Hydrology and Earth System Sciences

The ability of a GCM-forced hydrological model to reproduce global discharge variability

F. C. Sperna Weiland[1,2], **L. P. H. van Beek**[1], **J. C. J. Kwadijk**[2], **and M. F. P. Bierkens**[1,3]

[1]Department of Physical Geography, Utrecht University, P.O. Box 80115, 3508 TC, Utrecht, The Netherlands
[2]Deltares, P.O. Box 177, 2600 MH, Delft, The Netherlands
[3]Deltares, P.O. Box 80015, 3508 TA, Utrecht, The Netherlands

Abstract. Data from General Circulation Models (GCMs) are often used to investigate hydrological impacts of climate change. However GCM data are known to have large biases, especially for precipitation. In this study the usefulness of GCM data for hydrological studies, with focus on discharge variability and extremes, was tested by using bias-corrected daily climate data of the 20CM3 control experiment from a selection of twelve GCMs as input to the global hydrological model PCR-GLOBWB. Results of these runs were compared with discharge observations of the GRDC and discharges calculated from model runs based on two meteorological datasets constructed from the observation-based CRU TS2.1 and ERA-40 reanalysis. In the first dataset the CRU TS 2.1 monthly timeseries were downscaled to daily timeseries using the ERA-40 dataset (*ERA6190*). This dataset served as a best guess of the past climate and was used to analyze the performance of PCR-GLOBWB. The second dataset was created from the ERA-40 timeseries bias-corrected with the CRU TS 2.1 dataset using the same bias-correction method as applied to the GCM datasets (*ERACLM*). Through this dataset the influence of the bias-correction method was quantified. The bias-correction was limited to monthly mean values of precipitation, potential evaporation and temperature, as our focus was on the reproduction of inter- and intra-annual variability.

After bias-correction the spread in discharge results of the GCM based runs decreased and results were similar to results of the ERA-40 based runs, especially for rivers with a strong seasonal pattern. Overall the bias-correction method resulted in a slight reduction of global runoff and the method performed less well in arid and mountainous

regions. However, deviations between GCM results and GRDC statistics did decrease for $\overline{Q}$, $Q90$ and IAV. After bias-correction consistency amongst models was high for mean discharge and timing (*Qpeak*), but relatively low for inter-annual variability (IAV). This suggests that GCMs can be of use in global hydrological impact studies in which persistence is of less relevance (e.g. in case of flood rather than drought studies). Furthermore, the bias-correction influences mean discharges more than extremes, which has the positive consequence that changes in daily rainfall distribution and subsequent changes in discharge extremes will also be preserved when the bias-correction method is applied to future GCM datasets. However, it also shows that agreement between GCMs remains relatively small for discharge extremes.

Because of the large deviations between observed and simulated discharge, in which both errors in climate forcing, model structure and to a lesser extent observations are accumulated, it is advisable not to work with absolute discharge values for the derivation of future discharge projections, but rather calculate relative changes by dividing the absolute change by the absolute discharge calculated for the control experiment.

1 Introduction

Because runoff regimes might change significantly due to climate change, strategies for water management are sought for that either mitigate the undesired effects of a changing climate or gain from the positive effects. The search for these strategies relies on reliable assessment of the effect of climate change on river discharge. Consequently much research has been conducted investigating the hydrological response to climate change, both on local (Christensen and

Lettenmaier, 2007; Prudhomme and Davies, 2008; Buytaert et al., 2009), regional (Lehner et al., 2006; Strzepek and Yates, 1997; Hagemann et al., 2009) and global scale (Arnell, 1999, 2003; Alcamo et al., 2000; Alcamo and Henrichs, 2002; Milly, 2006).

Future projections can be calculated using change factors; long-term average changes derived from climate model data which are applied to observed meteorological time-series to obtain future meteorological time-series as input for hydrological models (Alcamo and Henrichs, 2007; Nijssen et al., 2001; Vörösmarty et al., 2000). However, with this method changes in temporal variability are ignored. Therefore climate datasets from General Circulation Models (GCM) are often directly used as input for hydrological models when investigating possible changes. Unfortunately, different GCM datasets produce varying and even contradicting results (Varis et al., 2004). GCM results of the current climate control run (20CM3) do not always agree with the observed climate, deviations are especially apparent for precipitation (Covey et al., 2003; Perkins and Pitman, 2009). Often too many days with light rain are simulated and the frequency and amount of heavy rain events are underestimated (Dai, 2006). Because of the variance amongst GCMs many studies concluded that a multi-model ensemble of GCMs should be used to obtain a reliable impression of the spread of possible regional changes and the uncertainties accompanying these changes (Murphy et al., 2004; Boorman and Sefton, 1997; IPCC, 2007). Furthermore, it has been widely recognized (Wood et al., 2004; Leander and Buishand, 2007; Fowler and Kilsby, 2007; Wilby et al., 1998) that precipitation data needs to be bias-corrected before they can be used.

Previous studies (Milly et al., 2006; Nohara et al., 2006) already investigated the correct reproduction of mean discharge and runoff regimes. Therefore we focus on the ability of a GCM-forced hydrological model to reproduce global discharge variability (extremes, seasonal and inter-annual variation), parameters relevant for water management. We realize that even if we obtain a correct reproduction of current discharge variability this is no guarantee that projected discharge variability is correct as well (Prudhomme and Davies, 2008). However, if discharge variability is biased for the current period, future projections of discharge variability are likely to be biased as well.

We restrict ourselves to a bias-correction of monthly mean values of GCM precipitation, temperature and potential evaporation (potential evaporation was derived using either Penman-Monteith or Blaney-Criddle, depending on data availability). Although a correction on monthly means does not guarantee that rainfall marginal distributions are well reproduced (Dai, 2006), no additional correction on GCM variability is applied for two reasons: First, because several climate change experiments have shown that GCM variability will change, especially for precipitation. Precipitation frequency and storm duration is likely to

decrease, while intensity will increase resulting in heavier rain events (Trenberth et al., 2003; Allan and Soden, 2008; Meehl et al., 2000). Therefore, corrections based on past deviations between modeled and observed variability, although often applied (Leaner and Buishand, 2007; Ines and Hansen, 2006; Wood et al., 2004), may not hold in future. Applying variability corrections might even mask important changes, since changes in rainfall distributions may have a much larger effect on the hydrological cycle than changes in mean precipitation (Allen and Ingram, 2002). Second, as discussed before, the very goal of this study is to examine the effect of GCM variability on modeled hydrological variability, thereby assessing the usability of GCM data for water management related climate effect studies. Replacing the differences in variability between GCMs to a single current observed variability (i.e. CRU) would rule out such an analysis.

Our evaluation of the ability of a GCM-forced hydrological model to reproduce global discharge variability consists of three steps. First, the performance of PCR-GLOBWB is assessed when forced directly with the CRU TS 2.1 monthly timeseries (New et al., 2000) which are downscaled to daily values with the ERA-40 re-analysis (Uppala et al., 2005) for the period 1961 to 1990. The CRU TS 2.1 timeseries are derived by interpolation of monthly meteorological observations while the ERA-40 reanalysis reflects the daily variations in the large-scale meteorological conditions over the globe. Thus, the down-scaled data set provides a best guess of the inter-annual and seasonal variability. River discharge statistics resulting from the PCR-GLOBWB run driven with this forcing dataset are then compared with observed river discharge statistics from the GRDC dataset (GRDC, 2007). Although this comparison is subject to any errors in the forcing dataset, evaluation of the model results from this run provides the best available information on the performance of PCR-GLOWB.

In a second step, the ERA-40 dataset is bias-corrected using the same method as used for the GCM datasets as a yearly month-by-month correction with the CRU TS2.1 monthly time-series, similar to step 1, is not possible for GCM datasets; GCM timeseries only provide a realization of a given climate and do not represent the observed meteorology of a specific year. Rather, each monthly GCM field is updated in a manner that scales the 30-year monthly average GCM quantity to the 30-year monthly average CRU TS2.1 quantity. With this method the inter-annual variability remains similar to that of the original GCM or ERA-40 timeseries. Results of the hydrological model run driven with this forcing dataset are compared with results of the hydrological model run of step 1 and with observed GRDC data in order to quantify the influence of the bias-correction method as this preserves the full temporal variability of the modeled meteorological product (ERA-40, GCMs) instead of imposing the observed seasonal and inter-annual signal of the CRU TS 2.1.

In the last step PCR-GLOBWB is forced with the meteorological datasets of 12 GCMs bias-corrected according to the method described in step 2. To quantify the spread between the results of the GCMs the coefficient of variation (CV) is calculated for a number of statistical quantities derived for the ensemble of GCMs. Combined with the findings on model performance (step 1) and loss of observed temporal variability (step 2), these results will show the usability of meteorological GCM datasets in hydrological impact studies with a focus on extremes and discharge variability.

2 Hydrological model

2.1 Existing global hydrological models

Obviously, the macro-scale hydrological model (MHM) PCR-GLOBWB follows in a long line of existing MHMs. Without attempting to be complete, we refer to short reviews given by Arnell (1999) and Döll et al. (2003) describing VIC (Nijssen et al., 2001), Macro-PDM (Arnell, 1999), WBM (Vörösmarty et al., 1998) and WGHM (Döll et al., 2003), four models frequently used in large-scale hydrological studies. Similar to PCR-GLOBWB, the last three models calculate for each time-step the water balance of all individual grid cells. The grids of WBM and WGHM have a resolution of 0.5 degrees, corresponding to the finest resolution of most climate datasets available, within Macro-PDM grid cells can either be regular or catchment shaped. All three models contain at least one soil water layer and total runoff consists of a fast overland and a slow groundwater component. Size and partitioning of these fluxes depend on the degree of saturation of the soil water layer(s) that is calculated either physically based or described by a statistical relation. The models all apply some form of routing to obtain realistic river discharge. WGHM (Döll et al., 2003) is the sub model of the global water use and availability model WaterGAP (Alcamo et al., 2003).

In addition to stand-alone hydrological models, global water balances have also been modeled by coupled vegetation water balance models and by land surface schemes (LSS) used in global climate models. Examples of global coupled vegetation water balance models are GEPIC (Liu et al., 2009) and LPJ (Gerten et al., 2004). GEPIC focuses on the calculation of crop yield and crop water productivity. It combines the extent of crop covered areas (with different properties for a variety of crop types), Hargreaves potential evaporation and soil water availability to calculate crop water use. LPJ is a global dynamic vegetation model that describes the interaction between the terrestrial biosphere and the water cycle. Runoff is one of the outputs of this vegetation model and exists of excess over field capacity from the upper two soil layers and percolation from the second soil layer. Runoff is translated to river discharges using a routing scheme. At this stage the global water balance and basin runoff is represented less reliable by most LSSs than by hydrological models (Gerten at al., 2004) and the resolution is too coarse for hydrological studies. A notable exception is global VIC (Nijssen et al., 2001). VIC is designed to be a land surface scheme for climate models and solves both the energy and the water balance. At the global scale it currently has a resolution of 1 degrees (Sheffield et al., 2009), where vegetation, soil moisture and the application of precipitation are modeled using sub-grid variability schemes. Runoff, existing of baseflow from the lower soil moisture store and fast response flow, is routed along a routing network using a convolution approach. VIC is one of the few LSSs that is frequently used in stand-alone hydrological studies.

The main focus of this study is the reproduction of variability in runoff regimes. Therefore the hydrological model used should be designed to calculate this variability, including a good representation of hydrology. In addition to the models described above the model used in this study, PCR-GLOBWB, contains an advanced scheme for the subgrid parameterization of surface runoff, interflow and baseflow and an explicit routing scheme for surface water flow using the kinematic wave approximation that includes retention in and evaporation loss from wetlands, lakes and reservoirs (Van Beek and Bierkens, 2009).

2.2 PCR-GLOBWB

PCR-GLOBWB is a global distributed hydrological model with a resolution of 0.5 degrees (Van Beek and Bierkens, 2009; Bierkens and van Beek, 2009). Each model cell consists of two vertical soil layers and one underlying groundwater reservoir. Sub-grid parameterization is used to represent short and tall vegetation, surface water and for calculation of saturated areas for surface runoff as well as interflow. Water enters the cell as precipitation and can be stored as canopy interception or snow. Snow melt or accumulation occurs depending on temperature. Melt water and throughfall are passed to the surface. Evapotranspiration is calculated from the potential evaporation and soil moisture conditions. Exchange of water is possible between the soil and groundwater layers in both downward and upward direction depending on soil moisture status and groundwater storage. Total runoff consists of non-infiltrating melt water, saturation excess surface runoff, interflow and base flow. For each time-step the water balance is computed per cell. Runoff is accumulated and transferred as river discharge along the drainage network using kinematic wave routing. The drainage network is taken from DDM30 (Döll and Lehner, 2002) and lakes, wetlands and large reservoirs are obtained from the GLWD dataset (Lehner and Döll, 2004). The model calculates natural flows. Water use is not considered and only recently, in a newer version of PCR-GLOBWB, reservoir management has been included.

2.3 Data

2.3.1 GCM data

In 1997 the IPCC developed a set of emission scenarios, representing possible future climate change and provided boundary conditions to be used in GCM runs. These scenarios are widely used in climate impact studies. In addition to boundary conditions for studies based on these scenarios, the IPCC also provided boundary conditions for a 20 century control experiment. Climate modelling centers around the world conducted GCM runs with this data. The Program for Climate Model Diagnosis and Intercomparison (PCMDI) has collected the results and made these available through the PCMDI data portal (https://esg.llnl.gov:8443/index.jsp). The PCMDI provides results on a daily time-step, whereas the IPCC data portal only provides derived monthly averages. Although it has been said that daily values are less reliable (Prudhomme et al., 2002), we prefer to use daily data, since they provides more information on extremes and climate variability.

We collected data from the GCMs for which complete model datasets are provided on a daily time step for both the 20C3M control experiment and the future scenario A2, which we will analyze in future research. We selected the period 1961 to 1990, assuming that a thirty year period would be long enough to represent inter-annual variability, and all GCMs provide data for this period. For those GCMs with multiple runs available the first run was selected. We worked with a 30-year period of day values assuming this period is that long that potential biases in the selected GCM run are averaged out. Although the data portal does not provide all required parameters for the (state-of-the-art) Hadley centre climate models, HADGEM2 has been included. Data were retrieved from the CERA-gateway, http://cera-www.dkrz.de. Table 1 gives an overview of the selected models.

2.3.2 ERA-40 re-analysis data

Precipitation, temperature and evaporation were collected from the ERA-40 reanalysis as well. The ERA-40 dataset is obtained with a numerical weather prediction system that includes assimilation of meteorological observations. The prediction system reasonably represents the observed climate. It has a relative high horizontal resolution (ca. 125 km) and a vertical resolution of 60 levels (Uppala et al., 2005). Still precipitation is poorly approximated in ERA-40 data, particularly in the tropics (Troccoli and Kålberg, 2004). The ERA-40 datasets slightly overestimates precipitation globally. Overestimations are largest in the Congo basin, the southern slope of the Himalaya and the Andes region, while precipitation is underestimated in the Murray, Danube and Mississippi basins. In an attempt to eliminate the cold Arctic bias present in the ERA-15 dataset, a warm bias during winter is introduced in the Arctic region

in ERA-40 dataset. Inter-annual variability is relatively high because the observational data included in the system vary over time (Hagemann et al., 2005). This variation increased, amongst others, the inter-annual variability of precipitation totals, which is relatively high after the first satellite data was included in the system for 1972.

2.3.3 CRU data

Precipitation, temperature and the available parameters required to derive Penman-Monteith reference potential evapotranspiration have been retrieved from the monthly CRU TS 2.1 timeseries (New et al., 2000). Since radiation and wind speed are not provided as monthly timeseries, the climatology from the CLIM 1.0 (New et al., 1999) was used. The CRU data is based on station data that is interpolated as a function of longitude, latitude and elevation above sea level using the thin-plate spline method. The CRU TS 2.1 timeseries have been validated in several studies. Unfortunately no benchmark meteorological dataset exists and validations are mainly based on differences between historical datasets. Available historical data sources for precipitation (FAO, GHCN and CRU) correspond well with each other (Beck et al., 2004). Uncertainties in precipitation amounts and seasonality are largest in dry regions (Fekete et al., 2004). In Arctic regions the CRU datasets underestimate precipitation amounts due to snow undercatch problems (Fiedler and Döll, 2007).

2.3.4 Hydrological data

We selected 19 large catchments to be evaluated in this study, covering a variety of climate zones, latitudes and continents. The selected catchments are shown in Fig. 1, together with basin characteristics. Discharge data are obtained from the Global Runoff Data Center (GRDC, 2007). For most rivers daily data was available for at least part of the period of interest. For the Yangtze, Yellow River and Parana we retrieved monthly GRDC data and for the Indus we could only retrieve monthly data from the RivDis database (Vörösmarty et al., 1998).

2.4 Derivation of potential evapotranspiration

Evaporation is derived with the Penman-Monteith equation (Monteith, 1965). For those GCMs where the required atmospheric surface pressure data were not available, surface pressure was derived from the pressure at sea level using a global DEM. Air humidity fields, required to calculate the actual vapor pressure, could not be retrieved from the data portal for the complete period for some of the GCMs. Therefore we used a simplified method to calculate the actual vapor pressure from the minimum air temperature (Allen et al., 1998). For arid regions the assumption that the air is saturated when the temperature is at its minimum might not hold and therefore the minimum temperature
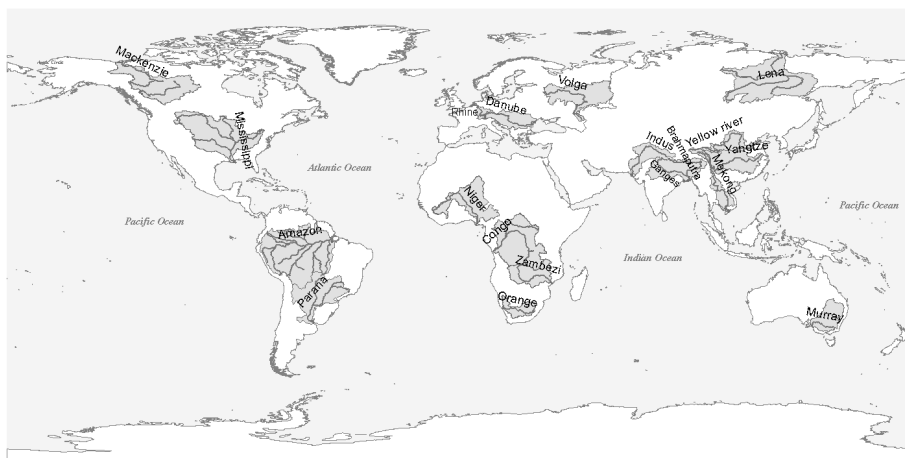
Catchment	Area (km^2)	Qavg (m^3/s)	Gauge	Catchment	Area (km^2)	Qavg (m^3/s)	Gauge
Amazone	6.915.000	190.000	Obidos	Murray	1.061.469	767	Wakool Junction
Brahmaputra	930.000	48.160	Bahadurabad	Niger	2.117.700	6.000	Dire
Congo River	3.680.000	41.800	Kinshasa	Orange river	973.000	365	Aliwal North
Danube	817.000	6.400	Ceatal Izmail	Parana	2.582.672	18.000	Corientes
Ganges	907.000	12.015	Hardinge Bridge	Rhine	65.683	2.200	Rees
Indus	1.165.000	6.600	Kotri	Volga	1.380.000	8.060	Volgograd
Lena	2.500.000	17.000	Kusur	Yangtze	1.800.000	31.900	Datong
MacKenzie	1.805.000	10.700	Norman Wells	Yellow river	752.000	2.571	Huayuankou
Mekong	795.000	16.000	Muhdahan	Zambezi	1.390.000	3.400	Katom a Mulilo
Mississippi	2.981.076	12.743	Vicksburg				

Fig. 1. Selected catchments.

will not equal the dew temperature. As suggested by Allen (1998) we subtracted 2 degrees from the minimum temperature in the arid regions. Arid regions have been selected using the climate moisture indices of the WWDRII (UN, 2006). For those models where time-series for other required parameters were missing or incomplete we used the Blaney-Criddle equation (Brouwer and Heibloem, 1986; Oudin et al., 2005) instead of Penman-Monteith. We realize this may have introduced additional noise between the model results (Kay and Davies, 2008). For several models we compared the potential evaporation we calculated using Penman-Monteith with evaporation calculated with the Blaney-Criddle equation. The full analysis has not been included in this article for brevity, however potential evaporation calculated with the Blaney-Criddle equation is relatively high for Europe, North-America and the north of

Asia during summer and relatively low for Africa and Asia. Especially the overestimation in the Northern continents can result in deviations because, particularly at the beginning of summer, evaporation will not be limited by water availability and actual evaporation will be too high as well. However, hydrological model studies are forced to use what has been reported by the GCMs host institutes and evaporation is not provided for most models.

2.5 Bias correction of ERA-40 and GCM data with 30-year mean month values

For the bias-correction of the GCM and ERA-40 datasets, monthly scaling factors were calculated from the difference (temperature) or ratio (precipitation and evaporation) in 30-year average monthly means between the CRU TS 2.1

Table 1. Overview of selected GCMs.

Model	Institute	Country	Acronym
BCM2.0	Bjerknes Centre for Climate Research	Norway	BCCR
CGCM3.1	Canadian Centre for Climate Modelling and Analysis	Canada	CCCMA
CGCM2.3.2	Meteorological Research Institute	Japan	CGCM
CSIRO-Mk3.0	Commonwealth Scientific and Industrial Research Organization	Australia	CSIRO
ECHAM5	Max Planck Institute	Germany	ECHAM
ECHO-G	Freie Universität Berlin	Germany	ECHO
GFDL-CM 2.1	Geophysical Fluid Dynamics Centre	USA	GFDL
GISS-ER	Goddard institute for Space Studies	USA	GISS
IPSL-CM4	Institute Pierre Simon Laplace	France	IPSL
MIROC3.2	Center of Climate System Research	Japan	MIROC
CCSM3	National Center for Atmospheric Research	USA	NCAR
HADGEM2	Met Office's Hadley Centre for Climate Prediction	UK	HADGEM

Table 2. Descriptive statistics. Where *Qday* is a daily discharge value, *k* corresponds to the day number, *K* is the total number of days in the 30-year period, $Q90$ is the discharge exceeded at 10% of the days during the 30-year period, $Q10$ is the discharge exceeded at 90% of the days during the 30-year period, j corresponds to the year number, *N* is the total number of years (30), *Qyear$_j$* is the mean annual discharge for the years 1961 to 1990, $\overline{Q}$ is the 30 year average mean discharge and $\overline{Q}_{\mathrm{month}}$ are the 30-year average mean monthly discharges.

Parameter	Description	Equation	Unit
$\overline{Q}$	Thirty year average mean annual discharge	$\overline{Q} = \frac{1}{K} \sum\limits_{k=1}^{K} Qday_k$	m^3/s
relative Q90	Discharge exceeded at 10% of the days during the thirty year period relative to mean discharge	$Q90/\overline{Q}$	–
relative Q10	Discharge exceeded at 90% of the days during the thirty year period relative to mean discharge	$Q10/\overline{Q}$	–
IAVrel	Inter annual variability derived from yearly average discharges	$IAV = \dfrac{\sqrt{\frac{1}{N} \sum\limits_{j=1}^{30} \left(Qyear_j - \overline{Q} \right)^2}}{\overline{Q}}$	–
Qpeak	Month of regime peak occurence	$Q_{\mathrm{peak}} = \max(\overline{Q}_{\mathrm{month}})$	month

* For the stations where no daily time-series were available (Indus, Parana, Yangtze and Yellow River) the observed Q90 and Q10 values could not be calculated and are therefore not included in the analysis.

timeseries (New et al., 2000) and GCM/ERA-40 timerseries for the period 1961 to 1990. Applying the monthly correction factors also implies downscaling GCM data spatially to the higher spatial resolution of the CRU dataset, since correction factors are calculated for all individual cells of the CRU grid. In this section both the GCM and ERA-40 data will be referred to as modeled dataset (*MOD*), the bias-corrected ERA-40 dataset created in this section will be referred to as *ERACLM*. For temperature an additive correction was used:

$$T_{\text{corrected\_MOD}} = T_{\text{MOD}} + (\overline{T}_{\text{CRU}} - \overline{T}_{\text{MOD}}) \qquad (1)$$

Where T is the daily temperature (K) and $\overline{T}$ is the 30-year average monthly temperature. For evaporation (*ETP*) a multiplicative correction is used, to avoid the occurrence of negative evaporation.

$$ETP_{\text{corrected\_MOD}} = ETP_{\text{MOD}} \frac{\overline{ETP}_{\text{CRU}}}{\overline{ETP}_{\text{MOD}}} \qquad (2)$$

Where *ETP* is the daily potential evaporation (mm/day) and $\overline{ETP}$ is the 30-year average monthly potential evaporation. For precipitation a similar multiplicative correction is used as for evaporation.

$$P_{\text{corrected\_MOD}} = P_{\text{MOD}} \frac{\overline{P}_{\text{CRU}}}{\overline{P}_{\text{MOD}}} \qquad (3)$$

Where P is the daily precipitation (mm/day) and $\overline{P}$ is the 30-year average monthly precipitation. For some regions (North-Africa, Amazon and Himalaya) differences between the modeled monthly precipitation amount and CRU TS 2.1 monthly precipitation amounts and number of wet days can be very large. In these regions a simple multiplicative correction resulted in unrealistic precipitation peaks in the bias-corrected precipitation time-series. Therefore the bias-correction of precipitation, Eq. (3), is extended with a minimum daily precipitation amount that has to be exceeded by the modeled total monthly precipitation amount before the multiplicative correction can be used (Van Beek, 2008). The threshold equals the monthly mean daily CRU precipitation amount:

$$P_{\text{crit}} = \frac{\overline{P}_{\text{CRU}}}{\overline{W}_{\text{CRU}}} \qquad (4)$$

Where $\overline{W}_{\text{CRU}}$ is the 30-year average number of wet days for the specific month. In addition a threshold value of 10 is set for the maximum value of the multiplicative correction factor ($\overline{P}_{\text{CRU}}/\overline{P}_{\text{MOD}}$). If 1) the monthly precipitation sum of the GCM does not exceed the threshold P_{crit} or if 2) the multiplicative correction factor is higher than 10, the days with precipitation occurrence are calculated from a temperature limit below which a day becomes wet. With this method the number of wet days is increased to avoid large rain events on the few days with rain in the GCM timeseries. The temperature limit is defined by:

$$T_{\text{crit\,GCM}} = T_{\text{min\,GCM}} + (T_{\text{max\,GCM}} - T_{\text{min\,GCM}}) \cdot \frac{\overline{W}_{\text{CRU}}}{N} \qquad (5)$$

Where T_{min} and T_{max} are the minimum and maximum temperature (Kelvin) of the given month and N is the total number of days in the specific month. With this formula the number of wet GCM days per month is calculated (W_{GCM}) and the precipitation amount for these days equals:

$$P_{\text{corrected\_GCM}} = \frac{\overline{P}_{\text{CRU}}}{W_{\text{GCM}}} \qquad (6)$$

With this equation, precipitation is equally distributed over the wet days and the original temporal distribution of the GCM precipitation time-series is lost. However, this correction method only had to be applied for a limited number of cells.

2.6 Downscaling of CRU monthly timeseries to daily timeseries with ERA-40

The CRU TS 2.1 monthly timeseries are downscaled to daily timeseries with the ERA-40 re-analysis dataset. The method used, resembles the bias-correction method described in Sect. 2.5. The only exception is that the 30-year average monthly precipitation, temperature, evaporation and number of wet days ($\overline{P}, \overline{T}, \overline{ETP}$ and $\overline{W}_{\text{CRU}}$) should be replaced in all equations by the monthly mean values of the specific month in the specific year, since for the ERA-40 dataset a year by year correction with the CRU TS 2.1 timeseries can be made. The resulting dataset will be referred to as *ERA6190* and is used in the first step of this study for the validation of the hydrological model.

2.7 Statistical analysis

2.7.1 Statistics

We calculated the thirty year average mean discharge, 10 (low flow) and 90 (high flow) percentile values for each catchment for results of all model runs and the observed discharge timeseries. From thirty year average quantities of the individual GCMs we calculated the GCM ensemble mean. When the mean modeled discharge deviates from the mean observed discharge, Q90 and Q10 values are likely to deviate as well. Therefore, instead of comparing the absolute Q10 and Q90 values, we compared the Q90 and Q10 values relative to the mean discharge (see Table 2). These quantities provide information on the shape of the discharge PDF, the positions of Q90 and Q10 are given relative to the mean discharge.

To compare the temporal behavior of the GCM and ERA-40 based model runs with observations, we also calculated the inter-annual discharge variability for the 30 year annual average discharges and the yearly month of regime peak occurrence (Table 2). Furthermore, the
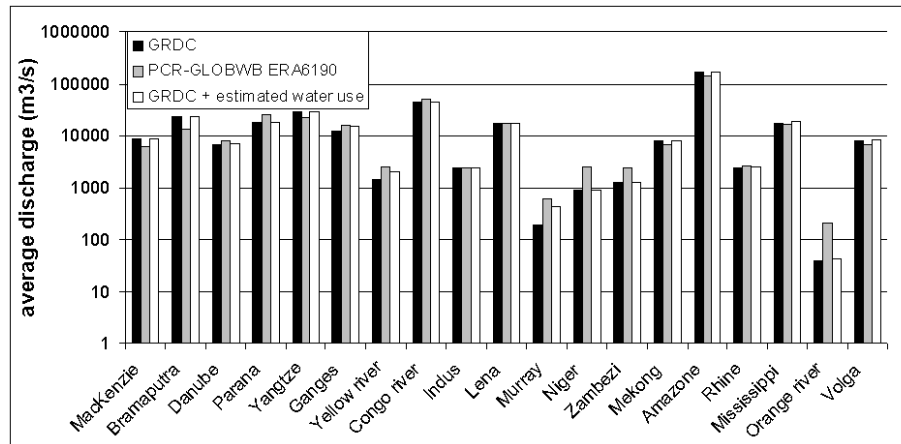
Fig. 2a. Annual mean observed discharge (GRDC), annual mean discharge calculated with PCR-GLOBWB forced with the *ERA6190* dataset (PCR-GLOBWB ERA6190) and the sum of observed annual mean discharge and estimated water use (GRDC + estimated water use) as an approximation of natural flow.
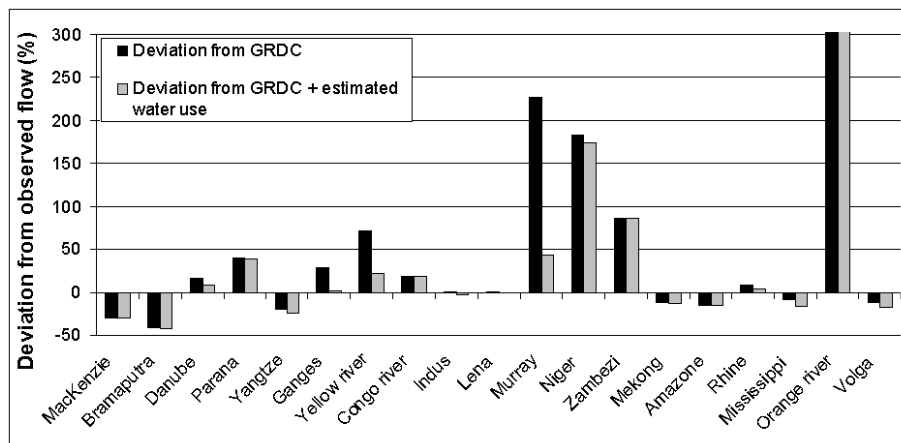

Fig. 2b. Percentage deviation of annual mean modeled discharge (PCR-GLOBWB, *ERA6190* run) from 1) observed annual mean discharge (deviation from GRDC) and 2) the sum of observed annual mean discharge and estimated water use as an approximation of natural flow (Deviation from GRDC + estimated water use).

coefficient of variation (CV) of the ensemble of 12 GCM results is calculated for all statistics according to:

$$CV = \frac{\sqrt{\frac{1}{M} \sum_{i=1}^{12} \left(\overline{Q}_i - \overline{Q}_{\mathrm{ENS}} \right)^2}}{\overline{Q}_{\mathrm{ENS}}} \qquad (7)$$

Where M is the number of GCMs (12), i corresponds to one of the 12 GCMs, $\overline{Q}_i$ is the 30-year average discharge for the specific GCM and $\overline{Q}_{\mathrm{ENS}}$ is the GCM ensemble mean 30-year average discharge; all discharges in $\mathrm{m^3 \cdot s^{-1}}$. In Eq. (7) $\overline{Q}$ is used as an example, however the same equation is used for relative $Q10$, relative $Q90$, $Qpeak$ and IAV. The CV is a measure for the ensemble spread and indicates the resemblance of the PCR-GLOBWB runs driven with the different bias-corrected GCM datasets.

2.7.2   Regime curves

The hydrological regime curve consists of the 30-year average mean monthly discharges, calculated for all twelve months individually from modeled discharge timeseries at the station locations (Fig. 1). Regime curves have been calculated from: (1) discharge simulated with the *ERA6190* dataset, (2) discharge simulated with the *ERACLM* dataset, (3) discharges simulated for each individual bias-corrected GCM dataset, (4) the ensemble mean of the discharges calculated by these individual GCMs and for comparison (5) the observed GRDC discharges. In addition to plots of the regime curves we calculated Nash-Sutcliffe (NS) coefficients (Nash and Sutcliffe, 1970) from the 30-year average monthly values displayed in the regime curves.

Table 3. Continental runoff based on data and model based estimates in km^3/year (Wada et al., 2008).

	Europe	Asia	Africa	N.-America	S.-America	Oceania	Global	Time Period
Data based estimates:								
Baumgartner and Reichel (1975)	2564	12467	3409	5840	11 039	2394	37 713	–
Korzun et al. (1978)	2970	14 100	4600	8180	12 200	2510	44 560	–
L'vovich (1979)	3110	13 190	4225	5960	10 380	1965	38 830	–
Shiklomanov (1997)	2900	13 508	4040	7770	12 030	2400	42 648	1921–1990
GRDC (2004)	3083	13 848	3690	6294	11 897	1722	40 533	1961–1990
Average	2925	13 423	3993	6809	11 509	2198	40 857	–
Model based estimates:								
Fekete et al. (2000)	2772	13091	4517	5892	11 715	1320	39 319	–
Vörösmarty et al. (2000)	2770	13 700	4520	5890	11 700	714	39 294	1961–1990
Nijssen et al. (2001)	–	–	3615	6223	10 180	1712	36 006	1980–1993
Oki et al. (2001)	2191	9385	3616	3824	8789	1680	29 485	1987–1988
Döll et al. (2003)	2763	11234	3592	5540	11 382	2239	36 687	1961–1990
Widén-Nilsson et al. (2007)	3669	13 611	3738	7009	9448	1129	38 605	1961–1990
Average	2833	12204	3933	5730	10536	1466	36 566	–
PCR-GLOBWB:								
ERA6190	2143	11 461	5573	5249	11 186	2633	36 245	1961–1990
ERACLM	2159	10 285	5223	4848	10 174	2384	35 075	1961–1990
GCM ensemble mean	2175	10 572	5099	4803	10 678	2371	35 699	1961–1990

(*) This table with continental discharges from previous studies has been obtained from Wada et al. (2008).

2.7.3 Map comparison

PCR-GLOBWB calculates daily maps of routed accumulated runoff per cell. For all statistical variables (e.g., $\overline{Q}$, relative $Q90$ and $Q10$, Qpeak, IAV) the GCM ensemble mean results and the results of the two ERA-40 based runs are compared. Maps with the CV are derived from the ensemble of GCM results for all variables as well. These maps show a spatial distribution of the spread between the results of the bias-corrected GCM runs.

3 Results

3.1 PCR-GLOBWB validation using *ERA6190*

To assess its performance we ran the model PCR-GLOBWB for the period 1961 to 1990 with the monthly CRU TS 2.1 timeseries, downscaled to daily values using the ERA-40 re-analysis data. This forcing dataset (*ERA6190*) provides the best available guess of the current climate. Despite its

coarse resolution PCR-GLOBWB can reproduce observed discharges reasonably well for most selected catchments. Other than through the hydrological model structure, errors can also be introduced by the model input, in particular through the meteorological forcing.

3.1.1 Continental discharges

In Table 3 the mean yearly continental discharges calculated with PCR-GLOBWB, are compared with continental discharges of previous studies calculated from both observations and model results (Wada et al., 2008). Continental discharges vary amongst studies. For Asia, North-America, South-America and Oceania the results of PCR-GLOBWB are within the range of continental discharges found in previous studies. Differences between the discharges modeled by PCR-GLOBWB and the average continental discharges of previous studies are small, globally the average difference is only 1%. For Europe PCR-GLOBWB slightly underestimates continental discharge. For Africa PCR-GLOBWB modeled discharge is on the high side,
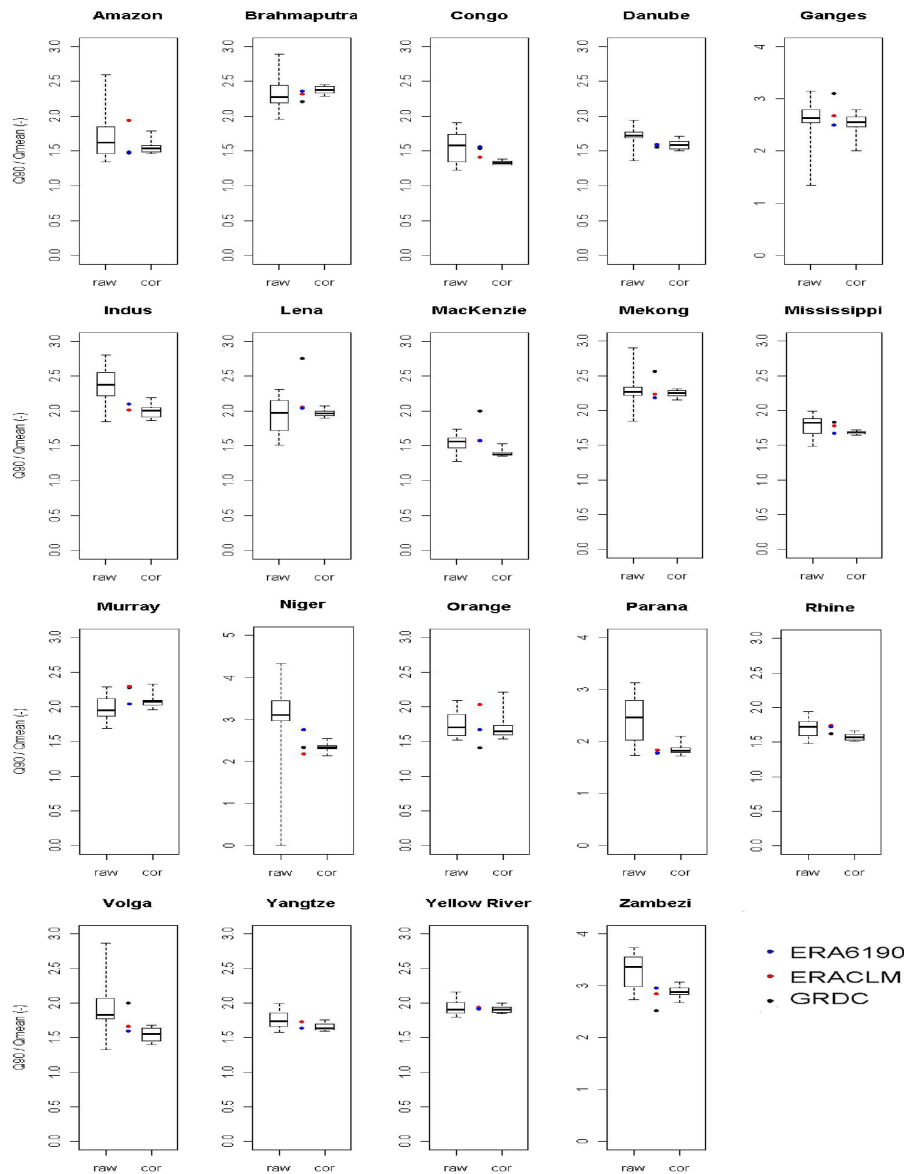
Fig. 3. Thirty year average relative Q90 values ($Q90/\overline{Q}$ (–)) for the ensemble of GCMs (boxes are drawn between the quartiles and whiskers represent minimum and maximum values), the *ERA6190* run (blue points), the *ERCLM* run (red points) and the GRDC observations (black points).

with PCR-GLOBWB overestimating the discharge for many African rivers. This is caused by the absence of water withdrawals in the model and limited river bed losses, like recharge to the groundwater and evaporation (which is particular relevant in the deltas of the Niger and the Nile). In addition, the density of meteorological observations is relatively low for this continent which limits the constraint of the CRU TS 2.1 and ERA-40 datasets (Fekete et al., 2004; Van Beek and Bierkens, 2009).

3.1.2 Discharge statistics

Figure 2a shows the annual mean modeled and observed discharges and Fig. 2b shows the modeled discharge as a percentage of the observed GRDC discharge. The annual mean modeled discharge shows little deviation from observed annual mean discharge for the Danube, Volga, Rhine, Mississippi and Lena. Deviations are relatively large for the drier basins; Murray, Niger, Zambezi and Orange. Modeled discharge for the MacKenzie is too low due to undercatch in the CRU snowfall amounts (Fiedler and Döll, 2007). Discharge is also relatively low for the Monsoon
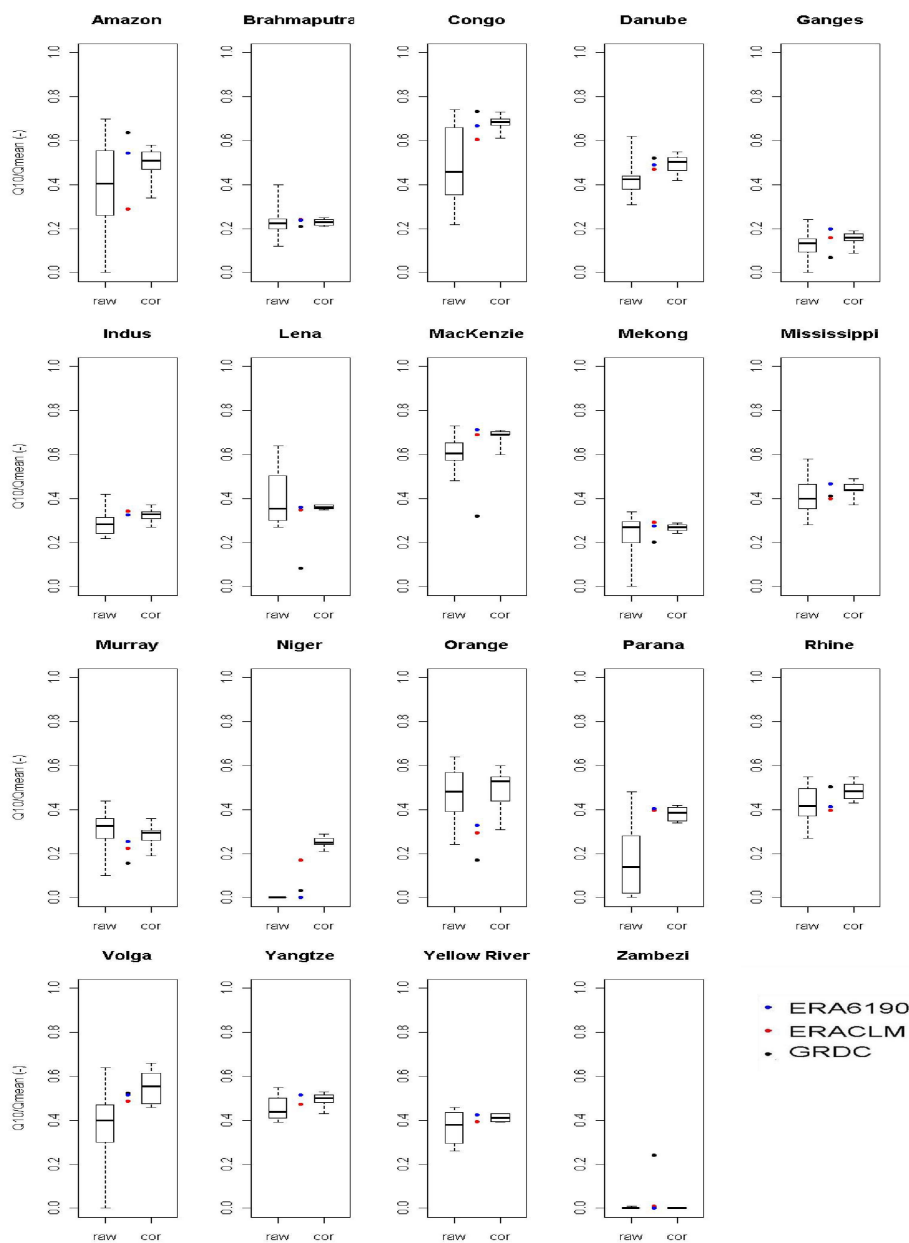
Fig. 4. Thirty year average relative Q10 values ($Q10/\overline{Q}(-)$) for the ensemble of non-bias corrected GCM runs (raw) and bias-corrected GCM runs (cor; boxes are drawn between the quartiles and whiskers represent minimum and maximum values), the *ERA6190* run (blue points), the *ERCLM* run (red points) and the GRDC observations (black points).

influenced Asiatic rivers Brahmaputra, Indus, Mekong and Yangtze, while for the Ganges and Yellow River discharges are on the high side.

To investigate the possible influence of water use, we made the simplifying assumption that water demand equals water use. In reality water demand will, especially in drier regions, be higher than water use. Water demand is estimated on a grid of 0.5 degrees. It is the sum of estimated industrial, agricultural and domestic water demand (Wada et al., 2008). Calculated discharge is compared with the sum of observed

discharge and water demand, which approximates natural discharge. Figure 2b shows that for the Murray the deviations between observed and calculated discharge can partly be assigned to the lack of inclusion of water use. To a lesser extent this also applies for the Danube, Ganges, Yellow River and Rhine.

Besides mean discharge statistics, the positions of discharge extremes (e.g. the positions of $Q90$ and $Q10$ values relative to $\overline{Q}$) have been compared for those basins where daily GRDC time-series were available, see Figs. 3
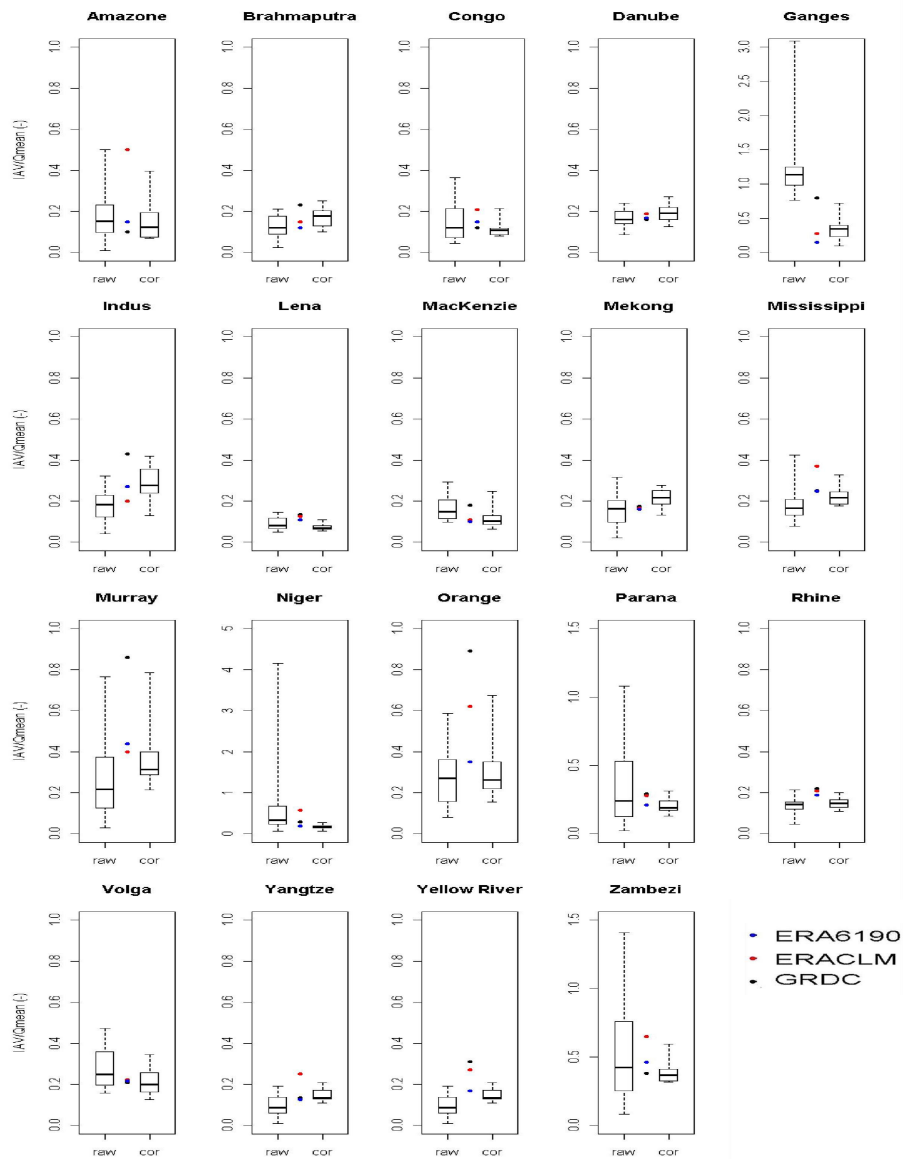
Fig. 5. Thirty year average relative inter-annual variability values (IAV$/\overline{Q}$(–)) for the ensemble of non-bias corrected GCM runs (raw) and bias-corrected GCM runs (cor; boxes are drawn between the quartiles and whiskers represent minimum and maximum values), the *ERA6190* run (blue points), the *ERCLM* run (red points) and the GRDC observations (black points).

and 4 where the blue dots represent the results of the *ERA6190* run and black dots are GRDC values. In this analysis we will focus on the catchments with the largest deviations.

The relative position of the $Q90$ discharge is higher in the model results than in the GRDC observations for the drier basins; Orange and Zambezi, the position of the $Q10$ discharge is too high as well for the Niger, Orange and Murray, but too low for the Zambezi. For the Ganges the relative $Q90$ discharge is too low and for both the Lena and the MacKenzie, the relative $Q90$ is underestimated as well due to the relative low snow melt driven discharge rise in

spring, while the position of $Q10$ flows are overestimated due to a higher baseflow during winter.

Deviations in relative inter-annual variability are small for European catchments (Fig. 5). IAV is underestimated in the drier catchments; Murray, Orange and Niger, in the Arctic catchments and in the Monsoon influenced basins; Brahmaputra, Indus, Ganges and Yellow River. For the remaining catchments differences are small.

Month of peak discharge occurrence is modeled well for the Arctic and most monsoon influenced basins (Fig. 6). Large deviations in timing are present for all dry catchments. For the Rhine the discharge peak is modeled in April and
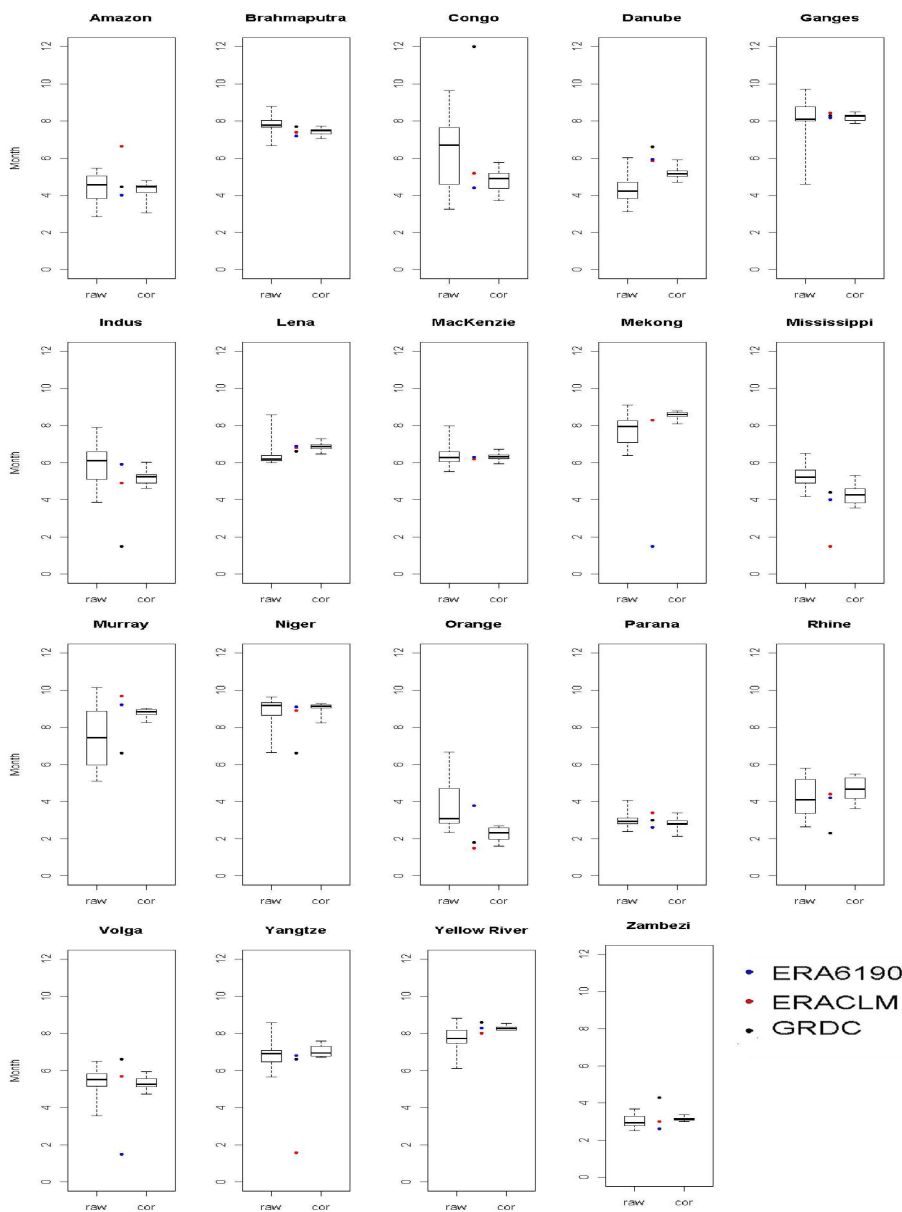
Fig. 6. Thirty year average month of peak discharge occurrence (*Qpeak* (–)) for the ensemble of non-bias corrected GCM runs (raw) and bias-corrected GCM runs (cor; boxes are drawn between the quartiles and whiskers represent minimum and maximum values), the *ERA6190* run (blue points), the *ERCLM* run (red points) and the GRDC observations (black points).

observed in February due to a too late snow melt driven discharge peak. For both the Amazon and Congo the modeled regime shows a discharge rise in March and April which is not present in the observed discharges. For the Congo Zaitchik et al. (2010) found a similar shift in peak timing, which they assign to the complexity of the bimodal runoff regime and the poor quality of the GRDC gauge data for the Congo.

3.1.3 Regime curves

Finally, 30 year average modeled hydrological regimes are compared with observed regimes (GRDC) for all catchments, see the blue (*ERA6190*) and black (GRDC) lines in Fig. 7 blue line. The plots show that difference between the modeled and observed regime is especially large for the MacKenzie as mentioned before. Although for the Lena annual average modeled discharge is close to observed, the regime curve shows deviations as well. The observed
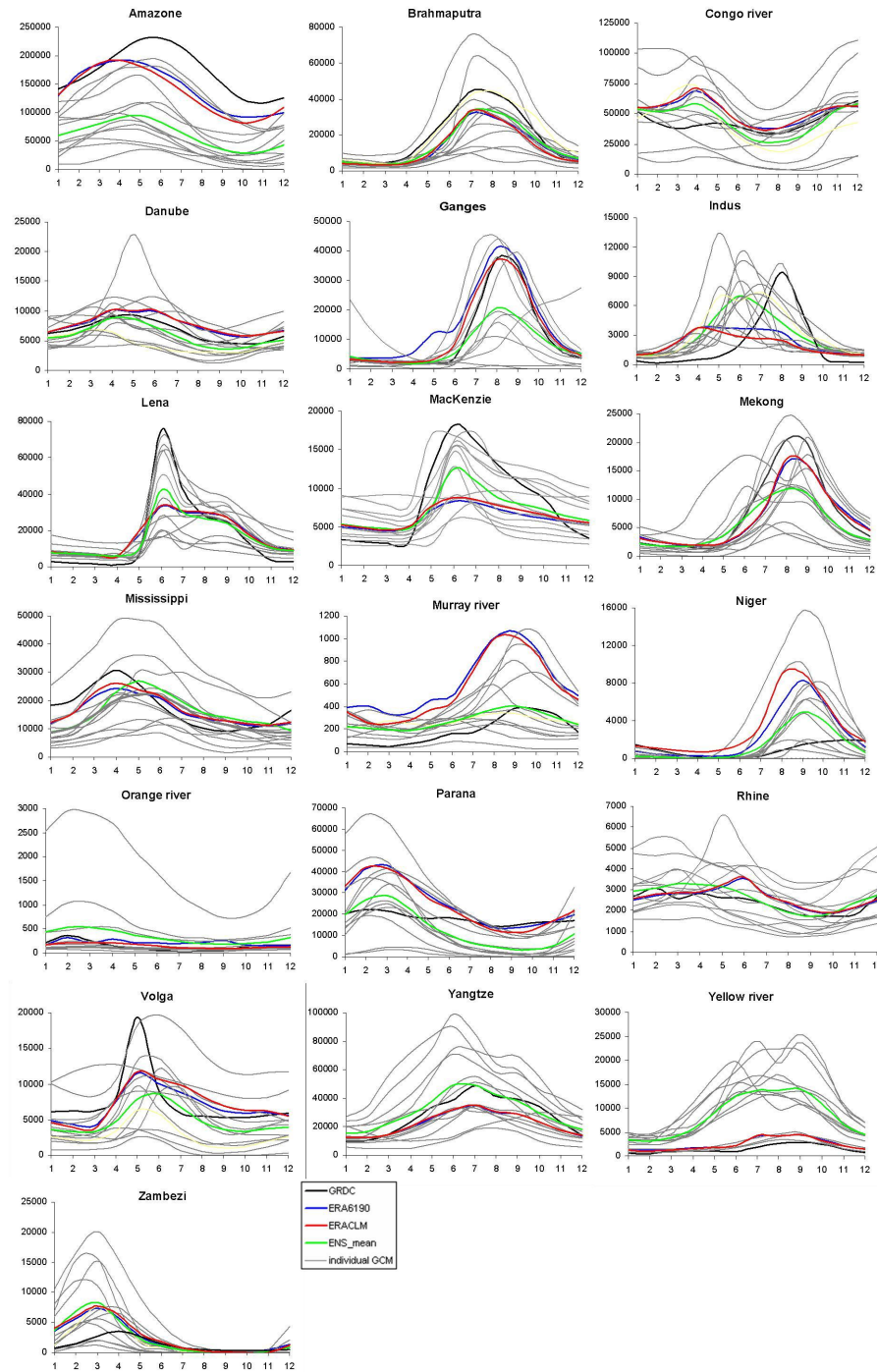
Fig. 7. Hydrological regimes for all 19 catchments derived from discharges calculated for the period 1961–1990, with the average monthly discharge (m^3/s) on the y-axis and the month numbers on the x-axis. Black line is annual mean observed discharge (GRDC), grey lines are PCR-GLOBWB based on individual *non bias-corrected* GCM datasets, green line is ensemble mean result, blue line is derived from PCR-GLOBWB run forced with the ERA6190 dataset and red line is derived from PCR-GLOBWB run forced with the ERACLM dataset.

snowmelt driven discharge peak is steeper than modeled and modeled baseflow is too high. Modeled discharge for the Murray, Niger and Zambezi is higher throughout the

year. The regime curves of monsoon influenced rivers are modeled relatively well, except for the regime curve of the Indus, where the snow and glacier melt driven discharge
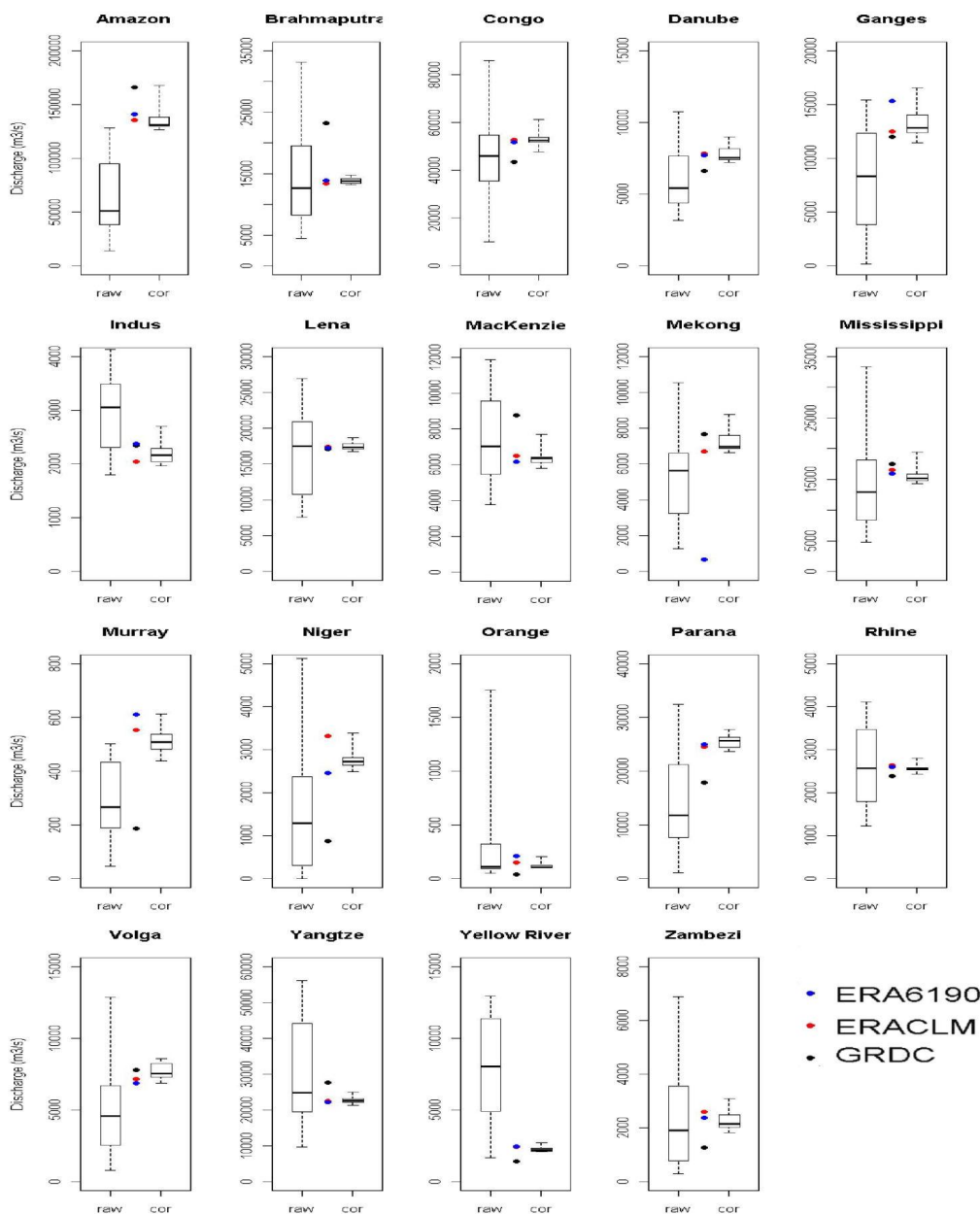
Fig. 8. Boxplots obtained form the 30-year average mean discharges of the ensemble of non-bias corrected GCM runs (raw) and bias-corrected GCM runs (cor; boxes are drawn between the quartiles and whiskers represent minimum and maximum values). The points show 30-year average mean discharges of the *ERA6190* run (blue), *ERACLM* run (red) and GRDC observations (GRDC).

rise is too early in spring and does not coincide with the monsoon influenced discharge peak, as in the observed regime (Immerzeel et al., 2010). For the Parana observed river flow is more constant through time, because of the presence of large reservoirs. The NS coefficients (Table 4) show that the modelled regime curves show large deviations, due to biases in absolute discharge quantities, for the arid catchments (Murray, Orange, Zambezi and Niger) and for

the Yellow river. Furthermore, negative NS coefficients are obtained for the Parana and Congo where the shape of the modeled regime deviates from the observed regime. NS coefficients of more than 0.5 are obtained for the Brahmaputra, Ganges, Mekong and Yangtze and for the Mississippi and Lena.
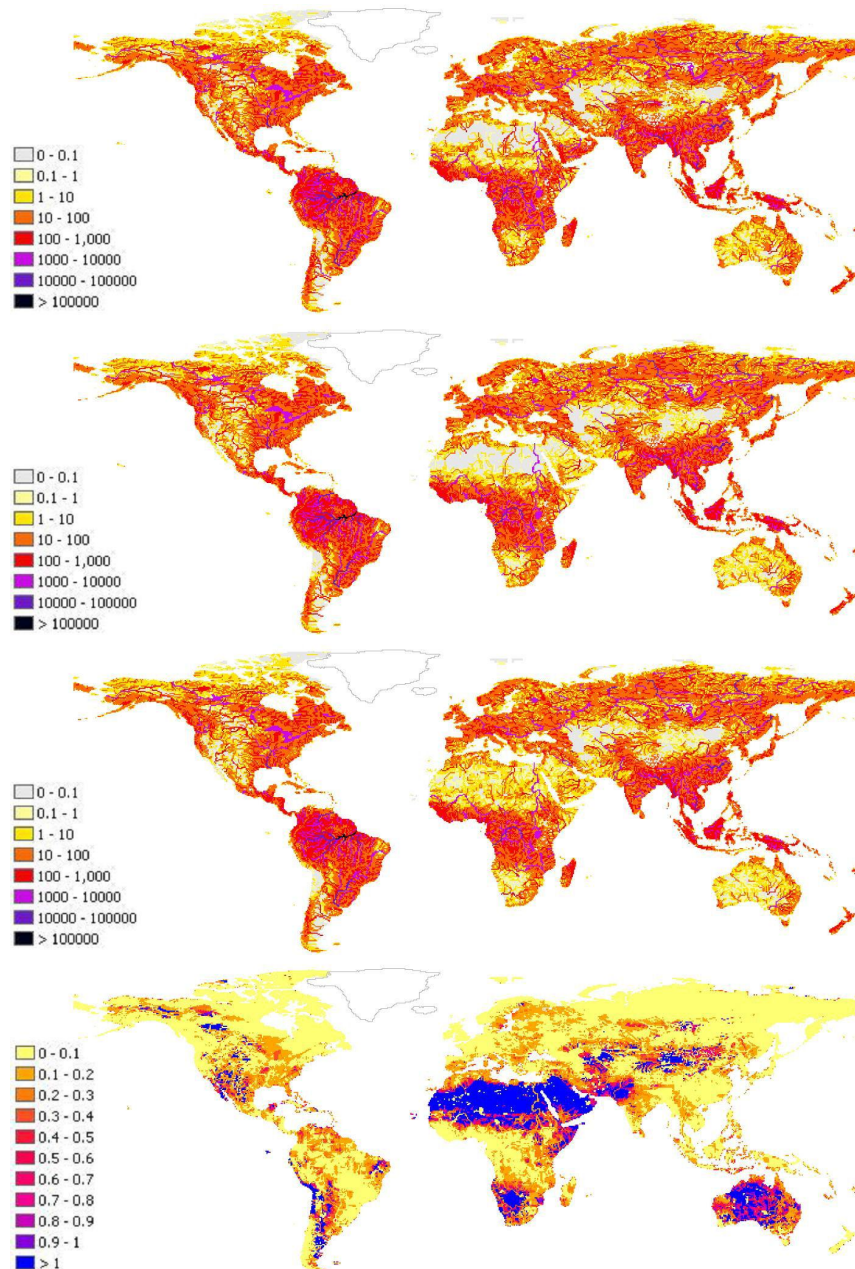
Fig. 9. Thirty year average mean discharge (m³/s) calculated from **(a)** the PCR-GLOBWB run forced with the *ERA6190* dataset, **(b)** the PCR-GLOBWB run forced with the *ERACLM* dataset, **(c)** the ensemble mean discharge results of the 12 GCM based PCR-GLOBWB runs and **(d)** the CV (–) of the thirty year average discharges of the ensemble of GCMs.

3.2 Evaluation of the bias-correction method using the *ERACLM* dataset

The results of the *ERACLM* run are compared with the results of the *ERA6190* run and observed GRDC discharges in order to assess the influence of the bias-correction method applied to the GCM datasets.

3.2.1 Continental discharges

The continental discharges of the *ERACLM* and the *ERA6190* run in Table 3 are comparable in size. Globally the difference is only 3%. The difference is especially small for Europe (0.07%). For all continents, except Europe, discharge is lower in the *ERACLM* run. The *ERACLM* run underestimates the global average modeled runoff obtained from previous studies by 4%.
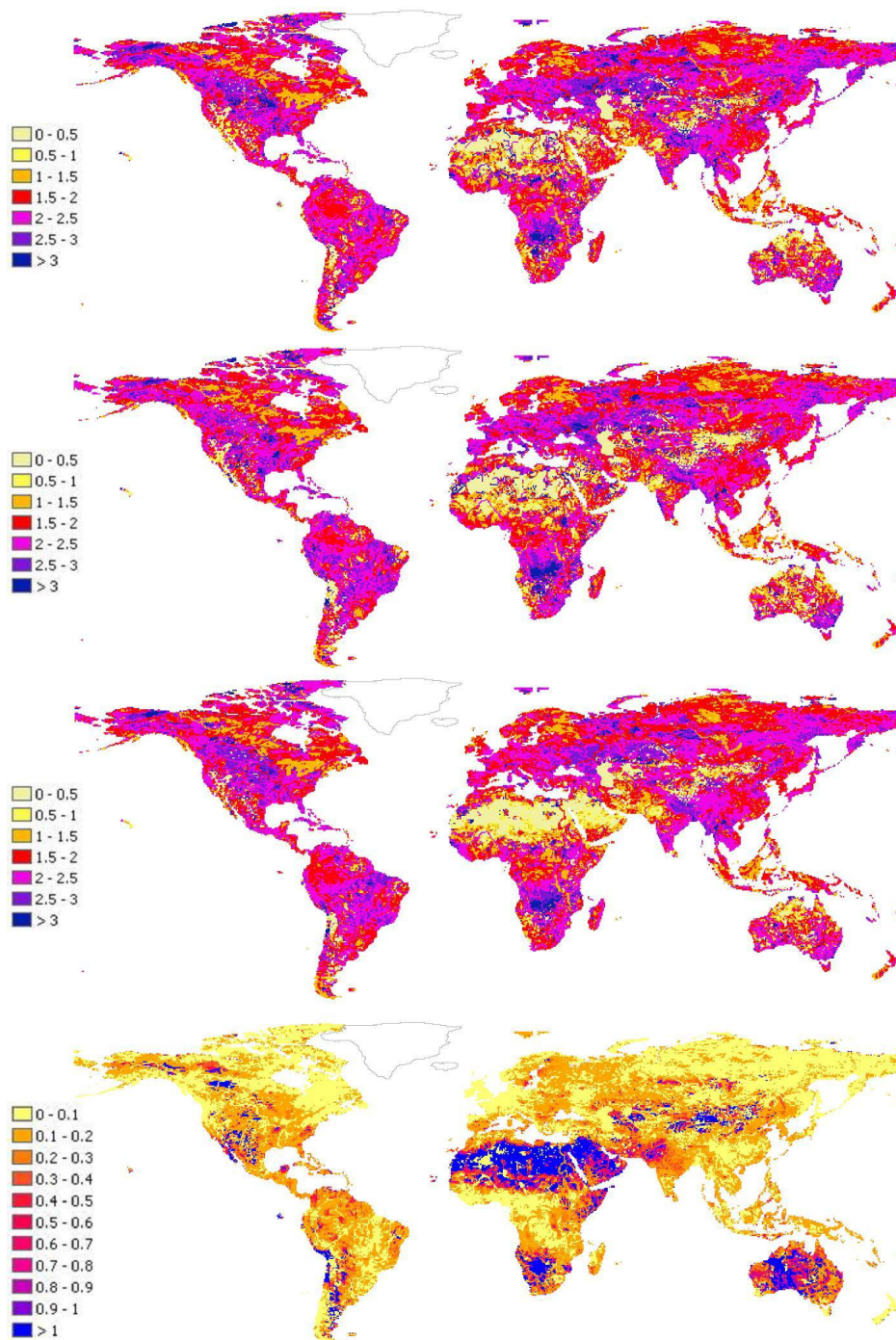
Fig. 10. Thirty year average relative $Q90$ values: $Q90/\overline{Q}$ (–) calculated from **(a)** the PCR-GLOBWB run forced with the *ERA6190* dataset, **(b)** the PCR-GLOBWB run forced with the *ERACLM* dataset, **(c)** the ensemble mean discharge results of the 12 GCM based PCR-GLOBWB runs and **(d)** the CV (–) of the thirty year average $Q90$ of the ensemble of GCMs.

3.2.2 Discharge statistics

Differences in results of the *ERACLM* and *ERA6190* are small, especially for mean discharge (see Fig. 8). Deviations of both the GCM runs and the ERA-40 based runs from the observed GRDC discharge statistics are larger (Figs. 3–6). High and low-flow values are more extreme in the *ERACLM* run than in the *ERA6190* run; relative $Q90$ values are higher for 14 out of 19 basins and relative $Q10$ values are lower for
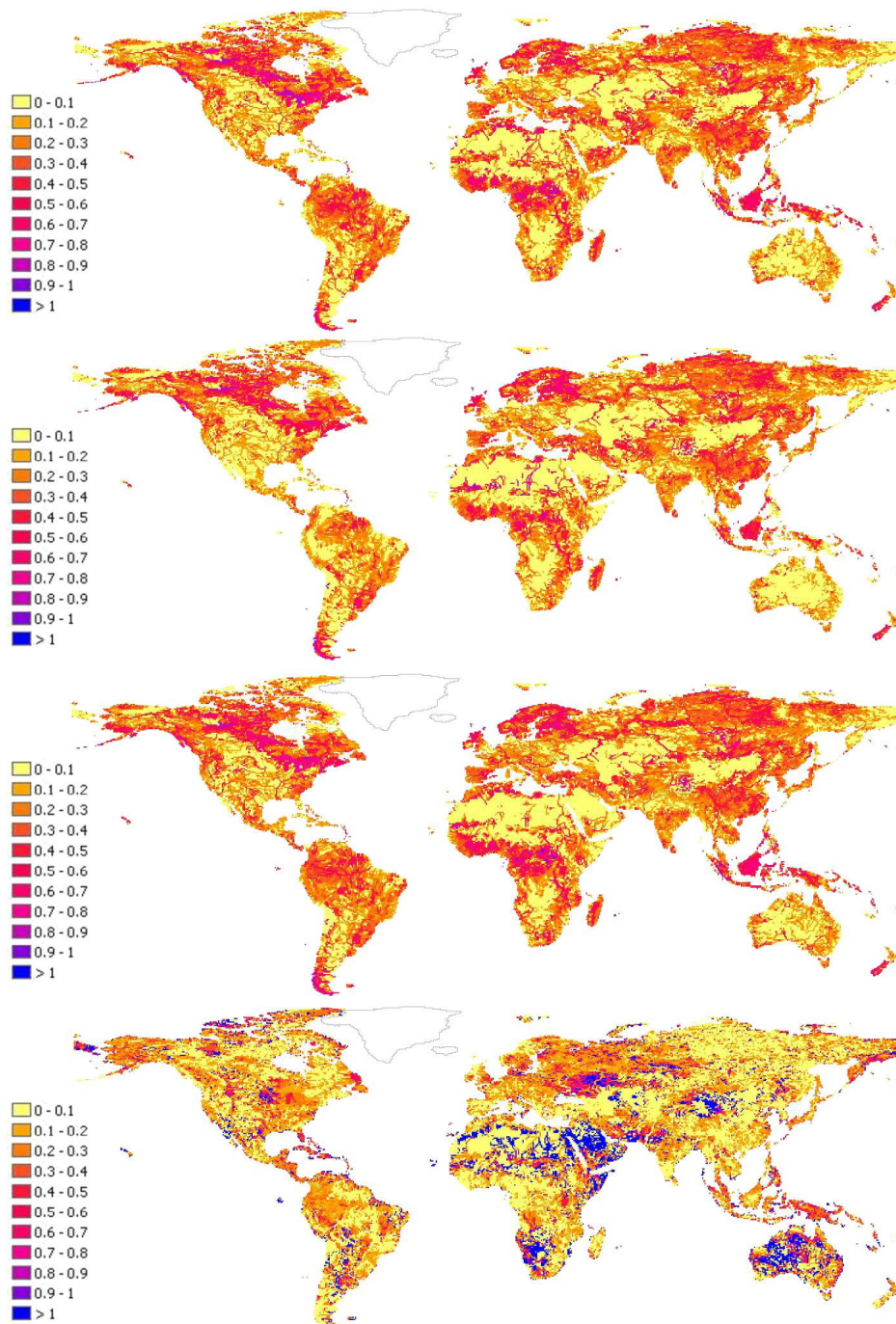
Fig. 11. Thirty year average relative Q10 values: $Q10/\overline{Q}$ (–) calculated from **(a)** the PCR-GLOBWB run forced with the *ERA6190* dataset, **(b)** the PCR-GLOBWB run forced with the *ERACLM* dataset, **(c)** the ensemble mean discharge results of the 12 GCM based PCR-GLOBWB runs and **(d)** the CV (–) of the thirty year average Q10 of the ensemble of GCMs.

14 out of 19 basins. Differences between the runs in both relative Q10 and relative Q90 are largest for Amazon and Niger.

Differences in relative IAV are more pronounced. For 16 out of 19 basins the *ERA6190* results have a lower IAV. The IAV of the *ERA6190* is prescribed by the CRU TS

2.1 timeseries, whereas the IAV of the *ERACLM* dataset is prescribed by the ERA-40 re-analysis dataset. The IAV of ERA-40 precipitation is known to be too high after 1972 due to gradual increasing inclusion of satellite observations in the system (Chen and Bosilovich, 2007; Hagemann et al., 2005; Troccoli and Kålberg, 2004), resulting in genrally higher IAV
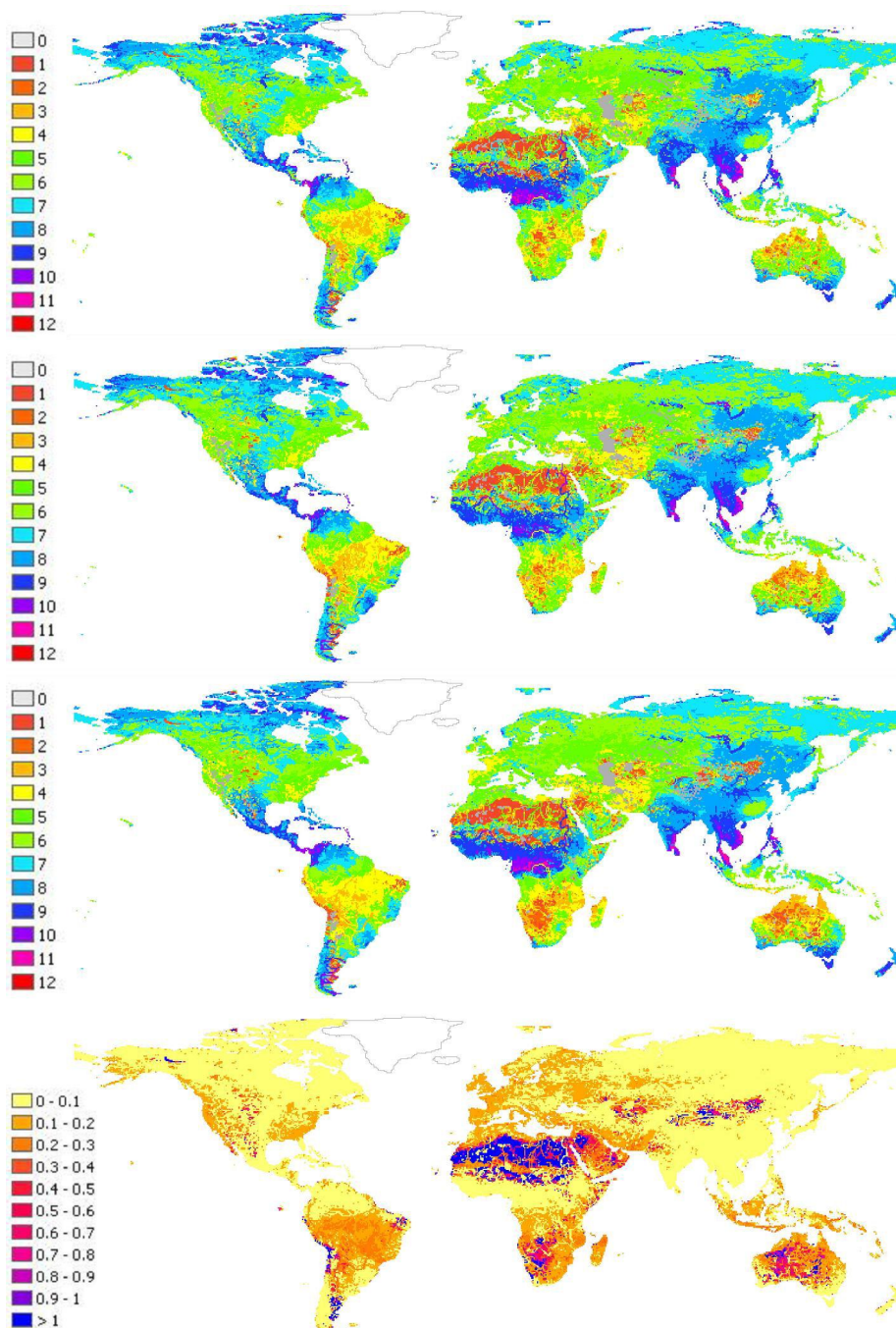
Fig. 12. Thirty year average timing of peak discharge (m^3/s) calculated from **(a)** the PCR-GLOBWB run forced with the *ERA6190* dataset, **(b)** the PCR-GLOBWB run forced with the *ERACLM* dataset, **(c)** the ensemble mean discharge results of the 12 GCM based PCR-GLOBWB runs and **(d)** the CV (–) of the thirty year average timing of peak discharge of the ensemble of GCMs.

values in the *ERACLM* than in the *ERA6190* run for 16 out of 19 basins. For 10 out of 19 basins (mainly located in monsoon and arctic regions) the IAV of the *ERACLM* dataset is closer to the IAV of the GRDC than the *ERA6190* dataset. For these basins IAV is underestimated by both runs and the high ERA-40 IAV compensates for this underestimation.

3.2.3 Regime curves

Timing of peak discharge and regime curves are comparable for the two forcing datasets, see for comparison the blue (*ERA6190*) and red (*ERACLM*) lines in Fig. 7. Exceptions are (1) the Indus, where the *ERA6190* dataset results in a later
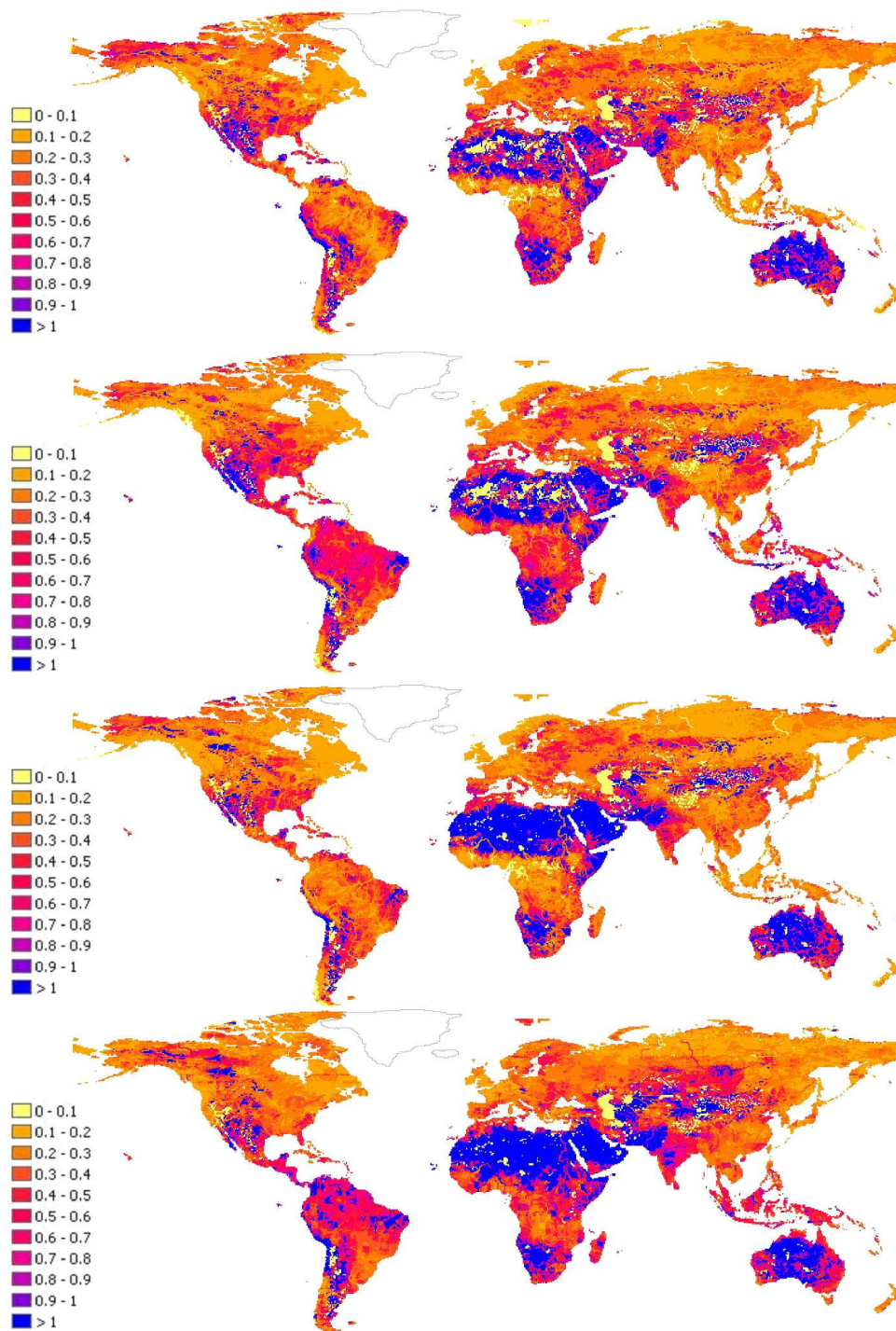
Fig. 13. Thirty year average relative inter-annual variability: $\text{IAV}/\overline{Q}$ (–) calculated from **(a)** the PCR-GLOBWB run forced with the *ERA6190* dataset, **(b)** the PCR-GLOBWB run forced with the *ERACLM* dataset, **(c)** the ensemble mean discharge results of the 12 GCM based PCR-GLOBWB runs and **(d)** the CV (–) of the IAV of the ensemble of GCMs.

timing of peak discharge that coincides with the monsoon instead of snow melt in spring, and (2) the Orange, where the average *ERA6190* regime is more constant throughout the year and the seasonality of the regime, with a low end of summer baseflow, is less pronounced.
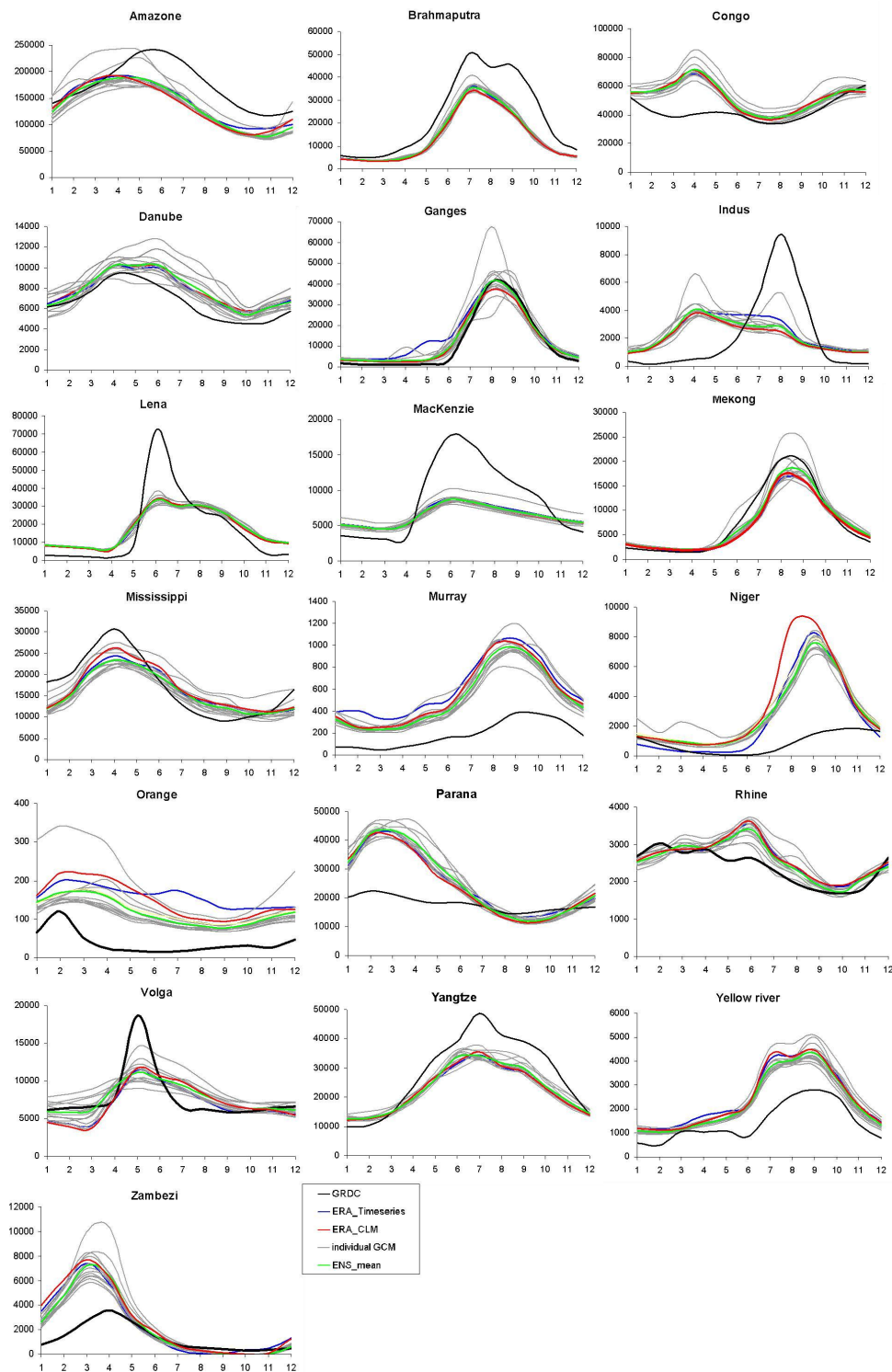
Fig. 14. Hydrological regimes for all 19 catchments derived from discharges calculated for the period 1961–1990, with the average monthly discharge (m³/s) on the y-axis and the month numbers on the x-axis. Black line is annual mean observed discharge (GRDC), grey lines are PCR-GLOBWB based on individual *bias-corrected* GCM datasets, green line is ensemble mean result, blue line is derived from PCR-GLOBWB run forced with the ERA6190 dataset and red line is derived from PCR-GLOBWB run forced with the ERACLM dataset.

Table 4. Nash-Sutcliffe coefficients for the regimes curves (30-year average monthly mean discharges) of the *ERACLM*, *ERA6190*, ensemble mean non bias-corrected (RAW) and ensemble mean bias-corrected (BiasCor) GCM runs relative to the GRDC regime.

NS	RAW	BiasCor	ERA6190	ERACLM
Amazon	−5.64	−0.08	0.43	0.50
Brahmaputra	−2.10	0.59	0.55	0.59
Congo river	−0.07	−2.01	−1.95	−1.48
Danube	0.87	0.33	0.38	0.50
Ganges	0.65	0.97	0.96	0.86
Indus	0.12	0.05	0.02	0.15
Lena	0.72	0.61	0.62	0.62
MacKenzie	0.58	0.27	0.30	0.21
Mekong	0.68	0.94	0.89	0.87
Mississippi	0.06	0.63	0.71	0.66
Murray	0.03	−8.34	−10.53	−13.26
Niger	−4.61	−14.98	−28.91	−16.78
Orange	−106.11	−7.86	−15.92	−17.75
Parana	−9.84	−23.61	−19.36	−19.44
Rhine	0.54	0.39	0.15	0.25
Volga	−0.05	0.50	0.43	0.50
Yangtze	0.87	0.68	0.66	0.63
Yellow River	−88.06	−0.70	−1.25	−1.27
Zambezi	−4.05	−2.17	−3.79	−2.86

3.2.4 Spatial statistics

The maps with $\overline{Q}$, relative $Q90$ and relative $Q10$ (Figs. 9, 10 and 11 upper two maps) show only a few differences between the *ERACLM* and *ERA6190* runs. Both the relative $Q90$ and $\overline{Q}$ are lower in western Australia for *ERACLM*. The south of Africa is relatively wet in the *ERA6190* run and relative $Q90$ values are higher in the *ERA6190* run for rivers and streams west of the Sahara. Mean discharge is slightly lower in the dry Western USA. Spatial patterns of timing of peak discharge are similar for the two runs (Fig. 12). Differences in IAV are also small (Fig. 13). Apparent are the relative high IAV values in the *ERACLM* run in Oman, Yemen and the Northern Sahel, where precipitation variability is influenced by the African Monsoon (Cook and Vizy, 2006). In dry regions the relative IAV values are the highest but of little importance as they are associated with mean discharge values that are close to zero.

3.3 Evaluation of discharge variability obtained with bias-corrected GCM datasets

3.3.1 Continental discharges

Differences in continental discharges derived from the *ERA6190, ERACLM* and ensemble of GCM runs are small (Table 3) as expected. The values of the GCM derived discharges are closer to the *ERACLM* discharges than to the *ERA6190*, globally the difference between *ERACLM* and the mean GCM discharge is less than 2%.

3.3.2 Discharge statistics

After bias-correction, deviations of the GCM ensemble discharge statistics from the GRDC observed discharge statistics have decreased for the majority of catchments (Fig. 8). The bias-correction decreased the deviation for $\overline{Q}$ for 11 out of 19 basins, for relative $Q90$ for 13 out of 19 basins and for the IAV for 12 out of 19 basins. However, for relative $Q10$ the deviation increased for 11 out of 19 basins and for $Qpeak$ the deviation increased for 10 out of 19 basins.

Furthermore, the spread between the individual GCM results decreased. The CV of the ensemble of GCMs decreased for $\overline{Q}$ and $Q10$ for all catchments, for $Qpeak$ and IAV for 18 out of 19 catchments and for relative $Q90$ for 12 out of 19 catchments, On average, after bias-correction, the CV of $\overline{Q}$ is lower than the CV of $Q10$ for 11 out of 19 basins and lower than the CV of relative $Q90$ for 14 out of 19 basins (Table 5). However, within the ensemble of bias-corrected GCM results, several GCMs still show deviations for $\overline{Q}$, especially for the Orange and Zambezi. The highest CV values are obtained for IAV. Low CV values and relatively small outliers of $Q10$, $Q90$ and $\overline{Q}$ are obtained for the Rhine, Lena, Yangtze and Brahmaputra. For the Lena and Brahmaputra the range between minimum and maximum values for IAV and timing of peak are also relatively small. This indicates the bias-correction reduced the spread between the GCMs best in these catchments. For the dry catchments; Orange and Zambezi, CV values and the range between minimum and maximum GCM results for $Q90$, $Q10$ and $\overline{Q}$ remain large. CV values for timing of peak discharge occurrence are small, especially for the Arctic rivers and the Monsoon influenced rivers Brahmaputra, Ganges, Mekong and Yellow river.

The results of the ensemble of bias-corrected GCMs are closer to the GRDC statistics than the results of the ERA-40 based runs for the majority of basins for $\overline{Q}$, $Q10$ and $Qpeak$ (Figs. 3–6). For relative $Q90$ the results of the *ERACLM* run are closer to the discharge observations for 11 out of 19 basins and the IAV of both ERA-40 based runs is closer to the GRDC values for more than half of the basins (*ERACLM* for 10 and *ERA6190* for 11 basins). On average the IAV is higher in discharge results of hydrological model runs based on the

Table 5. Ensemble coefficients of variation of the GCM results for the statistics of interest.

CV	Qmean		Q90		Q10		IAV		Qpeak	
	raw[*]	cor[*]	raw	cor	raw	cor	raw	cor	raw	cor
Amazon	0.56	0.08	0.21	0.13	0.50	0.12	0.83	0.72	0.17	0.12
Brahmaputra	0.54	0.04	0.10	0.05	0.26	0.05	0.71	0.29	0.06	0.02
Congo river	0.45	0.06	0.14	0.08	0.34	0.06	0.65	0.38	0.30	0.12
Danube	0.35	0.08	0.08	0.11	0.18	0.05	0.32	0.26	0.20	0.06
Ganges	0.59	0.11	0.16	0.10	0.51	0.11	0.88	0.56	0.15	0.02
Indus	0.24	0.10	0.11	0.08	0.19	0.06	0.46	0.37	0.17	0.07
Lena	0.34	0.03	0.12	0.04	0.30	0.04	0.55	0.24	0.11	0.03
MacKenzie	0.33	0.07	0.07	0.10	0.12	0.04	0.43	0.50	0.09	0.03
Mekong	0.51	0.08	0.12	0.08	0.39	0.06	0.68	0.21	0.11	0.02
Mississippi	0.53	0.09	0.08	0.10	0.19	0.05	1.16	0.28	0.12	0.11
Murray	0.46	0.09	0.09	0.13	0.27	0.13	0.92	0.49	0.21	0.02
Niger	0.95	0.08	0.47	0.07	3.32	0.14	0.96	0.38	0.09	0.03
Orange	1.45	0.23	0.11	0.37	0.35	0.10	1.08	0.80	0.34	0.15
Parana	0.65	0.05	0.19	0.09	0.91	0.07	0.97	0.27	0.14	0.10
Rhine	0.35	0.04	0.08	0.06	0.20	0.07	0.44	0.19	0.23	0.14
Volga	0.64	0.07	0.19	0.10	0.50	0.12	0.59	0.34	0.14	0.14
Yangtze	0.49	0.04	0.07	0.06	0.12	0.06	0.68	0.20	0.10	0.05
Yellow River	0.48	0.07	0.06	0.07	0.20	0.05	0.54	0.28	0.09	0.01
Zambezi	0.78	0.16	0.10	0.15	3.32	2.56	0.73	0.26	0.12	0.03

(*) raw refers to the statistics of the non bias-corrected GCM runs and cor refers to the statistics of the bias-corrected GCM runs.

forcing datasets where the IAV is determined by the model generated meteorological datasets (e.g. the datasets obtained by bias-correction).

3.3.3 Regime curves

Overall the simple monthly bias-correction reduces the differences between the regimes derived from the GCM based runs (compare Fig. 7 and Fig. 14, where grey lines represent individual GCM runs and the green line is ensemble mean regime). After bias-correction, the resemblance of the GCM runs is large for rivers with a strong seasonal pattern, like the monsoon influenced and Arctic rivers. For other catchments that are influenced by both rainfall and snowmelt, like the Amazon, Rhine, Volga, Danube and Indus, differences in regime pattern are larger.

After bias-correction the NS coefficient of the regime curves decreased for 10 out of 19 basins; for several basins, such as the Indus, MacKenzie, Murray, Niger and Parana, the mean regime curves of the non bias-corrected GCMs is much

closer to the GRDC regime than the mean regime curves of the bias-corrected GCMs. Still, the NS coefficient of the GCM ensemble mean is higher for 12 out of 19 basins than the NS coefficients derived for the regimes of both ERA-40 based runs.

For basins in the drier regions of the world, such as the Murray, Zambezi and Orange river, the GCM ensemble mean regime curve is slightly lower and closer to the GRDC regime than the regime curves of both ERA-40 runs. For the remaining basins the deviations of the GCM ensemble mean regime from the GRDC regime are comparable those of the *ERACLM* and *ERA6190* regimes.

3.3.4 Spatial statistics

Differences in ensemble mean spatial discharge patterns from both ERA-40 runs are small. $\overline{Q}$ is relatively high in the Sahara resulting in lower relative Q90 values. In the Australian desert GCM mean discharge is relatively low and in the Southwestern USA the region with lower discharge

values is slightly more extended. Ensemble uncertainties (CV) are large for arid African regions, the arid southwestern USA and for the desert region of Australia where the CV of precipitation of the ensemble of GCMs is large as well (Johnson and Sharma, 2009). High CV values are also obtained for mountainous regions, like the Andes ridge and parts of Asia, where the low GCM resolutions are unable to capture the high spatial heterogeneity (Viviroli et al., 2010). CV values are higher, and regions with high CV values are more extended, for relative $Q90$ and $Q10$ than for $\overline{Q}$. For all statistics the maps obtained from the ensemble of GCMs resemble the maps of the *ERACLM* run most, except for the Northern Sahel.

4 Discussion

The analysis in this study was performed in three steps, i.e. to quantify the influence of both the hydrological model deficiencies and the bias-correction method before evaluating the discharge variability derived from bias-corrected GCM datasets. In this case PCR-GLOBWB was used as a means to assess the hydrological discharge variability obtained from meteorological GCM datasets but the use of any other macro-scale hydrological model probably would have resulted only in minor differences. As assessed by the validation exercise in the first step of this study, the use of a hydrological model introduces uncertainty and bias in the discharge that propagates in the results derived from the bias-corrected GCM datasets. Although the model errors found here are specific for PCR-GLOBWB, they are not unlike those found in other global model studies (e.g., Döll et al., 2003; Nijssen et al., 2001; Arnell, 1999). These errors can be reduced when the model is tuned or calibrated but this is only possible at the expense of a loss of its physical basis and the result is highly depended on data availability (Beven, 2001). Calibration would also amend for errors in the meteorological input, thus compromising the applicability of the model to different GCMs.

The performance assessment is hampered by the lack of a perfect representation of the historical meteorology. By downscaling the CRU TS 2.1 timeseries to daily values with the ERA-40 reanalysis we created a 'best guess'. However, biases present in both the CRU and ERA-40 datasets will also be included in this new dataset. Nonetheless, PCR-GLOBWB, forced with this dataset, reproduces global runoff quantities and regimes relatively well. For the European continent and the rivers Danube, Rhine and Volga mean discharges and inter-annual variability are relatively close to the GRDC observed values. Differences in total continental discharges from previous studies are small. Globally the difference from the average of existing model studies is less than 1% (Table 3). Still, for the African continent as well as the Murray river PCR-GLOBWB overestimates discharge. Regime curves are also reproduced

well for Monsoon influenced rivers but mean discharge and inter-annual variability are underestimated for most of the Asian monsoon rivers (Fig. 7).

Other than through model deficiencies, the biases between modeled and observed discharge statistics are caused by errors in 1) the observed discharges, 2) the ERA-40 dataset and 3) the CRU TS 2.1 dataset, which is known to be inaccurate for parts of Africa and suffers from undercatch in snow dominated areas (Fiedler and Döll, 2007). By using the CRU dataset as reference for the bias-correction, deviations present in the CRU set are introduced in the bias-corrected GCM datasets as well.

The influence of the bias-correction method is obtained form the difference in results between the *ERA6190* and the *ERACLM* runs. Differences between these two runs are small, overall relative IAV and $Q90$ values are higher in the *ERACLM* run (Fig. 3 and 5) and $\overline{Q}$ values are lower (Fig. 8). The IAV of ERA-40 precipitation is relatively high after 1972 when satellite observations are gradually included in the system (Chen and Bosilovich, 2007). This high precipitation IAV compensates for the underestimation of observed IAV obtained with the *ERA6190* run.

The average annual continental runoff is lower for the *ERACLM* than for the *ERA6190* run, especially over Africa (Table 3). For the *ERACLM* dataset the bias-correction is not performed on a year-by-year basis, i.e. adjusting all monthly precipitation, temperature and evaporation values to the monthly observations from the CRU for the specific year. Instead, in the *ERACLM* (and the GCM) dataset the 30-year average annual mean precipitation, temperature and evaporation quantities are equalized to the 30-year average observed values of the CRU observations. The inter-annual variability for potential evaporation is relatively small compared to precipitation. Therefore, the correction of the 30-year mean potential evaporation has a similar increasing or decreasing effect on all 30 years. Yet, inter-annual variability for precipitation is larger and the multiplicative correction results in larger absolute changes for wet years. Herewith consistency between precipitation and evapotranspiration is lost over the wetter periods. Especially in wet years, with precipitation overestimations in the ERA-40 dataset, the correction results in relatively large precipitation decreases, while potential evaporation is only slightly adjusted. Here the ERACLM run will yield less effective precipitation with a consequent decrease in discharge.

Bias-correction decreases the deviations of GCM based model results from the observed discharges for $\overline{Q}$, relative $Q90$ and IAV for most basins, but increases deviations for Qpeak and relative $Q10$for most basins. After bias-correction, continental discharges obtained from the ensemble of GCMs are close to the *ERACLM* results and lower than discharges modeled with the *ERA6190* dataset (Table 3). The results of the bias-corrected GCM runs are on average closer to the GRDC data for $\overline{Q}$, $Q10$ and *Qpeak*,

while for $Q90$ the *ERA6190* run is closer to the GRDC values and for IAV both ERA-40 based runs are closer to the GRDC for a majority of the basins.

Application of the bias-correction method to the GCM datasets decreases the spread between the GCM ensemble members and lower CV values are obtained for the majority of basins for all discharge statistics. The largest ensemble CV values remain for drier catchments. Furthermore, bias-correction has less influence on extremes than on mean discharge; CV values are lower for $\overline{Q}$ than for $Q10$ and $Q90$ (Table 5). The CV of the inter-annual variability decreased for 17 out of 19 basins. Still, the CV values for IAV remain the highest of all statistics.

Bias-correction decreased the spread between the individual GCM derived regime curves. Especially, the resemblance of regimes for the monsoon influenced and Arctic rivers is large (Fig. 14). The NS coefficient decreased for 10 out of 19 basins after bias-correction. Yet, it can be questioned, whether this decrease is a result of better performance of the non bias-corrected datasets or a result of biases in the not corrected GCM datasets that compensate for inadequacies of the hydrological model (Fig. 7).

5   Conclusions

In order to evaluate the usability of bias-corrected GCM datasets in hydrological impact studies with a focus on changes in extremes and variability, we calculated hydrological regime curves and various discharge statistics for the period 1961–1990. We did this by simulating discharge with PCR-GLOBWB for a selection of 12 GCMs and comparison of the results with observed GRDC discharge data and two reference runs based on the ERA-40 and CRU TS 2.1 datasets.

As expected, the spread between the results of the different GCM based runs decreases after bias-correction, especially for rivers with a strong seasonal pattern. Overall, the bias-correction decreased the deviation between GCM ensemble mean and GRDC observed statistics for $\overline{Q}$, $Q90$ and IAV. However, the bias-corrected datasets result in slightly lower mean annual discharge amounts and for arid and mountainous basins the method performs less well.

Observed discharge statistics ($Q10$, $Q90$ and $\overline{Q}$) are better approximated by the bias-corrected GCM datasets than inter-annual variability (IAV) and agreement amongst models is higher for mean discharge and timing than for IAV. This suggests that GCMs are mostly useful in global hydrological impact studies in which persistence is of less consequence (e.g. in case of flood studies rather than those on droughts).

Furthermore, the bias-correction influences mean discharges more than discharge extremes, which has the positive consequence that changes in daily rainfall distribution and subsequent changes in discharge extremes will also be preserved when the bias-correction method is applied to future GCM datasets. However it also shows that agreement between GCMs remains small for discharge extremes

Finally, it can be concluded that because of the large deviations between observed and simulated discharge in which both errors in climate forcing, model structure and to a lesser extent observations are accumulated, it is advisable not to work with absolute discharge values for the derivation of future discharge projections, but rather calculate relative change by dividing the absolute change by the discharge calculated for the control experiment.

References

Alcamo, J., Henrichs, T., and Rösch, T.: World water in 2025: Global modeling and scenario analysis for the world commission on water for the 21st century, Kassel World Water Series Report No. 2, Center for Environmental Systems Research, University of Kassel, Germany, 2000.

Alcamo, J. and Henrichs T.: Critical regions: A model-based estimation of world water resources sensitive to global changes, Aq. Sci., 64, 352–263, 2002.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., and Siebert, S.: Development and testing of the WaterGAP 2 global model of water use and availability, Hydrol. Sci. J., 48(3), 317–337, 2003.

Allan, R. P. and Soden, B. J.: Atmospheric warming and the amplification of precipitation extremes, Science, 321, 1481–1484, doi:10.1126/science.1160787, 2008.

Allen, R. G., Pereira, L. S., Raes, D., and Smith M.: Crop evapotranspiration: FAO Irrigation and drainage paper 56, FAO, Rome, Italy, 1998.

Allen, M. R. and Ingram W. J.: Constraints on the future changes in climate and the hydrological cycle, Nature, 419, 224–232, 2002.

Arnell, N. W.: A simple water balance model for the simulation of streamflow over a large geographic domain, J. Hydrol., 217, 314–355, 1999.

Arnell, N. W.: Effects of IPCC SRES* emissions scenarios on river runoff: a global perspective, Hydrol. Earth Syst. Sci., 7, 619–641, doi:10.5194/hess-7-619-2003, 2003.

Baumgartner, A. and Reichel, E.: The world water balance: Mean annual global, continental and maritime precipitation, evaporation and runoff, Elsevier, Amsterdam, 179 pp., 1975.

Beck, C., Grieser, J., and Rudolf, B.: A new monthly precipitation climatology for the global land areas for the period 1951 to 2000,

Climate Status Report 2004, pp. 181–190, German Weather Service, Offenbach, Germany, 2004.

Beven, K.: How far can we go in distributed hydrological modelling?, Hydrol. Earth Syst. Sci., 5, 1–12, doi:10.5194/hess-5-1-2001, 2001.

Bierkens, M. F. P. and van Beek, L. P.: Seasonal predictability of european discharge: NAO and hydrological response time, J Hydrometeorol, 10, 953–968, doi:10.1175/2009JHM10341, 2009.

Boorman, D. B. and Sefton, C. E. M.: Recognizing the Uncertainty in the Quantification of the Effects of Climate Change on Hydrological Response, Climatic Change, 35(4), 415–434, 1997.

Brouwer, C. and Heibloem, M.: Irrigation water management: Irrigation water needs, FAO, Rome, Italy, 1986.

Buytaert, W., Célleri, R., and Timbe, L.: Predicting climate change impacts on water resources in the tropical Andes: Effects of GCM uncertainty, Geophys. Res. Lett., 36, L07406, doi:10.1029/2008GL037048, 2009.

Chen, J. and Bosilovich, B. G.: Hydrological variability and trends in global reanalyses, 19th Conference on Climate Variability and Change, San Antonio, Texas, January 2007, JP4.4, 2007.

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.

Cook, K. H. and Vizy, E. K.: Coupled Model Simulations of the West African Monsoon System: Twentieth- and Twenty-First-Century Simulations, J. Climate, 19, 3681–3703, 2006.

Covey, C., AchutaRao, K. M., Cubasch, U., Jones, P., Lambert, S. J., Mann, M. E., Phillips T. J., and Taylor K. E.: An overview of results from the coupled model intercomparison project, Global Planet. Change., 37, 103–133, doi:10.1016/S0921-8181(02)00193-5, 2003.

Dai, A.: Precipiation characteristics in eighteen coupled climate models, Q. J. Am. Meteor. Soc., 19, 4605–4630, 2006.

Döll, P. and Lehner, B.: Validating of a new global 30-minute drainage direction map, J. Hydrol., 258, 214–231, 2002.

Döll, P., Kaspar, F., and Lehner, B.: A global hydrological model for deriving water availability indicators: model tuning and validation, J. Hydrol., 270, 105–134, 2003.

Fekete, B. M., Vörösmarty, C. J., and Grabs W.: Global, Composite Runoff Fields Based on Observed River Discharge and Simulated Water Balances, Technical Report 22, Global Runoff Data Centre, Koblenz, Germany, 2000.

Fekete, B. M., Vörösmarty, C. J., Roads, J. O., and Willmott, C. J.: Uncertainties in precipitation and their impacts on runoff estimates, J. Climate, 17, 294–304, 2004.

Fiedler, K. and Döll, P.: Global modelling continental water storage changes – sensitivity to different climate data sets, Adv. Geosci., 11, 63–68, 2007,
http://www.adv-geosci.net/11/63/2007/.

Fowler, H. J. and Kilsby, C.G.: Using regional climate model data to simulate historical and future ruver flows in northwest England, Climatic Change, 80, 337–367, 2007.

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., and Sitch, S.: Terrestrial vegetation and water balance – hydrological evaluation of a dynamic global vegetation model, J. Hydrol., 286, 249–270, 2004.

GRDC: Long Term Mean Annual freshwater Surface Water Fluxes into the World Oceans, Comparisons of GRDC freshwater flux estimate with literature, http://grdc.bafg.de/servlet/is/7083, 2004

GRDC: Major River Basins of the World / Global Runoff Data Centre, D – 56002, Federal Institute of Hydrology (BfG), Koblenz, Germany, 2007.

Hagemann, S., Arpe, K., and Bengtsson, L.: Validation of the hydrological cycle of ERA-40, Reports on earth system science, Max Planck Institute, Hamburg, Germany, 2005.

Hagemann, S., Göttel, H., Jacob, D., Lorenz, P., and Roeckner, E.: Improved regional scale processes reflected in projected hydrological changes over large European catchments, Clim. Dynam., 32, 767–781, doi:10.1007/s00382-008-0403-9, 2009.

Immerzeel, W. W., van Beek, L. P. H. and Bierkens, M. F. P.: Climate change will affect the Asian water towers, Science, 328, 5984, 1382–1385, doi:10.1126/ science.1183188, 2010.

Ines, A. V. M. and Hansen, J. W.: Bias correction of daily GCM rainfall for crop simulation studies, Agr. Forest Meteorol., 138, 44–53, 2006.

IPCC: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007.

Johnson, F. and Sharma, A.: Measurement of GCM skill in predicting variables relevant for hydroclimatological assessments, J. Climate, 22, 16, 4373–4382, doi:10.1175/2009JCLI2681.1, 2009.

Kay, A. L. and Davies, V. A.: Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts, J. Hydrol., 358, 221–239, 2008.

Kay, A. L., Davies, H. N., Bell, V. A., and Jones R. G.: Comparison of uncertainty sources for climate change impacts: flood frequency in England, Climatic Change, 92, 41–63, 2009.

Korzun, V. I., Sokolov, A. A., Budyko, M. I., Voskresensky, K. P., Kalinin, G. P., Konoplyantsev, A. A., Korotkevich, E. S., and L'vovitch, M. I.: World Water Balance and Water Resources of the Earth, UNESCO, 663 pp., 1978.

L'vovich, M. I.: World water resources and their future, American Geophysical Union, Washington DC, 1979

Leander, R. and Buishand, A. T.: Resampling of regional climate model output for the simulation of extreme river flows, J. Hydrol., 332, 487–496, 2007.

Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, J. Hydrol., 296, 1–22, 2004.

Lehner, B., Döll, P., Alcamo, J., Henrichs, T., and Kaspar, F.: Estimating the impact of global change on flood and drought risks in Europe: a continental, integrated analysis, Climatic Change, 75, 273–299, 2006.

Liu, J., Wiberg, D., Zehnder, A. J. B., and Yang, H.: Modelling the role of irrigation in winter wheat yield, crop water productivity, and production in China, Irrig. Sci., 26(1), 21–33, 2007.

Liu, J., Zehnder, A. J. B. and Yang, H.: Global consumptive water use for crop production: The importance of green water and virtual water, Water Resour. Res., 45, W05428, doi:10.1029/2007WR006051, 2009.

Meehl, G. A., Zwiers, F., Evans, J., Knutson, T., Mearns, L., and Whetton, P.: Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future

climate change, B. Am. Meteorol. Soc., 81(3), 413–416, 2000.

Monteith, J. L.: Evaporation and environment, Symp. Soc. Exp. Biol., 19, 205–234, 1965.

Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, 2004.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 3, 282–290, 1970.

New, M., Hulme, M., and Jones, P.: Representing Twentieth-Century space-time climate variability. Part 1: Development of a 1961–90 mean monthly terrestrial climatology, J. Climate, 12(3), 829–856, 1999.

New, M., Hulme, M., and Jones, P.: Representing Twentieth-Century Space–Time Climate Variability. Part II: Development of 1901–96 Monthly Grids of Terrestrial Surface Climate, J. Climate, 13, 13(2217–2238), 2000.

Nijssen, B., O'Donnel, G. M., and Lettenmaier, D. P.: Predicting the discharge of global rivers, Am. Meteorol. Soc., 3307–3323, 2001.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 – Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modeling, J. Hydrol., 303, 290–306, 2005.

Perkins, S. E. and Pitman A. J.: Do weak AR4 models bias projections of future climate changes over Australia? Climatic change, 93, 527–558, doi:10.1007/s10584-008-9502-1, 2009.

Prudhomme, C., Reynard, N., and Crooks, S.: Downscaling of global climate models for flood frequency analysis: where are we now?, Hydrol. Process., 16, 1137–1150, doi:10.1002/hyp.1054, 2002.

Prudhomme, C. and Davies, H.: Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: Future climate, Climatic Change, 93, 197–222, doi:10.1007/s10584-008-9461-6, 2008.

Sheffield, J., Andreadis, K. M., Wood, E. F., and Lettenmaier, D. P.: Global and continental drought in the second half of the 20th century: severity-area-duration analysis and temporal variability of large-scale events, J. Climate, 22(8), 1962–1981, 2009.

Shiklomanov, I. A.: Assessment of water resources and availability in the world. Comprehensive assessment of the freshwater resources of the world, Stockholm, Stockholm Environment Institute, 88 pp., 1997.

Strzepek, K. M. and Yates, D. N.: Climate change impacts on the hydrologic resources of Europe: a simplified continental scale analysis, Climatic Change, 36, 79–92, 1997.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.: The changing character of precipitation, B. Am. Meteor. Soc., 84, 1205–1217, 2003.

Troccoli, A. and Kallberg, P.: Precipitation correction in the ERA-40 reanalysis. ERA-40 Project Rep. Series 13, 6 pp., ECMWF, Reading, UK, 2004.

UN, 2nd UN World Water Development Report: WWDRII data download page, http://wwdrii.sr.unh.edu/download.html, 2006.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., et al.: The ERA-40 re-analysis, Q. J. Roy. Meteor. Soc., 131, 2961–3012, 2006.

Van Beek, L. P. H.: Forcing PCR-GLOBWB with CRU meteorological data, Utrecht University, Utrecht, Netherlands: http://vanbeek.geo.uu.nl/suppinfo/vanbeek2008.pdf, 2008.

Van Beek, L. P. H. and Bierkens, M. F. P.: The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification, Report Department of Physical Geography, Utrecht University, Utrecht, Netherlands, available at: http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf, 2009.

Varis, O., Kajander, T., and Lemmela, R.: Climate water: from climate models to water resources management and vice versa, Climatic Change, 66, 321–344, 2004.

Viviroli, D., Archer, D. R., Buytaert, W., Fowler, H. J., Greenwood, G. B., Hamlet, A. F., Huang, Y., Koboltschnig, G., Litaor, M. I., López-Moreno, J. I., Lorentz, S., Schädler, B., Schwaiger, K., Vuille, M., and Woods, R.: Climate change and mountain water resources: overview and recommendations for research, management and politics, Hydrol. Earth Syst. Sci. Discuss., 7, 2829–2895, doi:10.5194/hessd-7-2829-2010, 2010.

Vörösmarty, C. J., Fekete, B., and Tucker, B. A.: River Discharge Database, Version 1.1 (RivDIS v1.0 supplement), available through the Institute for the Study of Earth, Oceans, and Space/University of New Hampshire, Durham NH, USA, 1998.

Vörösmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B.: Global Water Resources: Vulnerability from Climate Change and Population Growth, Science 289, 284–288, 2000.

Wada, Y., Van Beek, L. P. H., Viviroli, D., Dürr, H. H., Weingartner, R., and Bierkens, M. F. P.: Water Stress over the Year: Quantitative Analysis of Seasonality and Severity on a Global Scale, MSc Thesis, University Utrecht, Utrecht, Netherlands, available at: http://igitur-archive.library.uu.nl/student-theses/2010-0308-200229/UUindex.html, 2008.

Widén-Nilsson, E., Halldin, S., and Xu, C.: Global water-balance modelling with WASMOD-M: Parameter estimation and regionalization, J. Hydrol. 340, 105–118, 2007.

Wilby, R. L., Wigley, T. M. L., Conway, D., Jones, P. D., Hewitson, B. C., Main, J., and Wilks, D. S.: Statistical downscaling of general circulation model output: A comparison of methods, Water Resour. Res., 34, 11, 2995–3008, 1998.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, Climatic Change, 62, 189–216, 2004.

Zaitchik, B. F., Rodell, M., and Olivera, F.: Evaluation of the global land data assimilation system using river discharge data and a source-to-sink routing scheme, Water Resour. Res., 46, W06507, doi:10.1029/2009WR007811, 2010.