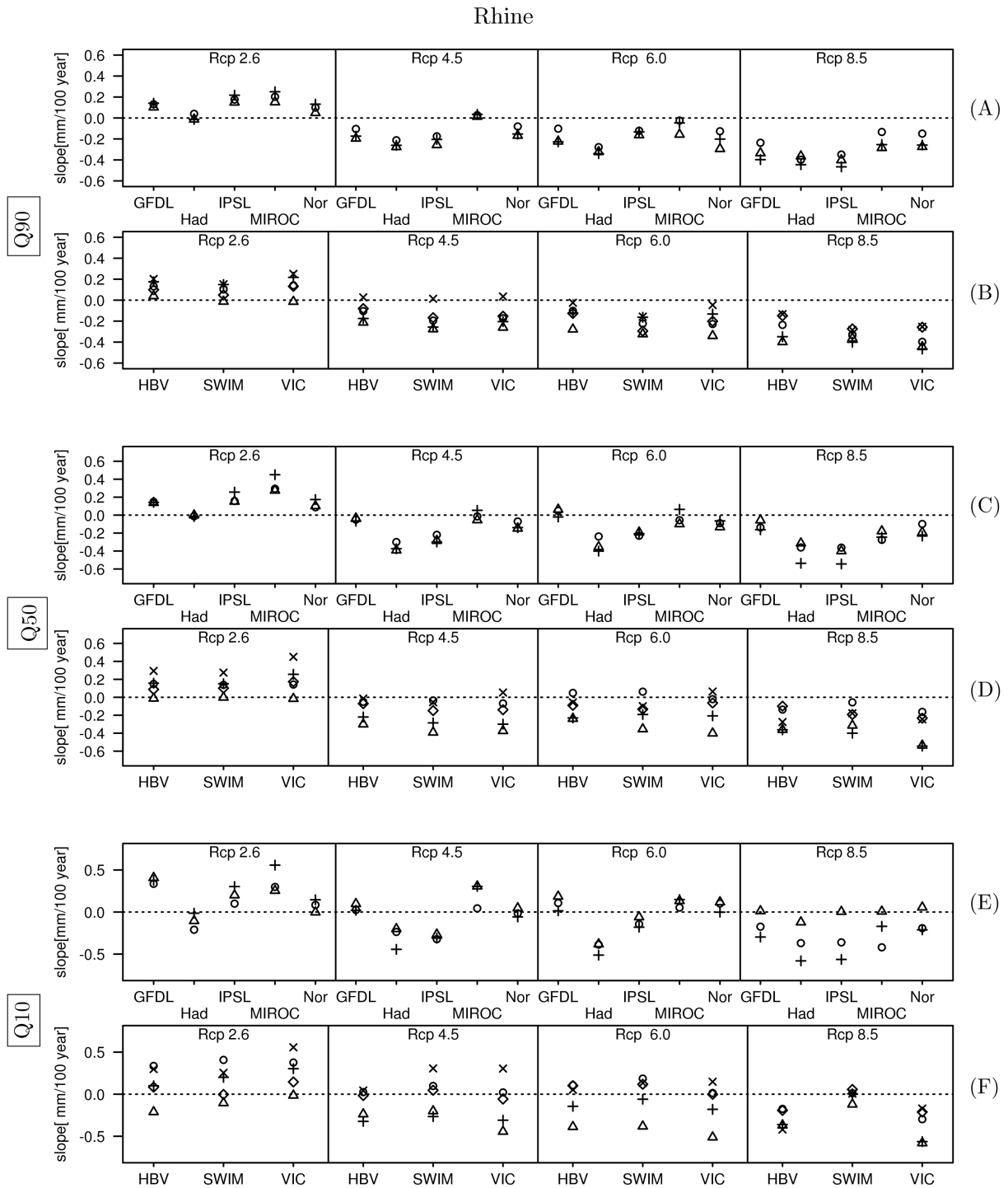
Appendix A



**Figure A1.** Slopes of trends in (a, b) low-flow percentile  $Q_{90}$ , (c, d) medium discharge  $Q_{50}$  and (e, f) high-flow percentile  $Q_{10}$  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  $Q_{90}$ , (c, d) medium discharge  $Q_{50}$  and (e, f) high-flow percentile  $Q_{10}$  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  $Q_{90}$ , (c, d) medium discharge  $Q_{50}$  and (e, f) high-flow percentile  $Q_{10}$  grouped by climate models (a, c, e) and by hydrological models (b, d, f) for the Rhine.

**Acknowledgements.** This work was financed by the project PROGRESS (Potsdam Research Cluster for Georisk Analysis, Environmental Change and Sustainability). For their roles in processing the climate data we like to thank the the ISI-MIP modeling group.

Edited by: R. Pavlick

## References

- Abbott, M. B. and Refsgaard, J. C.: Distributed hydrological modeling, Distributed hydrological modeling. Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.
- Aich, V., Liersch, S., Vetter, T., Huang, S., Tecklenburg, J., Hoffmann, P., Koch, H., Fournet, S., Krysanova, V., Müller, E. N., and Hattermann, F. F.: Comparing impacts of climate change on streamflow in four large African river basins, *Hydrol. Earth Syst. Sci.*, 18, 1305–1321, doi:10.5194/hess-18-1305-2014, 2014.
- Andersson, L., Samuelsson, P., and Kjellström, E.: Assessment of climate change impact on water resources in the Pungwe river basin, *Tellus A*, 63, 138–157, 2011.
- Arnold, J. G., Allen, P. M., and Bernhardt, G.: A comprehensive surface-groundwater flow model, *J. Hydrol.*, 142, 47–69, 1993.
- Bartholomé, E. and Belward, A.: GLC2000: a new approach to global land cover mapping from Earth observation data, *Int. J. Remote Sens.*, 26, 1959–1977, 2005.
- Bergström, S.: Principles and confidence in hydrological modelling, *Nord. Hydrol.*, 22, 123–136, 1991.
- Bergström, S. and Forsman, A.: Development of a conceptual deterministic rainfall-runoff model, *Nord. Hydrol.*, 4, 147–170, 1973.
- Bergström, S. and Singh, V. P.: The HBV model, Computer models of watershed hydrology, Water Resources Publications, Highlands Ranch, Colorado, USA, 443–476, 1995.
- Blöschl, G. and Montanari, A.: Climate change impacts—throwing the dice?, *Hydrol. Process.*, 24, 374–381, 2010.
- Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., and Schaer, C.: Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections, *Water Resour. Res.*, 49, 1523–1536, doi:10.1029/2011WR011533, 2013.
- 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., and Viney, N. R.: Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use, *Adv. Water Resour.*, 32, 129–146, doi:10.1016/j.advwatres.2008.10.003, 2009.
- Chen, X., Yang, T., Wang, X., Xu, C.-Y., and Yu, Z.: Uncertainty Intercomparison of Different Hydrological Models in Simulating Extreme Flows, *Water Resour. Manage.*, 27, 1393–1409, 2013.
- Christensen, N. S. and Lettenmaier, D. P.: A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin, *Hydrol. Earth Syst. Sci.*, 11, 1417–1434, doi:10.5194/hess-11-1417-2007, 2007.
- Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE T. Evolut. Comput.*, 6, 182–197, doi:10.1109/4235.996017, 2002.
- Déqué, M., Rowell, D., Lüthi, D., Giorgi, F., Christensen, J., Rockel, B., Jacob, D., Kjellström, E., De Castro, M., and van den Hurk, B.: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections, *Climatic Change*, 81, 53–70, 2007.
- Doherty, J.: PEST: Model Independent Parameter Estimation, 5th Edn. of user manual, Watermark Numerical Computing, Brisbane, Australia, 2005.
- FAO, IIASA, ISRIC, and ISSCAS: JRC (2009) Harmonized world soil database (version 1.1), FAO, Rome and IIASA, Laxenburg, 2009.
- Finger, D., Heinrich, G., Gobiet, A., and Bauder, A.: 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 Resour. Res.*, 48, W02521, doi:10.1029/2011WR010733, 2012.
- Graham, P.: Modelling runoff to the Baltic basin, *Ambio*, 28, 328–334, 1999.
- GRDC: Global Runoff Data, The Global Runoff Data Centre, Koblenz, Germany, 1998.
- Greene, A. M., Goddard, L., and Lall, U.: Probabilistic multimodel regional temperature change projections, *J. Climate*, 19, 4326–4343, doi:10.1175/JCLI3864.1, 2006.
- 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., and Yeh, P.: Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results, *J. Hydrometeorol.*, 12, 869–884, doi:10.1175/2011JHM1324.1, 2011.
- Hagemann, S., Chen, C., Clark, D. B., Folwell, S., Gosling, S. N., Haddeland, I., Hanasaki, N., Heinke, J., Ludwig, F., Voss, F., and Wiltshire, A. J.: Climate change impact on available water resources obtained using multiple global climate and hydrology models, *Earth Syst. Dynam.*, 4, 129–144, doi:10.5194/esd-4-129-2013, 2013.
- Hamlet, A. F. and Lettenmaier, D. P.: Effects of Climate Change on Hydrology and Water Resources in the Columbia River Basin, *J. Am. Water Resour. Assoc.*, 35, 1597–1623, 1999.
- Hattermann, F. F., Weiland, M., Huang, S., Krysanova, V., and Kundzewicz, Z. W.: Model-supported impact assessment for the water sector in Central Germany under climate change – a case study, *Water Resour. Manage.*, 25, 3113–3134, 2011.
- Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in projections of regional precipitation change, *Clim. Dynam.*, 37, 407–418, doi:10.1007/s00382-010-0810-6, 2011.
- Hempel, S., Frieler, K., Warszawski, L., Schewe, J., and Piontek, F.: A trend-preserving bias correction – the ISI-MIP approach, *Earth Syst. Dynam.*, 4, 219–236, doi:10.5194/esd-4-219-2013, 2013.
- hoc AG-Boden: Bodenkundliche Kartieranleitung, 5. verbesserte und erweiterte Auflage, Bundesanstalt für Geowissenschaften und Rohstoffe, Stuttgart, 2006.
- Huang, S., Krysanova, V., Österle, H., and Hattermann, F. F.: Simulation of spatiotemporal dynamics of water fluxes in Germany under climate change, *Hydrol. Process.*, 24, 3289–3306, doi:10.1002/hyp.7753, 2010.
- Huang, S., Hattermann, F. F., Krysanova, V., and Bronstert, A.: 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, 631–663, 2013.
- Huber, P. J.: Robust regression: asymptotics, conjectures and Monte Carlo, *Ann. Stat.*, 1, 799–821, 1973.
- IPCC: IPCC Special Report Emission Scenarios – Summary for Policymakers. A Special Report of IPCC Working Group III, Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2000.
- IPCC: 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, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007.
- IPCC: Climate Change 2013: The Physical Science Basis, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.
- Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-CORDEX: new high-resolution climate change projections for European impact research, *Reg. Environ. Change*, 14, 1–16, 2013.
- Jarvis, A., Reuter, H., Nelson, A., and Guevara, E.: Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database, <http://srtm.csi.cgiar.org> (last access: 8 January 2015), 2008.
- Kay, A., Davies, H., Bell, V., and Jones, R.: Comparison of uncertainty sources for climate change impacts: flood frequency in England, *Climatic Change*, 92, 41–63, doi:10.1007/s10584-008-9471-4, 2009.
- Koller, M. and Stahel, W. A.: Sharpening Wald-type inference in robust regression for small samples, *Comput. Stat. Data Anal.*, 55, 2504–2515, doi:10.1016/j.csda.2011.02.014, 2011.
- Krysanova, V., Meiner, A., Roosaare, J., and Vasilyev, A.: Simulation modelling of the coastal waters pollution from agricultural watershed, *Ecol. Model.*, 49, 7–29, 1989.
- Krysanova, V., Müller-Wohlfeil, D.-I., and Becker, A.: Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds, *Ecol. Model.*, 106, 261–289, 1998.
- Krysanova, V., Bronstert, A., and Müller-Wohlfeil, D.: Modelling river discharge for large drainage basins: from lumped to distributed approach, *Hydrolog. Sci. J. – Journal des Sciences Hydrologiques*, 44, 313–331, doi:10.1080/02626669909492224, 1999.
- Krysanova, V., Hattermann, F., Huang, S., Hesse, C., Vetter, T., Liersch, S., Koch, H., and Kundzewicz, Z.: Modelling climate and land use change impacts with SWIM: lessons learnt from multiple applications, *Hydrolog. Sci. J.*, accepted, 2015.
- Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res.-Atmos.*, 99, 14415–14428, 1994.
- Liang, X., Wood, E. F., and Lettenmaier, D. P.: Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification, *Global Planet. Change*, 13, 195–206, 1996.
- Lidén, R., Harlin, J., Karisson, M., and Rahmberg, M.: Hydrological modelling of fine sediments in the Odzi River, Zimbabwe, *Water SA*, 27, 303–314, 2001.
- Liersch, S., Cools, J., Kone, B., Koch, H., Diallo, M., Reinhardt, J., Fournet, S., Aich, V., and Hattermann, F. F.: Vulnerability of rice production in the Inner Niger Delta to water resources management under climate variability and change, *Environ. Sci. Policy*, 34, 18–33, 2013.
- Menzel, L., Thielen, A. H., Schwandt, D., and Bürger, G.: Impact of climate change on the regional hydrology–scenario-based modelling studies in the German Rhine catchment, *Nat. Hazards*, 38, 45–61, 2006.
- Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, 1970.
- Ott, I., Duethmann, D., Liebert, J., Berg, P., Feldmann, H., Ihringer, J., Kunstmann, H., Merz, B., Schaedler, G., and Wagner, S.: High-Resolution Climate Change Impact Analysis on Medium-Sized River Catchments in Germany: An Ensemble Assessment, *J. Hydrometeorol.*, 14, 1175–1193, doi:10.1175/JHM-D-12-091.1, 2013.
- Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., and Participants, D.: Overall distributed model intercomparison project results, *J. Hydrol.*, 298, 27–60, doi:10.1016/j.jhydrol.2004.03.031, 2004.
- Refsgaard, J. C. and Knudsen, J.: Operational Validation and Intercomparison of Different Types of Hydrological Models, *Water Resour. Res.*, 32, 2189–2202, doi:10.1029/96WR00896, 1996.
- Saelthun, N. R.: The Nordic HBV model, Norwegian Water Resources and Energy Administration Publication 7, Norwegian Water Resources and Energy Administration, Oslo, Norway, 1–26, 1996.
- 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., and Kabat, P.: Multimodel assessment of water scarcity under climate change, *P. Natl. Acad. Sci. USA*, 111, 3245–3250, doi:10.1073/pnas.1222460110, 2014.
- Smith, M. B., Seo, D.-J., Koren, V. I., Reed, S. M., Zhang, Z., Duan, Q., Moreda, F., and Cong, S.: The distributed model intercomparison project (DMIP): motivation and experiment design, *J. Hydrol.*, 298, 4–26, doi:10.1016/j.jhydrol.2004.03.040, 2004.
- Su, F. and Xie, Z.: A model for assessing effects of climate change on runoff in China, *Prog. Nat. Sci.*, 13, 701–707, 2003.
- Velázquez, J. A., Schmid, J., Ricard, S., Muerth, M. J., Gauvin St-Denis, B., Minville, M., Chaumont, D., Caya, D., Ludwig, R., and Turcotte, R.: 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, doi:10.5194/hess-17-565-2013, 2013.
- Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., and Schewe, J.: The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP): Project framework, *P. Natl. Acad. Sci.*, 111, 3228–3232, 2014.

- Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century, *J. Hydrometeorol.*, 12, 823–848, 2011.
- Wilby, R. L. and Harris, I.: A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK, *Water Resour. Res.*, 42, WR004065, doi:10.1029/2005WR004065, 2006.
- Yip, S., Ferro, C. A. T., Stephenson, D. B., and Hawkins, E.: A Simple, Coherent Framework for Partitioning Uncertainty in Climate Predictions, *J. Climate*, 24, 4634–4643, doi:10.1175/2011JCLI4085.1, 2011.
- Yohai, V. J.: High Breakdown-point and High-efficiency Robust Estimates For Regression, *Ann. Stat.*, 15, 642–656, doi:10.1214/aos/1176350366, 1987.
- Yu, P.-S. and Wang, Y.-C.: Impact of climate change on hydrological processes over a basin scale in northern Taiwan, *Hydrol. Process.*, 23, 3556–3568, 2009.
- Zhang, X. and Lindström, G.: A comparative study of a Swedish and a Chinese hydrological model, *J. Am. Water Resour. Assoc.*, 32, 985–994, doi:10.1111/j.1752-1688.1996.tb04067.x, 1996.
- Zhu, J., Forsee, W., Schumer, R., and Gautam, M.: Future projections and uncertainty assessment of extreme rainfall intensity in the United States from an ensemble of climate models, *Climatic Change*, 118, 469–485, doi:10.1007/s10584-012-0639-6, 2013.