Hydrol. Earth Syst. Sci. Discuss., 7, 687–724, 2010
www.hydrol-earth-syst-sci-discuss.net/7/687/2010/

Hydrology and Earth System Sciences Discussions

This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

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

Received: 19 January 2010 – Accepted: 19 January 2010 – Published: 29 January 2010

Correspondence to: F. C. Sperna Weiland (frederiek.sperna@deltares.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Data from General Circulation Models (GCMs) are often used in studies investigating 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 was tested by applying bias-corrected daily climate data of the 20CM3 control experiment from an ensemble of twelve GCMs as input to the global hydrological model PCR-GLOBWB. Results are compared with discharges calculated from a model run based on a reference meteorological dataset constructed from the CRU TS2.1 data and ERA-40 reanalysis time-series. Bias-correction was limited to monthly mean values as our focus was on the reproduction of runoff variability. The bias-corrected GCM based runs resemble the reference run reasonably well, especially for rivers with strong seasonal patterns. However, GCM derived discharge quantities are overall too low. Furthermore, from the arctic regimes it can be seen that a few deviating GCMs can bias the ensemble mean. Moreover, the GCMs do not well represent intra- and inter-year variability as exemplified by a limited persistence. This makes them less suitable for the projection of future runoff extremes.

1 Introduction

Because runoff regimes might change significantly as a result of 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). Climate datasets from General

HESSD

7, 687–724, 2010

Circulation Models (GCM) are often used as input for hydrological models to investigate the possible change. Unfortunately, different GCM model results can produce quite varying and even contradictory results (Varis et al., 2004) and the modelled past climate does not always agree with the observed one. Deviations are especially apparent for precipitation, when evaluating historical predictions against observations (Covey et al., 2003; Perkins and Pitman, 2009). Climate models tend to simulate too many days with light rain, at the same time they underestimate the frequency and amount of heavy rain events (Dai, 2006). Many studies concluded that a multi-model ensemble of GCM's should be used to obtain a reliable impression of the spread of possible regional changes and their accompanying uncertainties (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 it can be used.

In this study we force a single global scale hydrological model with results of the 20 century control experiment from an ensemble of 12 GCMs for which complete climate datasets with a daily time-step are provided by the IPCC-PCMDI data portal. We compare discharges calculated from the GCM data with observed discharges, as well as with discharges calculated from the CRU TS2.1 time-series (New et al., 2000) downscaled by ERA-40 (Uppala et al., 2005). Contrary to previous studies (Milly et al., 2006; Nohara et al., 2006) we do not focus on the correct reproduction of mean discharge or river regimes, but on the ability of a GCM-forced hydrological model to reproduce global discharge variability (extremes, seasonal variation and inter-annual correlation), parameters that are of importance to water management. We realize that a correct reproduction of current discharge variability is no guarantee that projected discharge variability is correct as well (Prudhomme and Davies, 2008). However we want to investigate the influence of deviating GCM variability on the resulting hydrological variability. And an incorrect projection of current discharge variability is likely to preclude accurate projections, which makes the comparison useful.

We restrict ourselves to a bias-correction of monthly averages of GCM precipitation, temperature and derived potential evaporation, where monthly scaling factors are calculated from the difference in long term monthly mean between CRU and GCM quantities. 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, several climate change experiments have shown that GCM variability will change, especially for precipitation, where frequency and storm duration is likely to decrease and intensity will increase resulting in heavier rain events (Trenberth et al., 2003; Allan and Soden, 2008; Meehl et al., 2000). Corrections based on past deviations between modelled and observed variability, although often applied (Leaner and Buishand, 2007; Ines and Hansen, 2006; Wood et al., 2004), may therefore not hold in future. Applying variability corrections might 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, i.e. assessing the ability of GCM-forced hydrological models to reproduce extremes and intra- and inter-year variability in discharge, thereby assessing the usability of these 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.

2 Methods

The distributed global hydrological model PCR-GLOBWB (van Beek and Bierkens, 2009; Bierkens and van Beek, 2009) is run on a daily time-step with bias-corrected meteorological time series from 12 GCMs for the period 1961–1990. To evaluate the quality of the hydrological model on reproducing observed discharge when using bias-corrected GCM data as input, resulting runoff regimes are compared with observed regimes. Furthermore calculated maps for the ensemble mean thirty year average,

Q_{10}, Q_{90}, timing of peak discharge, inter-annual variability and lag-1 correlation are compared with maps derived from a run based on a reference meteo data set, created from ERA-40 (Uppala et al., 2005) and CRU (New et al., 2000) data.

2.1 Hydrological model

2.1.1 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 (2003) describing VIC (Nijssen, 2001), Macro-PDM (Arnell, 1999), WBM (Vörösmarty, 1998) and WGHM (Döll, 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°, 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, 2003) is the sub model of the global water use and availability model WaterGAP (Alcamo, 2003).

Besides by stand-alone hydrological models, global water balances have also been modelled by coupled vegetation water balance models 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. However, no lateral flow between cells or river routing is applied. 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° (Sheffield et al., 2009), where vegetation, soil moisture and the application of precipitation are modelled 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 river regimes. Therefore the hydrological model used should be designed to calculate this variable, 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 and evaporation loss from wetlands, lakes and reservoir (van Beek and Bierkens, 2009).

2.1.2 PCR-GLOBWB

PCR-GLOBWB is a global distributed hydrological model with a resolution of 0.5° (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 rainfall

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) which includes lakes, wetlands and large reservoirs. Information on model performance is given in Appendix A.

2.2 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. Besides boundary conditions for studies based on these scenarios, the IPCC also provided boundary conditions for a 20th 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, while 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 it 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 (1961–1990) and the future scenario A2, which we will analyze in future research. For those models where multiple ensemble runs were available the first run was selected. Although the data portal does

not provide all required parameters for the Hadley centre climate models, HADGEM1 has been included for we are interested in its performance. HADGEM1 data has been retrieved from the CERA-gateway, http://cera-www.dkrz.de. Table 1 gives an overview of the selected models.

2.2.1 Derivation of potential evaporation

Both temperature and precipitation are directly available at the data portal for all selected climate models. Evaporation had to be derived from other variables, using the Penman-Monteith equation (Monteith, 1965). For those models where the required surface pressure fields were not directly available, surface pressure is 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 will not equal the dew temperature. As suggested by Allen (1998) we subtracted 2° 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). Potential evaporation calculated with the Blaney-Criddle equation is too high for Europe, North-America and the north of Asia during summer and too low for Africa and Asia. Especially the first difference 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 on by the GCMs host institutes and evaporation is not provided for most of the models.

2.2.2 Bias correction and downscaling

Rainfall time-series from GCMs are often biased in monthly amounts, frequency and intensity. For other parameters deviations from observed quantities are present as well. Furthermore the resolution of GCMs is relatively course compared to the desired resolution for hydrological modelling. Therefore previous studies concluded that some kind of bias-correction and downscaling is needed before the data can be used in climate impact studies (Wilby et al., 1998; Ines and Hansen, 2006; Wood et al., 2004). We calculated discharge time-series from the uncorrected GCM meteo data and concluded that deviations between modelled and observed discharges were that large that for our datasets bias-correction was necessary as well. A bias-correction was applied to monthly averages of GCM precipitation, temperature and derived potential evaporation. Adjustment of temporal distributions of time-series has been avoided where possible, in order to keep the GCM variability unchanged. Monthly scaling factors are calculated from the difference (temperature) or ratio (precipitation and evaporation) in long term monthly mean between CRU and GCM quantities. Applying these correction factors also implies downscaling of the GCM data to CRU resolution. For temperature an additive correction was used:

$$T_{\text{corrected\_GCM}} = T_{\text{GCM}} - (\overline{T}_{\text{CRU}} - \overline{T}_{\text{GCM}}) \quad (1)$$

Where T is the daily temperature and $\overline{T}$ is the 30 year average monthly temperature.

For evaporation a multiplicative correction is used, to avoid the occurrence of negative evaporation.

$$ETP_{\text{corrected\_GCM}} = ETP_{\text{GCM}} \frac{\overline{ETP}_{\text{CRU}}}{\overline{ETP}_{\text{GCM}}} \quad (2)$$

For precipitation a similar multiplicative correction is used as in Eq. (2). For some regions (North-Africa, Amazone and Himalaya) differences between GCM and CRU

monthly rainfall amount and number of wet days can be very large. A simple multiplicative correction in these cases resulted in unrealistic rainfall peaks in the bias-corrected rainfall time-series. Time-series obtained with additive monthly correction contained negative rainfall values. Therefore the bias-correction of precipitation is extended with a minimum daily precipitation amount that has to be exceeded by the GCM precipitation before the multiplicative correction can be used (van Beek, 2008). The threshold equals the mean daily CRU rainfall amount:

$$P_{\text{crit}} = \frac{\overline{P}_{\text{CRU}}}{\overline{W}_{\text{CRU}}} \tag{3}$$

With $\overline{W}_{\text{CRU}}$ the 30 year average number of wet days for the specific month. When the precipitation threshold is not exceeded, the days with precipitation occurrence are calculated from a temperature limit below which a day becomes wet.

$$T_{\text{critGCM}} = T_{\text{minGCM}} + (T_{\text{maxGCM}} - T_{\text{minGCM}}) \cdot \frac{\overline{W}_{\text{CRU}}}{N} \tag{4}$$

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 rainfall amount for these days equals:

$$P_{\text{corrected\_GCM}} = \frac{\overline{P}_{\text{CRU}}}{W_{\text{GCM}}} \tag{5}$$

With this equation, rainfall is equally distributed over the wet days and the original temporal distribution of the GCM rainfall time-series is lost. However, this correction method is only applied for those months where relative differences between CRU and GCM time-series were that large that applying simple scaling factors resulted in larger deviations in discharge time-series.

2.3 Reference data

2.3.1 Meteorological data

PCR-GLOBWB was run for the period 1961–1990 with a reference meteorological dataset to create reference modelled discharge maps and time-series for comparison with discharges calculated form GCM data. This dataset was created from the ERA-40 re-analysis dataset from the ECMWF (Uppala et al., 2005), the CRU TS 2.1 monthly time-series (New et al., 2000) and the CRU CLIM 1.0 climatology (New et al., 2002) according to the same method as described above for the GCM data (van Beek, 2008). The ERA-40 dataset is a representation of the weather system on a daily time-step, but precipitation is poorly approximated, particularly in the tropics (Troccoli and Kålberg, 2004). Therefore, we preferred the CRU monthly quantities, which are interpolated from observed meteorological time-series. The hydrological model run based on this reference meteorological dataset will be referred to as reference run.

2.3.2 Hydrological data

We selected 19 large catchments (Fig. 1) to be included in this study, which cover a variety of climate zones, latitudes and continents. For most of these catchments average monthly discharges are available from the Global Runoff Data Centre (GRDC, 2007). For the remaining catchments data from the Global River Discharge Database were used (Vörösmarty et al., 1998).

2.4 Statistical analysis

2.4.1 Statistics

For all models we calculated the thirty year average modelled and the thirty year average observed mean discharge, 10 (low flow) and 90 (high flow) percentile values for each catchment (see Table 2). From thirty year average quantities of the individual

models we calculated the ensemble mean. To compare the temporal behavior of the GCM based model runs with the reference model run and the observations, we calculated the inter-annual discharge variability for the 30 year annual average discharges, the auto-correlation for a time-lag of one year and the yearly timing of the discharge peak occurrence. Inter-annual variability is of importance when GCMs are used to derive extreme value distributions. Reproduction of the lag-one autocorrelation is important for analyzing multi-year droughts related to persistent weather anomalies, for example those related to ENSO.

2.4.2 Map comparison

PCR-GLOBWB calculates daily maps with accumulative routed discharge per cell. For all statistical variables we calculated the GCM ensemble mean values and compared those with the statistics calculated from the results of the reference run. The performance is quantified by the relative difference between, respectively the yearly average mean discharge, the Q_{90}, Q_{10} and the inter-annual variability. For these variables we used relative values, because our main interest was to assess whether the GCM derived discharge statistics are relatively high or low compared to the calculated reference. For the yearly average month of peak occurrence we calculated absolute differences, revealing whether and where the timing of the peak discharge is in accordance with the results of the reference run.

$$\overline{Q}_{\text{rel\_diff}} = \frac{\overline{Q}_{\text{GCM}} - \overline{Q}_{\text{REF}}}{\overline{Q}_{\text{REF}}} \tag{6}$$

$$\overline{Q}_{\text{diff}} = \overline{Q}_{\text{GCM}} - \overline{Q}_{\text{REF}} \tag{7}$$

Where $\overline{Q}_{\text{REF}}$ and $\overline{Q}_{\text{GCM}}$ are, respectively the thirty year average discharge statistics derived from the reference run and the thirty year average statistics derived from the ensemble hydrological runs based on the different GCMs.

Furthermore we tested the significance of the calculated lag-1 correlation with a t-test for a significance level of 90%:

$$t = r\sqrt{\frac{n-2}{1-r^2}} \tag{8}$$

Where r is the calculated lag-correlation coefficient and n is the number of lagged pairs that can be formed. To visualize the ensemble results, for each individual cell the number of models, for which significant correlation was calculated, was displayed on a map.

To reveal the areas where the spread between the 12 models is largest, we calculated maps with the standard deviation of the ensemble of 12 GCM derived fields for the Q_{90}, Q_{10}, Q_{mean}, Q_{peak} and IAV according to:

$$\text{sdv} = \sqrt{\frac{1}{M}\sum_{i=1}^{12}(Q_i - \overline{Q})^2} \tag{9}$$

Where i corresponds to the model number, M is the total number of models (12), Q can be Q_{mean}, Q_{90}, Q_{10}, Q_{peak}, IAV or Cor-lag1 and $\overline{Q}$ is the ensemble mean value of the 12 GCMs for the given parameter.

3 Results and discussion

3.1 Catchment results

3.1.1 Hydrological regimes

For all catchments hydrological regimes have been calculated for: 1) simulated discharges for each individual bias-corrected GCM dataset, 2) the ensemble mean of the discharges calculated by these individual GCMs and 3) the reference run (see Fig. 2).

The calculated regimes are compared with the observed regimes derived from the GRDC time-series. Overall the simple monthly bias-correction reduces the differences between the reference run and the GCM based runs and it also reduces the spread between the modelled discharge for the ensemble of GCMs. For rivers with a strong seasonal pattern, for example the Monsoon driven rivers Brahmaputra, Ganges and Mekong, the resemblance of the GCM runs is large. Also for both the Lena and the MacKenzie, which have a snowmelt driven discharge peak that rises as temperature increases in spring, differences are small for most GCMs, with the exception of two models (IPSL-CM4 and HADGEM1) where discharge is relatively constant throughout the year, indicating that temperatures in the GCMs are too high for too many days in the colder half of the year. For other catchments, that are influenced by both rainfall and snowmelt like the Rhine and the Danube, differences in regime pattern remain. For basins in the more arid regions of the world, such as the Murray, Zambezi, Niger and Orange River, the ensemble mean discharge is lower than the mean discharge calculated by the reference run for most of the year. This is a result of the bias-correction method. For months with too little rain in the GCM dataset, the CRU rain quantities are divided over the number of days that have temperatures below the calculated temperature threshold. Resulting in a larger number of rain days with relatively small equal rain quantities. From such a rainfall distribution more water can evaporate than from a more realistic distribution with rainfall peaks. One of the GCMs (IPSL-CM4) shows exceptional behavior. Deviations in regime are apparent for the Parana, Yellow river, Yangtze, Brahmaputra, Volga, Lena, MacKenzie and present for almost all rivers for this GCM.

From the regime plots it can be seen that there are also differences between observed discharge and discharge results of the reference run. For example for the MacKenzie and Lena, calculated discharges are too low due to snow undercatch problems (Fiedler and Döll, 2007). Discharges for the more arid basins in Africa and Australia are overall too high. In the calculated regime of the Indus the snow and glacier melt driven discharge occurs too early in spring, while the Monsoon influenced

discharge peak is relatively small, as the melt water peak that in reality more or less coincides with the Monsoon maximum, has already occurred. Besides by model deficiencies these differences are caused by errors in the observed discharges and in the CRU dataset, which is known to be inaccurate for parts of Africa and suffers from undercatch in snow dominated areas. By using the CRU dataset as reference for the bias-correction, the deviations present in the CRU set are introduced in the corrected GCM data as well.

3.1.2 Catchment statistics

Discharge statistics of the hydrological model runs are given in Table 3. For most rivers, discharge calculated from the ensemble of GCM results is lower than discharge calculated with the reference run. This originates from the fact that GCMs in general have more days with rain (Dai, 2006; Ines and Hansen, 2006) and therefore daily rainfall amounts will be lower than in the bias-corrected ERA-40 dataset. With this GCM rainfall distribution a larger part of rainfall can evaporate or infiltrate. Relative differences in Q_{mean}, Q_{90} and Q_{10} for the reference and the ensemble mean derived discharge for the individual rivers range between −57 and +31% (leaving out the extreme outliers of 100% for the driest rivers). However 47% of the GCM derived values (Q_{90}, Q_{10} and Q_{mean} statistics for 19 catchments) deviate less than 10%. Ensemble standard deviations for the Q_{mean}, Q_{90} and Q_{10} are comparable in size.

For the Lena, MacKenzie and Volga, GCM derived mean discharge is higher than the reference mean discharge. For these three rivers there are two GCMs with deviating regime curves. These GCMs also cause the Q_{90} to be too high for the Volga and MacKenzie. This illustrates that a few deviating GCMs can bias the discharge statistics derived from the multi-model ensemble. The ensemble mean Q_{10} values exceed the reference value for half of the catchments. GCM calculated Q_{10} values are relatively high, because the GCMs have a lower inter-annual variability and prolonged droughts occur less often.

Inter-annual variability is smaller for almost all catchments in the ensemble GCM results than in the results of the reference run (up to 10 times). The MacKenzie is the only exception, but the ensemble standard deviation, and thus the uncertainty, is relatively large for this basin. Furthermore the MacKenzie is one of the few rivers for which GCM derived discharge is higher than discharge from the reference run and for which several regime curves show large deviations.

The lag-1 correlation coefficients calculated from the ensemble of GCMs show large deviations from the coefficients derived from the reference run. There is no clear trend in differences being positive or negative. Uncertainty between GCMs is large, the ensemble standard deviation of the lag-1 correlation exceeds the ensemble mean for more than half of the catchments. The reference run correlation coefficients also deviate from correlation coefficients derived from the GRDC data. In general GRDC values are higher. This is most likely due to regulation and presence of reservoirs in some of the large rivers.

3.2 Spatial differences

3.2.1 Mean, 10th and 90th percentile discharge

Because the bias-corrections are based on thirty year average monthly precipitation and evaporation quantities, it can be expected that thirty year average mean annual accumulated discharges, derived from the ensemble of GCM based hydrological model run, are comparable to results of the reference run. However from Fig. 3 it can be seen that this does not hold for the entire globe. The Middle-East, Southern Africa and Australia are drier in the results of the GCM ensemble. This is caused by the equal division of rain over potential rain days, applied when GCM time-series contain not enough rain in a month. Overall GCM derived discharges are higher for the Northern Hemisphere and lower for the Southern Hemisphere. The maps with the standard deviation of the results of the ensemble of GCMs show that the spread between GCMs is largest for regions with high discharge. This is particularly the case in large rivers where biases are

accumulated over the entire basin, resulting in large discharge deviations in the main rivers. The GCM based runs also show large deviations for Greenland, caused by both extreme biases and high correction factors for precipitation and evaporation and hydrological model deficiencies. Because of the limited information content of discharge changes in Greenland, this region has been excluded.

By correcting only monthly mean quantities, the daily rainfall distribution of the original GCM datasets remains practically unchanged. Herewith, the occurrence of extreme discharges should be less influenced than the average discharge. Figures 4 and 5 show maps of the Q_{10} and Q_{90} discharge values. Again differences are largest for Australia, North and South Africa and the Middle-East. The Q_{10} for the ensemble of GCMs is higher for large parts of the world, except for arid and Monsoon influenced regions. The Q_{90} is relatively high for the Amazone basin, Indonesia, India and parts of the Northern regions and too low for the same regions, but more extended, as for the Q_{10} values. Relative difference between the GCM based runs and the reference run are smaller for Q_{90} values than for Q_{10} values. Apparently the variability is smaller in (corrected) GCM datasets, resulting in less extreme values for Q_{10} and Q_{90}. This shows that a simple bias-correction leaves part of the GCM behavior, the behavior that influences extremes, unchanged. This is an argument for correcting the total cumulative distributions of daily discharge or precipitation. However, there is no guarantee that such corrections may hold for a future climate.

3.2.2 Timing of discharge peak

The calculated global pattern of timing of peak occurrence of monthly mean discharge (Fig. 6) is comparable for the reference run and the ensemble of GCMs. Although in the GCM results discharge peaks occur more often in January. For the southern drier regions, this can be explained by the difficulty of distinguishing the month with highest discharge when there is little variation in discharge. For Western Europe and North-America peak discharges occur to early in the GCMs. Differences between GCMs are largest in northern and in drier regions.

3.2.3 Inter-annual variability and persistence

Figure 7 shows that for most of the world the GCM derived inter-annual discharge variability is smaller than the inter-annual variability derived from the reference run. This difference is also present in discharges calculated from non-bias corrected GCM data, but is amplified in the runs using bias-corrected data. A lower inter-annual variability implies more resemblance between years and less extremes. This indicates that GCMs have difficulties in reproducing very dry and very wet years, possibly underestimating the frequency of floods and droughts.

To measure multi-year persistence, the lag-1 correlation is calculated. The variation between GCMs is large for the lag-1 correlation, especially in the more humid regions on the Northern Hemisphere. Overall positive correlation coefficients, ranging between 0.2 and 0.8, were found for both the reference run and the results of the ensemble of GCMs. Figure 8 shows maps with the significance of the calculated correlation for a significance level of 90%. For the reference run, regions with significant lag-1 correlation are indicated in black. When a significant correlation is present, predictability of next years hydrological conditions based on the current hydrological conditions is large. For regions where this is the case a same predictability, or correlation, should also be present in the GCMs. We stress however that the value of ERA-40 as a benchmark for correlation coefficients of individual rivers was found to be limited (see Table 4), probably because most large rivers are regulated, so that their results should be interpret with care.

For the ensemble of GCMs we calculated for each cell the number of models with significant correlation. Lag-1 correlation is significant in the Mississippi basin, Northern Europe and parts of Africa and Australia for both the reference runs and most of the GCMs. In the reference run significant correlation is also found for the Amazone and Lena basins, while only a few GCMs show this significance. There are only small areas where none of the GCMs give significant correlation. The consistency amongst models is small and while the values of the coefficient deviate from those calculated

from the reference run for most of the world we conclude that lag-1 correlation is not well reproduced by most GCMs.

4 Conclusions

In order to evaluate the ability of the hydrological model PCR-GLOBWB to reproduce global discharge variability when being forced with GCM datasets that are bias-corrected with a simple monthly correction method, we calculated the average hydrological regimes and various discharge statistics for the period 1961–1990 for the ensemble of 12 GCM datasets and compared the results with observed data and a reference hydrological model run based on the ERA-40 downscaled CRU dataset.

As expected, after bias-correction the spread between regimes calculated with the 12 corrected GCM datasets is decreased. Differences are smallest for basins with strong seasonal patterns; the Arctic and Monsoon driven rivers. GCM derived discharge is overall too low, as raw GCM data have too many rain days, resulting in many days with little rain from which a larger amount of rain can infiltrate or evaporate. For the Lena, MacKenzie and Volga deviations in regime curves are present for two of the GCMs (IPSL-CM4 and HADGEM1), these deviations increase the ensemble mean Q_{mean} and Q_{90} values. This illustrates that a few deviating GCMs can bias the discharge statistics derived from the multi-model ensemble.

All the GCM based runs show a lower inter-annual variability than both the reference run and observations, meaning an underestimation of the number of extreme years. And indeed, overall the Q_{90} values are lower and the Q_{10} values higher than in the reference run. Inter-annual runoff persistence as present in the reference run is reproduced by the GCM based runs only for a few areas, while results between GCMs vary greatly. This shows that multi-year droughts resulting from large weather anomalies (e.g. ENSO) cannot be produced either.

Overall the GCM based runs resemble the reference run reasonably well. However GCMs do, even after bias-correction, not represent Q_{10} values, lag-1 correlation and inter-annual variability well for the current period, which questions their usefulness for investigation of future changes in extremes.

Appendix A

Model performance PCR-GLOBWB

Despite its coarse resolution PCR-GLOBWB can reproduce observed discharges reasonably well for most selected catchments. To assess its performance we ran the model for the period 1961 to 1990 with CRU TS 2.1 monthly time series (New et al., 2000) downscaled to daily values with the ERA-40 re-analysis data (Uppala et al., 2005) and compared the yearly mean modelled discharge with observed discharges per catchment (see Fig. 9). For most of these catchments average monthly discharges are available from the Global Runoff Data Centre (GRDC, 2007) for the remaining catchments data from the Global River Discharge Database were used (Vörösmarty et al., 1998). Deviations are partly present because the hydrological model calculates natural flows and does not include water use. To investigate the influence of water use, calculated discharge are compared with the sum of observed discharge and estimated water demand. Water demand is estimated on a grid of 0.5° and is the sum of estimated industrial, agricultural and domestic water demand (Wada et al., 2010).

The 30-year mean observed ($\overline{Q}_{\mathrm{GRDC}}$) and modelled ($\overline{Q}_{\mathrm{CRU}}$) discharge are compared. Beside through inadequacies of the hydrological model, differences between modelled and observed discharge can also be caused by errors in discharge measurements or deviations in meteorological data. From Fig. 9 it can be seen 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 is also the case for the Danube, Ganges, Yellow River and Rhine. Deviations for the river Orange, Congo and

Zambezi are present as well, however this difference can not be explained by water use. For the Mississippi and the Volga the inclusion of water demand indicates that model results are worse than expected, deviations are larger than when comparing with observed discharge only. The Lena and MacKenzie river show deviations as well. These deviations are caused by undercatch in the CRU snowfall amounts (Fiedler and Döll, 2007).

Additionally 30 year average calculated hydrological regimes (CRU_ERA) are compared with observed regimes (GRDC) for each catchments, see Fig. 9. The plots show that the difference between the CRU derived and the observed discharge regime is especially large for the MacKenzie as mentioned before. Furthermore the observed snowmelt driven discharge peak for the Lena is steeper than modelled. Large differences are also present for the Murray because of water use. In the calculated regime of the Indus, the snow and glacier melt driven discharge increase is too large and too early in spring, while the Monsoon influenced discharge peak is relatively small. For the Parana observed river flow is more constant through time, because of the presence of large reservoirs.

Acknowledgements. We acknowledge the GCM modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modelling (WGCM) for their roles in making available the WCRP CMIP3 multi-model dataset. Support of this dataset is provided by the Office of Science, US Department of Energy.

References

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

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., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., and Siebert, S.: Development and testinf 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, M. R. and Ingram, W. J.: Constraints on the future changes in climate and the hydrological cycle, Nature, 419, 224–232, 2002.

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

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, 2003, http://www.hydrol-earth-syst-sci.net/7/619/2003/.

Bierkens, M. F. P. and van Beek, L. P.: Seasonal predictability of european discharge: NAO and hydrological response time, J. Hydrometeorol., 10(4), 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.

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, 2007, http://www.hydrol-earth-syst-sci.net/11/1417/2007/.

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-min 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.
Fiedler, K. and Döll, P.: Global modelling of continental water storage changes – sensitivity to different climate data sets, Adv. Geosci., 11, 63–68, 2007, http://www.adv-geosci.net/11/63/2007/.
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: Major River Basins of the World/Global Runoff Data Centre, 56002, Federal Institute of Hydrology (BfG), Koblenz, Germany, 2007.
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.
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.
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.
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., 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., 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.

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.
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), 427–436, 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.
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, 2000.
New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, Climate Res., 21, 1–25, 2002.
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. 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.
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.
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.

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. Meteorol. 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, last access: 20 March 2009, 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, The Netherlands, http://vanbeek.geo.uu.nl/suppinfo/vanbeek2008.pdf, last access: 14 January 2010, 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, The Netherlands, available at: http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf, last access: 14 January 2010, 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.
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.
Wada, Y., van Beek, L. P. H., Viviroli, D., Dürr, H. H., Weingartner, R., and Bierkens, M. F. P.: Variations in blue water stress: quantitative analysis of seasonality and severity at the global scale, in preparation, 2010.
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.

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
HADGEM1	Met Office's Hadley Centre for Climate Prediction	UK	HADGEM

HESSD
7, 687–724, 2010

Table 2. Description of statistical quantities of interest. Where j corresponds to the year number, N is the total number of years (30), Q_j is the mean annual discharge for the calculated for the years 1961 to 1990, $\overline{Q}$ is the 30 year average mean discharge, Q is an array containing all day-values for the thirty year period and k is the time-lag of one year.

Parameter	Description	Equation
Q_{mean}	Thirty year average mean annual discharge	$Q_{\text{mean}} = \frac{1}{N}\sum_{j=1}^{30} Q_j$
Q_{90}	Discharge exceeded at 10% of the days during the thirty year period	–
Q_{10}	Discharge exceeded at 90% of the days during the thirty year period	–
IAV	Inter annual variability derived from yearly average discharges	$\text{IAV} = \sqrt{\frac{1}{N}\sum_{j=1}^{30}\left(Q_j - \overline{Q}\right)^2}$
Cor-lag1	Auto-correlation of thirty annual average discharges for time lag of one year	$\text{Cor-lag1} = \frac{\frac{1}{N}\sum_{j=k}^{N-(k+1)}(Q_j - \overline{Q})(Q_{j+k} - \overline{Q})}{\frac{1}{N}\sum_{j=1}^{N-1}(Q_j - \overline{Q})^2}$

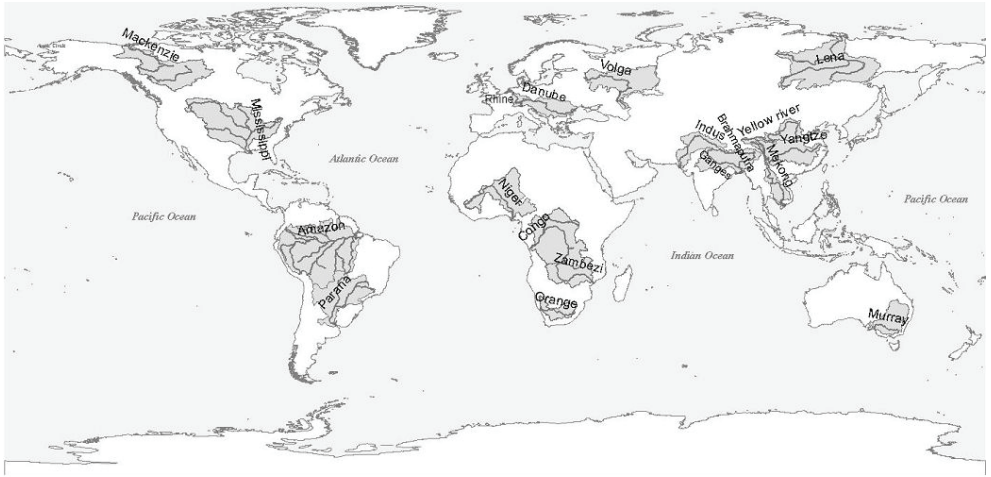
Table 3. Discharge time-series statistics derived from: observations (GRDC), PCR-GLOBWB run based on reference meteo dataset (ERA) and ensemble of GCM based discharge results.

	Q_{mean}				Q_{90}			Q_{10}		
	GRDC	ERA	ens_mean	ens_sdv	ERA	ens_mean	ens_sdv	ERA	ens_mean	ens_sdv
Amazone	171 296	141 007	127 413	1409	208 819	185 549	4698	76 583	70 398	1044
Brahmaputra	18 648	13 944	13 253	559	32 864	32 233	951	3320	3257	196
Congo river	43 366	51 752	46 990	1368	70 681	58 126	2928	34 576	34 012	780
Danube	6620	7696	7409	508	12 244	10 922	914	3762	4280	577
Ganges	10 996	15 338	10 886	869	38 209	30 087	2003	3047	1707	114
Indus	3037	2365	1913	81	4963	3839	273	769	730	33
Lena	17 051	17 235	17 462	1512	35 094	31 902	1286	6213	7574	2418
MacKenzie	8531	6174	6531	1378	8438	8963	1836	4400	4627	1029
Mekong	7472	6683	6619	200	14 609	14 919	446	1836	1878	45
Mississippi	17 466	15 961	14 152	1219	26 690	22 915	1415	7436	7527	1032
Murray	186	610	459	23	1246	884	62	155	159	9
Niger	971	2462	2130	129	6794	6400	428	0	0	0
Orange	145	212	96	16	356	152	36	70	57	5
Parana	17 919	25 034	22 113	961	44 645	39 601	2280	10 151	9698	601
Rhine	2377	2589	2541	97	4469	3819	148	1071	1357	134
Volga	7382	6885	7460	1512	10 976	11 308	1990	3540	4639	1174
Yangtze	27 670	22 145	21 337	1096	36 243	33 804	1563	11 422	11 688	808
Yellow river	1412	2458	2037	314	4706	3896	726	1043	911	97
Zambezi	1283	2387	1845	194	7059	5644	532	0	0	0

Table 3. Continued.

	Inter-annual variability				lag-1 correlation			bold=significant	
	GRDC	ERA	ens_mean	ens_sdv	GRDC		ERA	ens_mean	ens_sdv
Amazone	15 145	21 224	3612	2355		*	**0.59**	0.17	0.25
Brahmaputra	1801	1564	425	163		*	0.02	0.11	0.25
Congo river	5106	7977	1542	555		*	**0.56**	0.24	0.18
Danube	1087	1346	424	252	0.22		0.26	0.37	0.32
Ganges	2652	2416	624	322		*	−0.09	0.20	0.27
Indus	1190	655	65	12		*	0.25	0.12	0.14
Lena	1459	1930	778	693	**0.82**		0.22	0.27	0.41
MacKenzie	839	615	659	780		*	**0.47**	**0.61**	0.22
Mekong	1548	1066	312	141	**0.69**		−0.07	0.10	0.14
Mississippi	4083	4043	1106	626	0.41		**0.43**	**0.42**	0.18
Murray	162	275	24	13	0.41		**0.41**	0.29	0.21
Niger	212	475	97	43		*	**0.60**	0.17	0.28
Orange	106	117	8	7		*	−0.03	**0.44**	0.20
Parana	5237	5467	1121	454	0.40		0.10	0.19	0.30
Rhine	540	494	96	40	**0.62**		0.23	0.22	0.33
Volga	1257	1541	968	746	**0.76**		0.20	**0.49**	0.31
Yangtze	3595	2868	905	430	0.41		0.01	0.34	0.28
Yellow river	456	431	132	182	0.40		0.02	**0.52**	0.22
Zambezi	496	1113	243	106	**0.66**		0.32	0.09	0.24

* No lag-1 correlation calculated for observed discharge time-series with less than 25 consecutive years.



Catchment	Area (km2)	Qavg(m3/s)	Gauge		Catchment	Area (km2)	Qavg(m3/s)	Gauge
Amazone	6,915,000	190,000	Obidos		Murray river	1,061,469	767	Wakool Junction
Bramaputra	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	Farakka		Rhine	65,683	2,200	Rees
Indus	1,165,000	6,600	Attock		Volga	1,800,000	31,900	Datong
Lena	2,500,000	17,000	Aldan		Yangtze	752,000	2,571	Huayuankou
MacKenzie	1,805,000	10,700	Norman Wells		Yellow river	1,380,000	8,060	Volgograd
Mekong	2,981,076	12,743	Vicksburg		Zambezi	1,390,000	3,400	Katom a Mulilo
Mississippi	795,000	16,000	Mukdahan					

Fig. 1. Selected catchments.

HESSD
7, 687–724, 2010

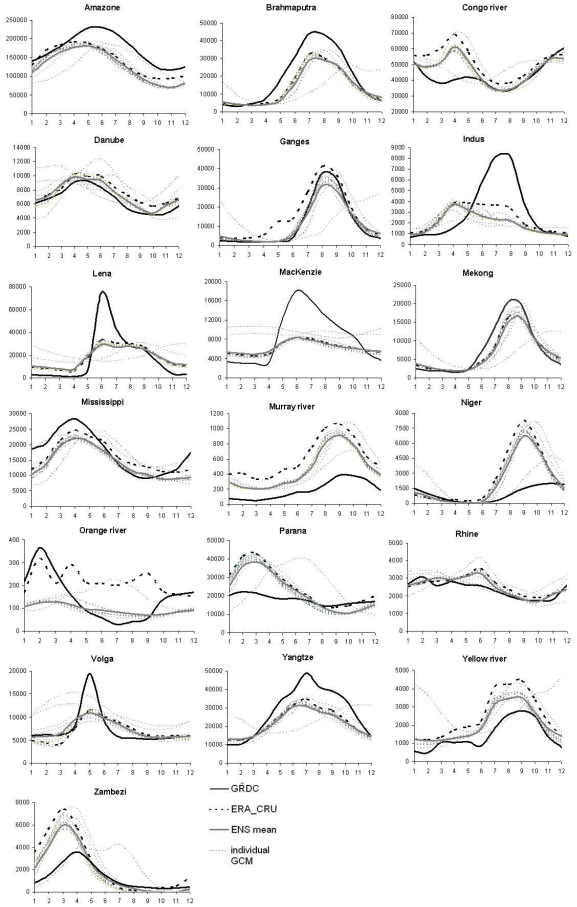


Fig. 2. Hydrological regimes for all 19 catchments derived from discharges calculated for the period 1961-1990, with the average monthly discharge (m^3/s) on the y-axis and the month numbers on the x-axis. For the GRDC, CRU_ERA, individual GCMs and ensemble mean.

HESSD
7, 687–724, 2010

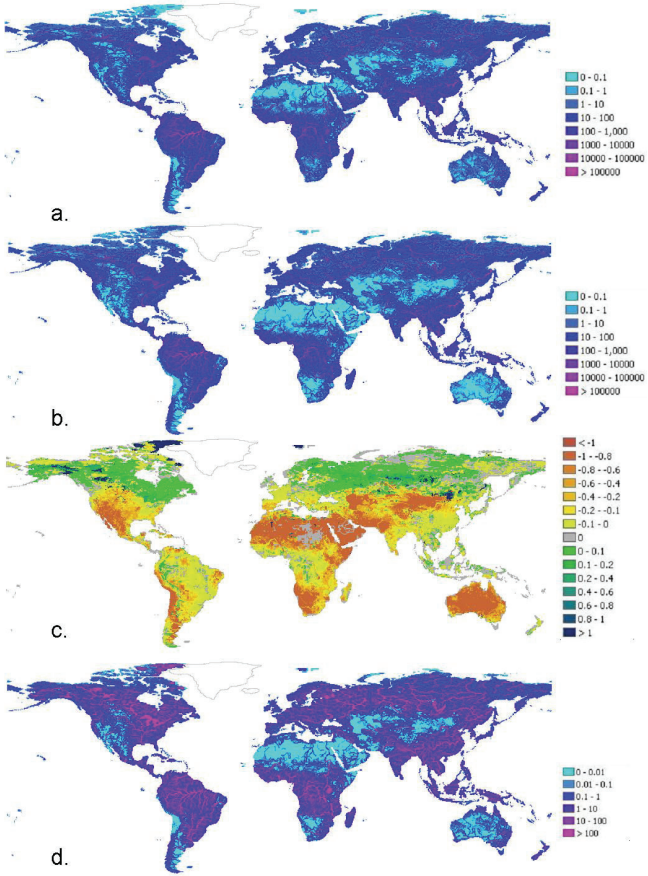


Fig. 3. Thirty year average discharge calculated from **(a)** the reference run and **(b)** the ensemble mean discharge results of the 12 GCMs, **(c)** the average difference and **(d)** ensemble standard deviation.

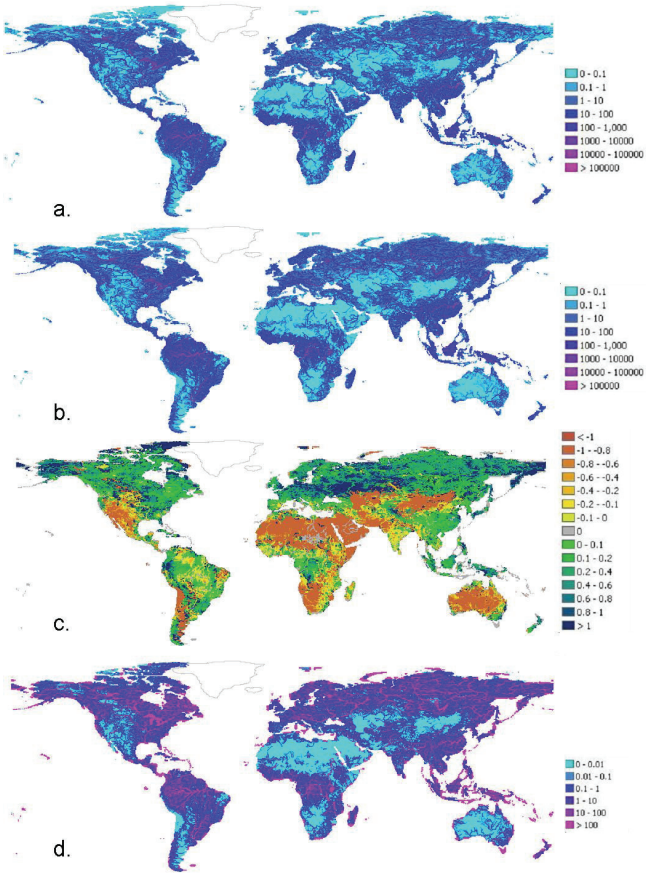


Fig. 4. Thirty year average Q10 values calculated from **(a)** the reference run and (b) the ensemble mean discharge results of the 12 GCMs, **(c)** the average difference and **(d)** ensemble standard deviation.

HESSD
7, 687–724, 2010

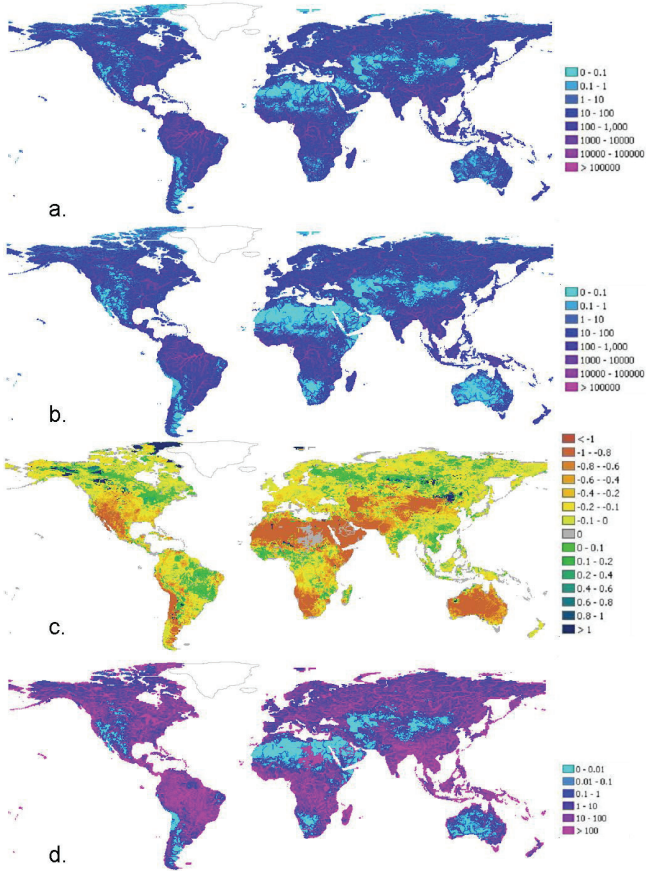


Fig. 5. Thirty year average Q90 values calculated from **(a)** the reference run and **(b)** the ensemble mean discharge results of the 12 GCMs, **(c)** the average difference and **(d)** ensemble standard deviation.

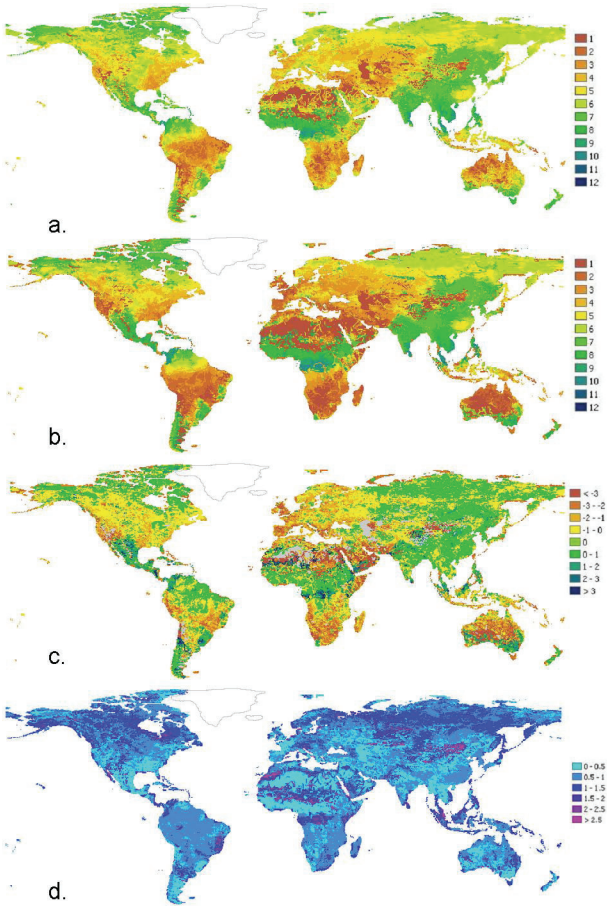


Fig. 6. Thirty year average month of peak discharge occurrence calculated from **(a)** the reference run and **(b)** the ensemble results of the 12 GCMs, **(c)** the absolute difference in months and **d)** ensemble standard deviation.

HESSD
7, 687–724, 2010

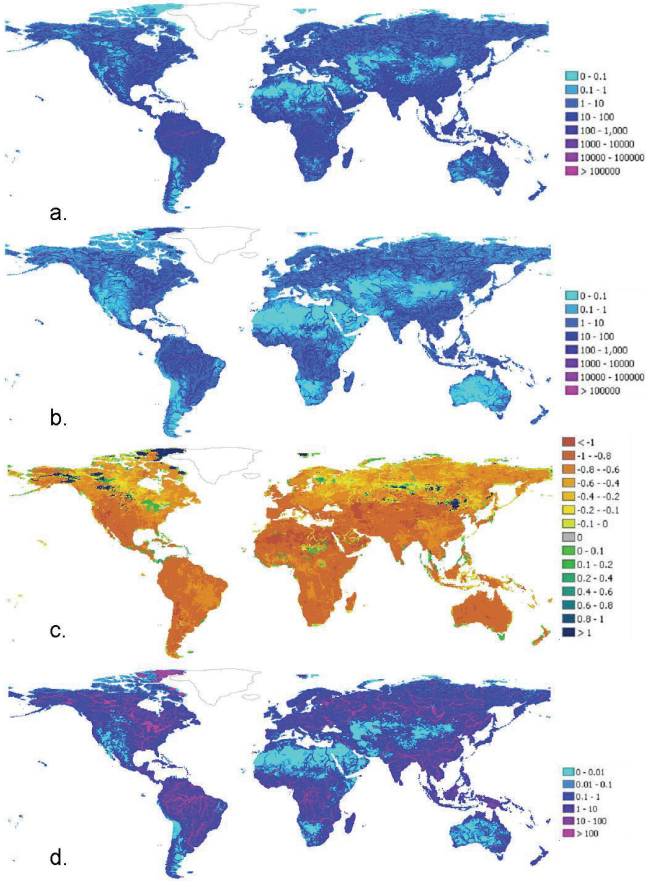


Fig. 7. Inter-annual variability calculated from **(a)** the 30 year reference run and **(b)** the ensemble discharge results of the 12 GCMs, **(c)** the relative difference and **(d)** ensemble standard deviation.

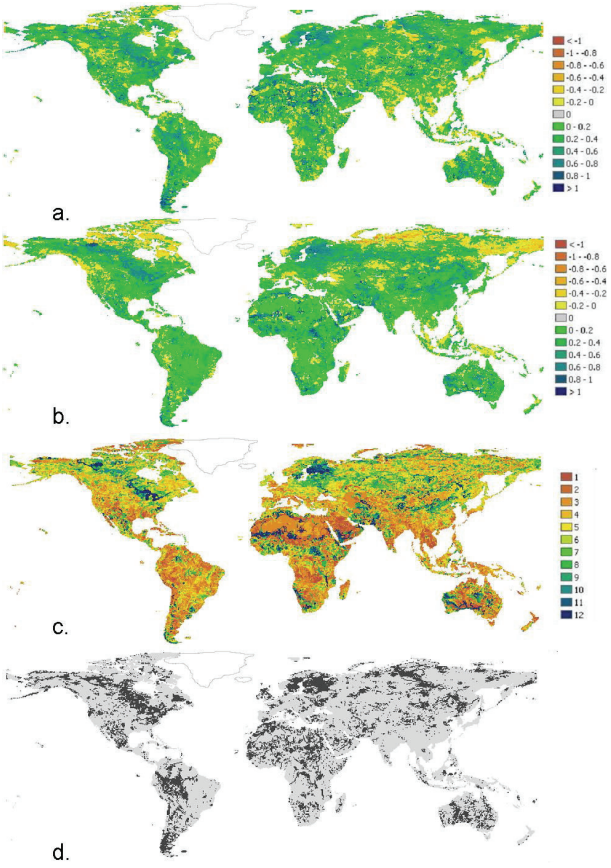


Fig. 8. First two maps show the lag-1 correlation calculated for **(a)** the 30 year reference run and **(b)** the ensemble of 12 GCMs. Map **(c)** shows per model cell the number of model predicting significant correlation. Map **(d)** shows regions with significant correlation (black) and regions with no significant correlation (grey) for the reference run.

HESSD
7, 687–724, 2010

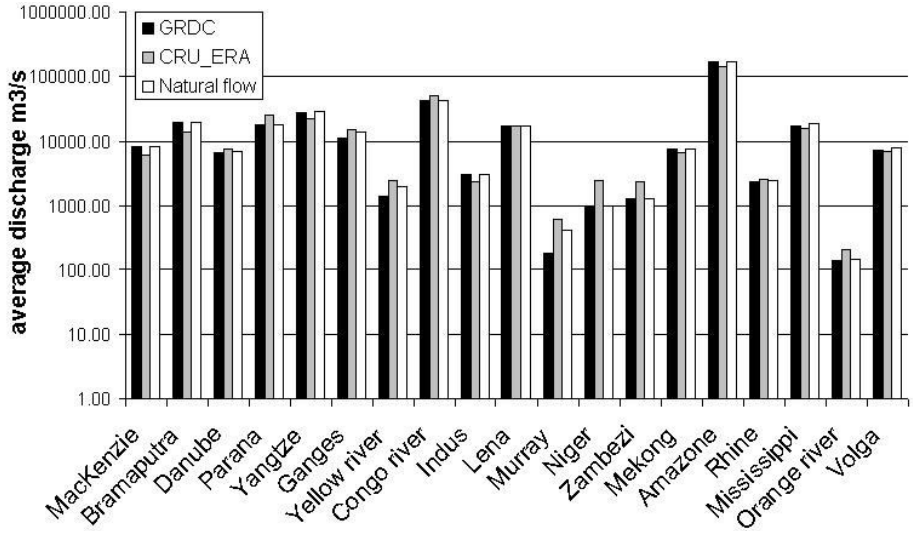


Fig. 9. Average yearly observed (GRDC), observed + demand, and modelled (CRU_ERA) discharge per catchment and deviations.

HESSD
7, 687–724, 2010