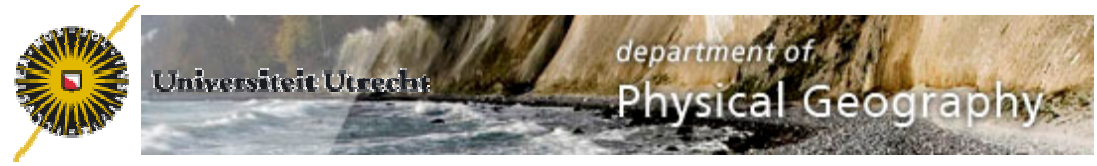
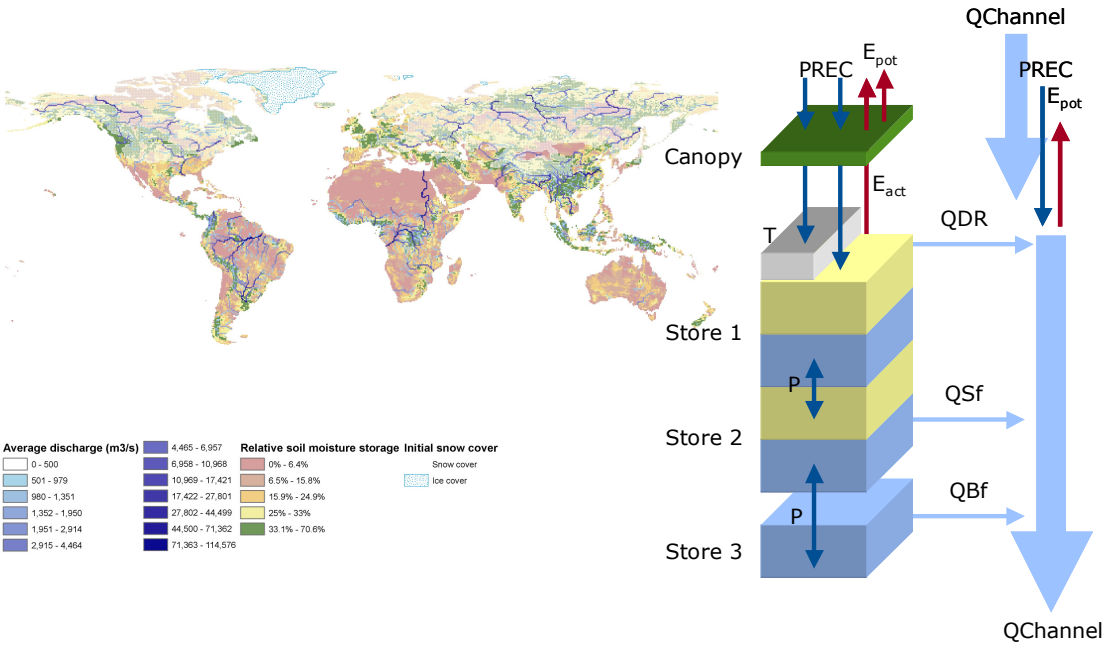



# The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification




## Report

**Department of Physical Geography**
**Faculty of Earth Sciences**
**Utrecht University**

**L.P.H. (Rens) van Beek**
**Marc F.P. Bierkens**

**P.O. Box 80115, 3508 TC, Utrecht, the Netherlands**

## Preface

Since 2003 we have been developing the global hydrological model PCR-GLOBWB, **PCR**aster **GLOB**al **W**ater **B**alance model. The model is used in seasonal prediction, modeling hydrological effects of climate change, water level modeling for $CH_4$ emission from wetlands, modeling of terrestrial water storage and modeling global water resources in the context of water stress. This report serves as a background document to present the details of the modeling concepts used as well as the way the model is parameterized. There is often no space or place to report all these details in papers that use the model. Hence, this background document to refer to. New features are continuously being added and new material added for validation. Consequently, the report is a living document that is altered regularly. Hence the distribution through internet, currently at:

http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf

Please refer to the model by referring to report as:

Van Beek, L.P.H. and M.F.P. Bierkens (2008), *The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification*, Report Department of Physical Geography, Utrecht University, Utrecht, The Netherlands, http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf

Rens van Beek (r.vanbeek@geo.uu.nl)
Marc Bierkens (m.bierkens@geo.uu.nl)

# The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification

*L.P.H. (Rens) van Beek and Marc F.P. Bierkens*

Department of Physical Geography, Utrecht University, P.O. Box 80115, 3508TC, Utrecht, the Netherlands.

## 1. Introduction

Models describing hydrological processes at a global scale are now frequently being used to assess the effect of global change on the world's water resources. Examples are the assessment of global water stress [Vörösmarty et al., 2000; Alcamo et al, 2000; Oki et al, 2001], continental runoff [Nijssen et al., 2001; Fekete et al., 2002], projected change in continental runoff [Milly, et al., 2005], soil moisture fields and global drought [Sheffield and Wood, 2007] and total continental water storage [Güntner et al., 2007].
Available models can be largely divided into hydrological models that are used off-line and land surface models forming the land component of general circulation models (GCMs). Intercomparisons of terrestrial water fluxes from a suite of GCM land surface models can be found in Dirmeyer et al. [1999, 2006]. Although global hydrological models (GHMs) are similar in nature to the GCM land surface models, they usually operate at longer temporal but smaller spatial scales, while terrestrial hydrological processes are represented in more detail. Known GHMs from the literature are VIC [Liang et al., 1994] applied at the global scale [Nijssen et al., 2001], LaD [Milly and Schmakin, 2002], WMB [Fekete et al., 2002], WGHM [Döll et al., 2003] and WASMOD-M [Widén-Nilsson, 2007]. WBM, WGHM and WASMOD-M are purely water-balance type models, while VIC and LaD also resolve the surface energy balance.
In this paper we describe a recently developed global hydrological model called PCR-GLOBWB, which is derived from PCRaster GLOBal Water Balance model. PCRaster [Wesseling et al., 1999] is the dynamic scripting language in which the model is coded. Although our model is similar to existing GHMs in many ways and uses quite some data sets in common, we also developed a number of new concepts that are worth presenting to the hydrological community. In particular we developed new and advanced schemes for subgrid parameterization of surface runoff, interflow and baseflow and added explicit routing of surface water flow using the kinematic wave approximation, dynamic inundation of floodplains and a reservoir scheme. Also, we added a routine for lateral water transport of latent heat from which we can calculate water temperature and river ice thickness. This makes the model suitable for runoff analyses at time-scales smaller than a month, as well as tailored to low flow analyses and nutrient transport..
The remainder of this paper starts with a global description of the model, followed by a more in depth description of the features that are specific to our model when compared to existing GHMs. Next, we describe the verification of the model using runoff data. Continental runoff as calculated from our model is compared to that from earlier studies.

## 2. General description of the model

Here we present the general outline of the PCR-GLOBWB model. Specific new features compared to existing GHMs are presented separately in the next section. A schematic representation of PCR-GLOBWB is given in Figure 2.1. PCR-GLOBWB is a grid-based model of global terrestrial hydrology. Similar to other GHMs it is essentially a leaky-bucket type of model [Bergström, 1976] applied on a cell-by-cell basis. PCR-GLOBWB calculates for each grid cell ($0.5^0 \times 0.5^0$ globally) and for each time step (daily) the water storage in two vertically stacked soil layers (max. depth 0.3 and 1.2 m) and an underlying groundwater layer, as well as the water exchange between the layers and between the top layer and the atmosphere (rainfall, evaporation and snow melt). The model also calculates canopy interception and snow storage.

Meteorological forcing is supplied at a daily time step and assumed constant over a grid cell. Precipitation falls either as snow or rain depending on atmospheric temperature. Snow accumulation and melt is only temperature driven and modeled according to the snow module of the HBV model [Bergström, 1976], including the storage of melt water in the snow pack which may be prone to refreezing and evaporation. Canopy interception store is finite and subject to open water evaporation.. Excess precipitation either adds to the snow pack (in case atmospheric temperature $T < 0$), adds to the liquid pore space in the snow pack (in case of snow present) or infiltrates into the first soil layer. Actual evaporation can either be prescribed directly (e.g. from the output of a GCM) or calculated by the model from potential evaporation and soil moisture status based on Penman-Monteith [Allen et al., 1998].

Sub-grid variability is taken into account by considering separately tall and short vegetation, open water, and different soil types. Short vegetation extracts water from the upper layer only, while tall vegetation extracts water from both soil layers. Calculation of the fraction of saturated soil to assess direct runoff is based on the improved ARNO scheme [Hageman and Gates, 2003] and the surface elevations of the 1×1 km Hydro1k data set [USGS EROS Data Center, 2006a].

The total specific runoff of a cell consists of saturation excess surface runoff, melt water that does not infiltrate, runoff from the second soil reservoir (interflow) and groundwater runoff (baseflow) from the lowest reservoir. Groundwater reservoir characteristic response time is parameterized based on a world map of lithology [Dürr et al, 2005] and drainage density derived from Hydro1k. River discharge is calculated by accumulating and routing specific runoff along the drainage network taken from the DDM30 and including dynamic storage effects and evaporative losses from the GLWD inventory of lakes, wetlands and plain [Lehner and Döll, 2004].

PCR-GLOBWB is coded in PCRaster [Wesseling et al., 1996]. The PCRaster Environmental Modeling language is a high level computer language that uses spatio-temporal operators with intrinsic functionality for constructing spatio-temporal models. It enables a very efficient manipulation of raster-based maps and has several in-built hydrological functions, such as accumulating and routing water and sediments over drainage networks.

# 3. Specific features of PCR-GLOBWB

## 3.1 Forcing PCR-GLOBWB with CRU-data

PCR-GLOBWB can be directly forced with precipitation and evaporation from climate models or re-analysis data such as ERA40 [Uppala et al., 2005]. Evaporation is then enforced as water abstraction from the first soil reservoir (short vegetation) or the first and second soil reservoir (tall vegetation). Alternatively, actual evaporation can be calculated by PCR-GLOBWB itself based on reference evaporation calculated from meteorological variables, crop factors and soil moisture status.

The Climate Research Unit the University of East Anglia provides global climate data that are potentially very useful as forcing for macro-scale hydrological models such as PCR-GLOBWB. The advantage of the CRU products is that they are based on observations, covering the global land mass (gridded data), and processed in a consistent manner. Two well-used products with a spatial resolution of 0.5° and a monthly resolution are the CRU TS 2.1 (New et al. [2000, 2002], time series between 1901 to 2002) and the CRU CLIM 1.0 (New et al. [1999], climatology over 1961-1990). We used these two products to construct daily fields of precipitation and reference evaporation for the ERA40 period (1957-2002).

Unfortunately, some relevant fields, such as radiation and wind speed are only available for the $0.5^0$ resolution as climatology. Other complications in the use of the CRU dataset arise from a varying extent of the landmass for some of the variables, from a discrepancy with the land mask used by the hydrological model, and from the monthly temporal resolution which may lead to undesired results when coupled with non-linear processes such as interception that operate on finer time scales. Van Beek [2008] describes in detail the downscaling of the CRU dataset to daily values with the assistance of the ERA40 reanalysis, its extrapolation to the desired land mask of PCR-GLOBWB and the methodology used to calculate potential evaporation for the different land surfaces in PCR-GLOBWB. In short, the steps taken in Van Beek [2008] are:

1. extrapolate the CRU data to a common land mask for all products and extrapolate all relevant products to the land mask employed by PCR-GLOBWB. Here extrapolation was performed by assigning to cells without values the value of the closes CRU cell belonging to the same Holridge [1967] Life Zone Classification (available at $0.5^0$ from Leemans [1989]);
2. calculate the monthly reference potential evaporation from the CRU dataset for the length of the time series. Here, reference evaporation is the evaporation under well-watered condition of a reference crop (crop height of 0.12 m, a fixed surface resistance of 70 s $m^{-1}$ and an albedo of 0.23) and calculated with the Penman-Monteith equation [Allen et al. 1998]. Temperature and vapour pressure were available from CRU TS 2.1, and average wind from CRI CLIM 1.0. Incoming short-wave radiation was calculated from extraterrestrial radiation and an empirical reduction factor due to cloud cover. Net longwave radiation was calculated from an empirical relation [Allen et al., 1998] involving air temperature, cloud cover, vapour pressure and incoming short-wave radiation;

3. develop a climatology of crop factors for the different land surfaces in PCR-GLOBWB that can be used to convert the monthly reference potential evaporation into vegetation specific values. Conform the land surface parameterisation in PCR-GLOBWB, these crop factors have to be specified for the fraction open water, short vegetation and tall vegetation respectively and effective values have been calculated for each 0.5° cell for each month. For the vegetated surface, the GLCC version 2 (USGS Eros Data Center [2002], mapped globally at 30 arcseconds) land cover types [Olsen 1994a,b] were divided into three categories: natural vegetation, rainfed crops and irrigated crops, where each category was subdivided into tall and short vegetation. Crop factors varying over the year for irrigated crops were obtained from Döll and Siebert [2002]. For natural vegetation and rainfed crops crop factors were calculated from average phenology. First, following the temperature and precipitation climatology (CRU CLIM 1.0) for each PCR-GLOBWB cell the Leaf Area Index (LAI) climatology for each GLCC-type was estimated, with LAI values per type for dormancy and growing season obtained from Hagemann et al. [1999]. Next, an empirical relation [Allen et al., 1998] was used to convert LAI to crop factor. The effective crop factor for short and tall vegetation per PCR-GLOBWB cell and per month was then calculated by averaging crop factors per GLCC-type based on their areal coverage within the cell. Finally, separate crop factors were used for deep water (channels, lakes and reservoirs) and shallow water (floodplains when inundated);
4. Break-down monthly values of CRU precipitation and temperature and the calculated reference potential evapotranspiration to daily values using daily surface fields from the ERA-40 reanalysis to drive the model. The temporal downscaling for temperature is performed by correcting ERA40 daily temperature by the difference between CRU and ERA40 mean monthly temperature. Daily precipitation was estimated by multiplying ERA40 precipitation by the ratio of monthly CRU precipitation to monthly ERA40 precipitation. Reference potential evapotranspiration was also downscaled multiplicatively, using the ration of ERA40 daily temperature to ERA40 monthly temperature as a proxy.

### 3.2 Surface runoff scheme, vertical fluxes and interflow

Figure 1 shows that specific runoff from a cell consists of overland flow, runoff from the second soil reservoir (interflow) and groundwater runoff (baseflow). Here we discuss in more detail how the surface runoff and interflow are parameterized, as well as the fluxes between reservoirs. Groundwater discharge is described in the following section.

*Surface runoff.* The input to the first soil reservoir consists of net precipitation (non-intercepted part) and snow melt. Melt water is first stored in the snow pack up to a maximum storage capacity that is related to snow depth (in snow water equivalent) and an average snow porosity. Water stored in the snow pack may refreeze or is subject to evaporation. Snow melt in excess of the snow water storage capacity is added to the precipitation. The sum of net precipitation and excess snow melt infiltrates into the first soil layer if the soil is not saturated, while surface runoff occurs if the soil is saturated. Variability of within grid soil saturation is accounted for by using the improved Arno scheme [Hagemann and Gates, 2003] where the fraction of surface precipitation and

excess snow melt turned into surface runoff is related to the fraction of saturated soil, given by:

$$x = 1 - \left( \frac{W_{max} - W}{W_{max} - W_{min}} \right)^{\frac{b}{b+1}} . \quad (1)$$

where $W$ [L] is the cell-average water storage in the upper reservoir, $W_{max}$ the maximum average storage and $W_{min}$ the minimum average storage . These parameters all refer to the pervious area in the cell (i.e. not open water) and are based on the FAO gridded soil map of the world [FAO, 1998]. The parameter $b$ is a dimensionless shape factor (-) that defines the distribution of soil water storage within the larger cell and is calculated based on the distribution of maximum rooting depths, which in turn is derived from the 1×1 km distribution of vegetation types from GLCC [Hagemann, 2002; USGS Eros Data Center, 2002]. Based on this parameterisation surface runoff $Q_s$ is related to cell-averaged moisture storage $W$ and net input (sum of net precipitation and excess snow melt) $P_n$ as [Hagemann and Gates, 2003]:

i) if $P_n < W - W_{\min}$ :

$$Q_s = 0 \quad (2)$$

ii) if $W - W_{\min} \le P_n < (b+1)(W_{max} - W_{min}) \left( \frac{W_{max} - W}{W_{max} - W_{min}} \right)^{\frac{1}{b+1}}$ :

$$Q_s = P_n - (W_{max} - W) + (W_{max} - W_{min}) \left( \left( \frac{W_{max} - W}{W_{max} - W_{min}} \right)^{\frac{1}{b+1}} - \frac{P_n}{(b+1)(W_{max} - W_{min})} \right)^{b+1} \quad (3)$$

iii) if $P_n \ge (b+1)(W_{max} - W_{min}) \left( \frac{W_{max} - W}{W_{max} - W_{min}} \right)^{\frac{1}{b+1}}$ :

$$Q_s = P_n - (W_{max} - W) \quad (4)$$

*Vertical water exchange between layers*. The vertical fluxes between the three layers are calculated as follows: The downward fluxes between the layers are equal to the unsaturated hydraulic conductivity of the top layer [$LT^{-1}$]:

$$\begin{aligned} q_{1\to2}(t) &= k_1(s_1(t)) \\ q_{2\to3}(t) &= k_2(s_2(t)) \end{aligned} \quad (5)$$

where $s_1$ and $s_2$ are respectively degree of saturation of layer 1 and 2, defined as $s = (\theta - \theta_r)/(\theta_s - \theta_r)$, with $\theta$ average soil moisture content of the layer, and $\theta_s, \theta_r$ soil moisture at saturation and residual soil moisture respectively. This relates to average soil moisture for the entire grid cell, e.g. $\theta, \theta_s, \theta_r$ can be obtained by dividing $W$, $W_{max}$, $W_{min}$

by the thickness of the layer. Soil physical relationships for each layer per $0.5^0$x$0.5^0$ grid cell were obtained from the distribution of soil types and associated tabulated moisture retention and unsaturated conductivity values within the grid cell, both based on the digital soil map of the world [FAO, 1998]. First for each degree of saturation average matric potential and average conductivity was calculated. Next, the relationships according to Clapp and Hornberger [1978] were fitted:

$$k(s) = k_s s^{2\beta+3} \tag{6}$$

$$\psi(s) = \psi_s s^{-\beta} \tag{7}$$

with $k_s$ saturated hydraulic conductivity ($ms^{-1}$), $\psi$ matric potential (m) and $\beta$ is a dimensionless empirical exponent that varies on average between 4 and 11 over the range from sand to clay.

If the relative degree of saturation of the top layer is smaller than that of the underlying second store, an upward flux (capillary rise) can be sustained driven by the soil moisture deficit in the top layer and proportional to the unsaturated hydraulic conductivity of the second layer [$LT^{-1}$]:

$$q_{2\to1}(t) = \begin{cases} k_2(s_2)\cdot(1-s_1) & \text{if } (s_1 < s_2) \\ 0 & \text{if } (s_1 \geq s_2) \end{cases} \tag{8}$$

For the second layer, the capillary rise is described in a similar way, except that the conductivity is the geometric mean of that of the conductivity of the second and the third (groundwater layer), that it only occurs given the proximity of the water table, and that the resulting moisture content of the second layer cannot rise above field capacity $s_{fc}$ (with $\psi$= 1.0 m):

$$q_{3\to2}(t) = \begin{cases} \sqrt{k_{s,3}\cdot k_2(s_2)}\cdot(1-s_2)\cdot 0.5 f_5 & \text{if } (s_2 < s_{fc}) \text{ and } W_3 > 0 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

with $W_3$ [L] the water storage in the third (groundwater) layer. Here $f_5$ is the fraction of the $0.5^0\times0.5^0$ cell with a groundwater depth within 5 m. It is thus assumed that capillary rise is at maximum if the groundwater level is at the surface and 0 if it 5 m below the surface or below. The factor "0.5" then occurs as an estimate of the average capillary flux over the area fraction $f_5$ with a groundwater table within 5 m depth. The fraction $f_5$ itself is determined as follows: First all 1×1 km cells are determined within the $0.5^0\times0.5^0$ cell that belong to perennial drainage network (see section about *calculating drainage density* hereafter and Figure 7). For each of these cells the catchment (upstream drainage area) is determined. Next, taking the actual water levels (obtained from the surface water routing; see section 3.5) of the perennial stream cells as a reference, the groundwater height $H = W_3 / f$ (with $f$ drainable porosity or specific yield) is added to arrive at a local groundwater level and groundwater depth. From the groundwater depth distribution for each catchment one can determine the area with a groundwater depth smaller than 5 m.

Adding these areas for all catchments and dividing by the total $0.5^0 \times 0.5^0$ cell area gives the fraction $f_5$.

*Intermezzo: calculating drainage density.* In order to model the interflow component, the local groundwater level (see above) and the groundwater discharge (hereafter), an estimate of the drainage density (of perennial streams) for each grid cell is needed. Here we shortly describe the procedure used to obtain the estimate. We start out by using the Hydro1K data set [USGS Eros Data Center, 2006a] to delineate catchments with upstream areas of approximately 1000 km$^2$. Next, we use VMAP0 [FAO, 1997], a vector map of perennial water courses of the world, to estimate the average drainage density $D$ (total length of water courses divided by catchment area, m$^{-1}$). Here, the area of each catchment is reduced with the area of lakes, wetlands and reservoirs taken from the GLWD inventory [Lehner and Döll, 2004]. Because the drainage density is scale-dependent and does not include the smallest streams that have not been mapped, we rescaled the estimated densities such that for each combination of Holdridge life zone [Holdridge, 1967] and lithology [Dürr et al., 2005] the average density is in accordance with field-based estimates reported by Gregory [1976]. Finally, the catchment-derived drainage density is converted into a $0.5^0$x$0.5^0$ drainage density map. From the drainage density we can also calculate the average flow path length to the nearest water course $B = 1/2D$ [L]. Figure 2 shows the global map of this quantity.

*Interflow.* In mountainous areas soils develop in regolith or in colluvium on top of bedrock. The jump from high to low conductivity from soil to bedrock results in the occurrence of perched groundwater bodies during wet periods, which will cause a fast downslope flux of water through the soils down to the water courses. If this "subsurface storm flow" or "interflow" component is not modeled explicitly, recharge to deeper groundwater (i.e. layer 3) will be overestimated and as a consequence hydrological response time will be overestimated as well. Here, we model the interflow to occur in the second soil layer. Figure 3 provides a schematic of the modeling scheme. We use the simplified approach from Sloan and Moore [1984] to calculate interflow as follows:

$$Q_i(t) = \left[1 - \frac{\Delta t}{TCL}\right] Q_i(t - \Delta t) + \frac{\Delta t}{TCL} L[q_{12}(t) - q_{23}(t)] \tag{10}$$

where $Q_i$ is interflow per m slope width, [L$^2$T$^{-1}$], $q_{12}$, $q_{23}$ fluxes from/to first soil layer and groundwater layer [LT$^{-1}$] and $L$ the average slope length or drainage distance [L] as denoted in Figure 3 and $\Delta t$ is the discrete time step taken (1 day in our case). The parameter $TCL$ [T] is some characteristic response time that is given by:

$$TCL = \frac{L(\theta_s - \theta_{\mathrm{fc}})}{2k_{s,2} \tan \beta} \tag{11}$$

With $\theta_s$ saturated soil moisture content $\theta_{\mathrm{fc}}$ soil moisture content at field capacity, $k_{s,2}$ the saturated hydraulic conductivity of layer 2 and $\beta$ the average slope. The average slope is determined from the average of calculated slopes from 1×1 km digital elevation model

(Hydro1K) within the larger $0.5^0$x$0.5^0$ grid cell, excluding the 10% lowest Hydro1K elevations which are assumed to be part of the floodplain.
Interflow is only calculated if the soil water content in the second layer is above field capacity (as denoted in Equation 10). Moreover, we only want to apply the interflow routine to areas with steep slopes and bedrock, i.e. mountainous areas. As a proxy for these, we use the soil thickness as derived from the digital soil map of the world (FAO, 1979). For each $0.5^0$x$0.5^0$ grid cell we determined the fraction of soils with a soil depth smaller than 1.5 m (maximum soil depth in our model). For this fraction of the cell interflow is taken into account, while it is assumed zero for the remaining part.

### 3.3 Bare soil evaporation and transpiration

In case forcing is by ERA40, actual evaporation is given and only limited by the availability of soil moisture in the upper two layers. Based on a fraction of vegetation cover ERA40 evaporation is further fractioned into bare soil evaporation, which is drawn from the upper soil layer after deduction of any evaporation of liquid water stored in the snow cover, and transpiration by vegetation, which is drawn from both soil layers in proportion to the relative root volume present after deduction of evaporation of intercepted rainfall.
In case forcing is by CRU potential evaporation (section 3.2) a more elaborate scheme is required to calculate actual bare soil evaporation and transpiration. First, potential reference evaporation $E_0$ is converted into potential soil evaporation $ES_0$ and potential transpiration $T_0$ ( all in [$LT^{-1}$] ) as:

$$ES_0 = k_{c\,min} \cdot E_0 \tag{12}$$

$$T_0 = T_c - ES_0 = k_c \cdot E_0 - k_{c\,min} \cdot E_0 = \left(k_c - k_{c\,min}\right)E_0 \tag{13}$$

where $T_c$ is the monthly crop specific potential transpiration [$LT^{-1}$] and $k_c$ and $k_{c\,min}$ [-], are the monthly crop factor and the minimum crop factor for bare soil evaporation respectively.
Reductions of the potential bare soil evaporation and the transpiration are directly or indirectly related to the available soil moisture storage. For the bare soil evaporation, no reduction is applicable for the saturated fraction, $x$ of each cell as obtained by the Improved Arno Scheme of Hagemann and Gates [2003] (Equation 1), except that the rate of evaporation cannot exceed the saturated hydraulic conductivity of the upper soil layer. Likewise, the potential bare soil evaporation over the unsaturated area, 1-$x$, is only limited by the unsaturated hydraulic conductivity of the upper soil layer:

$$ES = x \cdot \min\left(k_{s1}, ES_0\right) + \left(1 - x\right)\min\left(k\left(s_1\right), ES_0\right) \tag{14}$$

The lack of aeration prevents the uptake of water by roots under fully saturated conditions. So transpiration only takes place for the unsaturated fraction of the cell. For the non-saturated area transpiration by plants depends on the total available moisture in the soil layers. Actual transpiration thus becomes:

$$T = (1 - x) f_T T_0$$

with $f_T$ the ratio of actual to potential transpiration rate given by:

$$f_T = \frac{1}{1 + \left(S_E / S_{E50}\right)^{-3\beta}} \tag{15}$$

where $S_{E50}$ is the average degree of saturation at which the potential transpiration is halved and $\beta$ is the coefficient of the soil water retention curve. Following the improved Arno scheme, the average degree of saturation over the unsaturated fraction of the cell is [Van Beek, 2008]:

$$S_E = \frac{W_{max} + b(W_{max} - W_{min})\left[1 - \frac{b+1}{b}\left(\frac{W_{max} - W}{W_{max} - W_{min}}\right)^{\frac{1}{b+1}}\right]}{W_{max} + b(W_{max} - W_{min})\left[1 - \left(\frac{W_{max} - W}{W_{max} - W_{min}}\right)^{\frac{1}{b+1}}\right]} \tag{16}$$

Similar to the case of ERA40 forcing, bare soil evaporation is reduced with evaporation from snow if snow is present, while transpiration is reduced by interception evaporation.

### 3.4. Groundwater storage and groundwater discharge

Groundwater discharge contributes an important part to river flow in many parts of the world, in particular during low flow conditions. In PCR-GLOBWB groundwater storage and discharge is modeled by a first order linear reservoir approach. The third (groundwater) reservoir is of infinite capacity. However, the active groundwater storage, which is the part of the groundwater that is being drained by surface water, is computed by assuming a linear relationship between storage and outflow, $q_b(t) = W_3(t)/J$, with $J$ a reservoir coefficient which represents the average residence time of water in the groundwater reservoir. By this approach the groundwater storage can be modeled by the following equation for finite time steps $\Delta t$:

$$W_3(t) = \left[1 - \frac{\Delta t}{J}\right] W_3(t - \Delta t) + q_{23}(t) \tag{17}$$

In all GHMs that take account of base flow component, a single reservoir coefficient $J$ is used whose value is obtained by the calibration. In PCR-GLOBWB we attempted to use information about reservoir properties to regionalize the groundwater residence time $J$. The parameterization is based on drainage theory by Kraijenhoff van de Leur [1958]. Figure 4 provides a schematic of a catchment with an aquifer below. Based on this schematic, an approximation of the reservoir coefficient $J$ is given by [Kraijenhoff van de Leur, 1958]:

$$J = \frac{\pi^2 k_{s,3} D_c}{4 f B_c^2} \tag{18}$$

where $k_{s,3}$ is the saturated hydraulic conductivity of the aquifer, $f$ the drainable porosity (or specific yield), $D_c$ the aquifer depth and $B_c$ the drainage length. The parameter $B_c$ is obtained from the drainage density analysis (see Figure 2). Hydraulic conductivity and drainable porosity have been related to a simplified version (7 classes) of the lithological map of the world [Dürr et al., 2007] and a literature search. Figure 5 shows the lithological map, while Table 1 provides for each class the mean, minimum and maximum values of hydraulic conductivity and drainable porosity. Unfortunately, there is no reliable information about aquifer thickness in relation to e.g. drainage distance and lithology. Awaiting better information about this in the future, we arbitrarily assumed the aquifer thickness to be a constant of 50 m, this being the order of magnitude of the groundwater in contact with the surface water at the time scale of our simulations (45 years). By crossing the drainage length map (Figure 2) with the lithological map (Figure 5) and using the values from Table 1, a global map of the groundwater residence time can be estimated through Equation (18). From this map (Figure 6) an average groundwater residence time $J$ can be estimated for each $0.5^0$x$0.5^0$ grid cell. This parameterisation can be used as an initial estimate of global residence time, which can be further calibrated by comparing models results with low flows from discharge data [e.g. Vörösmarty et al., 1998] and tuning $k_{s,3}$ and $f$ for each class between the maximum and minimum values in Table 1. This has not yet been done in this version of the model, but will be subject to future research.

.

### 3.5 Surface water and surface water routing

Surface water constitutes a separate land cover class for which the process descriptions differ from the land surface proper. The main processes included in the model are the direct input to or withdrawals from the open water due to precipitation, potential evaporation and water consumption and the resulting lake storage and river discharge. Fundamental to these process descriptions is the subdivision of surface water into a network with river and lake stretches as mentioned above.

In addition to the above direct input and losses, also the total runoff from the land surface is fed without delay to the river network prior to routing. Discharge is calculated from the kinematic wave approximation of the Saint-Venant Equation [Chow et al., 1988]. The continuity equation is:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = q\,, \tag{19}$$

and the momentum equation can be expressed as:

$$A = \alpha Q^{\beta} \tag{20}$$

where $Q$ is the streamflow through the channel [$L^3$ $T^{-1}$], $A$ is the channel cross-section [$m^2$], $q$ is the inflow per length of channel [$L^2$ $T^{-1}$] $x$ is the length along the channel [L] and $t$ is the elapsed time [T]. Combining Equations (19) and (20) gives

$$\frac{\partial Q}{\partial x} + \alpha\beta Q^{\beta-1}\left(\frac{\partial Q}{\partial t}\right) = q\,. \tag{21}$$

The coefficients $\alpha$ and $\beta$ are obtained from Manning's equation [Dingman, 1995]

$$Q = \frac{R^{2/3}\sqrt{S}}{n} A \tag{22}$$

where $R$ is the hydraulic radius [L], $S$ is the gradient, in this case equal to that of the bed [$L \cdot L^{-1}$], and $n$ is Manning's roughness coefficient [$L^{5/6} \cdot T^{-1}$]. Substitution of $R$ by $A/P$, where $P$ is the wetted perimeter [L] allows Equation (22) to be rewritten in terms of A:

$$A = \left(\frac{nP^{2/3}}{\sqrt{S}}\right)^{3/5} Q^{3/5}, \tag{23}$$

which gives the following values for $\alpha$ and $\beta$ in Equation (22):

$$\alpha = \frac{nP^{2/3\,\,3/5}}{S^{1/2}}\quad, \tag{24}$$

and

$$\beta = 3/5\,. \tag{25}$$

A numerical solution of the kinematic wave approximation is available as an internal function in PCRaster in which the new discharge $Q_{t+1}$ at every point along the channel is calculated from the discharge from the previous time step, the lateral inflow and the coefficients $\alpha$ and $\beta$ and passed over the local drainage direction (LDD) to the downstream cells. The lateral inflow in the channel, $q$, is calculated from the total drainage from the land surface and any direct inputs to the freshwater surface, $I_w$ ($m \cdot d^{-1}$):

$$q = \frac{Q_{tot}}{\Delta L} = \frac{A_{Cell}}{\Delta L}\left(\left[1 - Frac_{Water}\right]\sum_i Q_i + Frac_{water} \cdot I_w\right) \tag{26}$$

At the end of the time step, the calculated discharge is used to retrieve the new stage $h$ [L], which is calculated under the assumption of a rectangular channel with known channel depth and width (see hereafter). The new stage is passed to the next time step to estimate the wetted perimeter $P$ for the calculation of $\alpha$. Equation (21) is solved with an explicit scheme using variable time stepping, making sure that the time step is smaller than the Courant number everywhere.

Floodplains and throughflow wetlands (i.e. Niger inner delta) are treated as regular river stretches except that flooding spans the entire floodplain and experiences a higher resistance that is the weighted average of the Manning's $n$ over the river bed and that over the floodplain proper.
Like river stretches, lakes and reservoirs can evaporate freely at a rate set by the potential evaporation. Both are treated as a body of water with a variable water height, $h_{\mathrm{Lake}}$ [m] over its extent $A_{\mathrm{Lake}}$ [$L^2$]. For each time step, the change in water level is instantaneous over the lake surface and given by:

$$\Delta h_{Lake} = \left( \frac{\Delta Q + \Delta L \cdot q}{A_{Lake}} \right) \Delta t \tag{27}$$

where $\Delta Q$ is the net influx into the lake [$L \cdot T^{-1}$] and $q$ is the total volume of water per unit channel length (lake shore) that is added or subtracted from the surface water network [$L^2 \cdot T^{-1}$] which comprises precipitation, potential evaporation, drainage from the soil compartment and any consumptive water subtractions. These fluxes are multiplied by the duration of the time step, $\Delta t$, to obtain the change in lake water height [L]. The net influx is the balance of inflow and outflow if the lake interrupts the river network. The inflow is the incoming river discharge [$L^3\ s^{-1}$], the outflow is calculated in analogy to a weir formula as the discharge through a rectangular cross-section [Bos, 1989]:

$$Q_{Out} = C \cdot \tfrac{2}{3}\sqrt{\tfrac{2}{3} g}\, b (h - h_0)^{3/2} \Delta t \approx 1.70 C \cdot b (h - h_0)^{3/2} \Delta t \tag{28}$$

where is the breadth of the outlet [L], $g$ is the gravitational acceleration [$LT^{-2}$] and $h$ and $h_0$ are respectively the actual lake level and the sill of the outlet [L]. $C$ is a prefactor ($L^{4/3} \cdot T^{-1}$) that corrects among others for the effects of back- and tail waters, viscosity, turbulence and deviations from the assumed uniform flow distribution, which is kept at unity here. The calculated discharge at the outlet is added within the time step to the lateral inflow in the kinematic wave approximation of (Equation 21) of the downstream rivers.
To estimate the fraction of the open water surface (lakes and floodplains) we developed two options: 1) a fixed area option and 2) a variable area option.

*1) Fixed area option:*
In the fixed area option the areas of lakes and floodplains are kept constant while routing. It is a preferred option when faster calculation times are required. To derive estimates of surface water areas we used various sources: first, a statistical relationship was established between climate indicators and the observed bank-full discharge for 296 stations of the RivDis dataset [Vörösmarty et al., 1998]. This relationship was then extrapolated over the world and the channel depth and width calculated using hydraulic relationships [Allen et al., 1994]. In addition, floodplains, lakes and reservoirs were selected from the GLWD inventory [Lehner and Döll, 2004]. Floodplains were further constrained by including only those elevations (from the Hydro1k data set) that can be flooded from the river and thus partake in the runoff process. The local critical elevation was estimated from using the bank-full discharge of the 1×1 km perennial river cells (see

the section on drainage density above and Figure 8, left panel) as reference and adding 1.0 m. All 1×1 km cells in the catchments of the perennial river cells with an elevation below the critical elevation are assumed to be subject to flooding and therefore belong to the floodplain area. Finally, floodplain width was calculated by dividing floodplain area by channel length.

For the lakes, potential lakes and reservoirs the following selection criteria were used: i) a lake or potential lake was always selected if the cell of its location was a sink or an land-locked basin; exorheic lakes and reservoirs we only selected if their total area exceeded 500 $km^2$ or the mean fraction of the total cell area taken up by a lake or reservoir was at least 5%. Parts of any lake, potential lake or reservoir were excluded if the local area did not exceed the area of the floodplain plus the channel (channel and floodplain width times length, the latter being the length of the drainage path multiplied by a tortuosity of 1.3) By this procedure we include 78% of the reservoir surface area and 85% of lake surface area in the drainage network.

For the routing, the freshwater surface was either treated as a river stretch or lake stretch, with floodplains treated as river stretches of increased resistance. The drainage network itself was taken from the DDM30 dataset [Döll and Lehner 2002] to represent the connections between river stretches, but this network was broken up by a lake or a reservoir. These lakes were assigned a single outlet wherever the flow accumulation along the drainage path was the longest and, in the case of multiple points, where the elevation according to the GTOPO30 DEM [USGS EROS Data Center, 2006b] was the lowest. For the kinematic wave routing scheme we used a gradient that was derived from the Hydro1k dataset [USGS EROS Data Center, 2006a], the estimated total floodplain width and length and a weighed Manning's *n* that was based on values of 0.04 and 0.10 for the channel and floodplain respectively.

*2) Variable area option*

In the variable area option both floodplains and lakes change their area during routing. The option is the preferred one when the temporal variation of flooded areas is of importance. However, the option comes at a price, as the routing becomes much slower than with the fixed area option.

Variable floodplain inundation (see Figure 8): as a first step, for each $0.5^o \times 0.5^o$ cell the 1×1 km perennial river cells are selected. For all 1×1 cells belonging to the catchment (upstream drainage area) of a perennial river cell, the relative elevation is calculated by subtracting from the cell's surface elevation the surface elevation of the corresponding perennial river cell. At each time step during routing it is checked whether the volume of open water stored in a $0.5^o \times 0.5^o$ cell is above the storage at bank-full discharge. If this is the case, the water stored in excess of bank-full discharge is the volume of water that will flood the surrounding areas. This volume of water is distributed over the 1×1 km cells based on the cumulative distribution of relative elevations, from which follows the flooded area for that time step (Figure 8, right panel). This area is subject to open water evaporation during routing. By this scheme, at each time step it is possible to calculate the fraction of flooded area within each $0.5^o \times 0.5^o$ cell for each time step. Figures 9 and 10 provide impressions of fraction inundated areas by flooding.

Variable lake extent: For lakes we assume the following relation between lake volume $V_{lake}$ and lake area $A_{lake}$, which is based on the assumption that the form of the lake is a half-pyramid [Liebe et al., 2005]:

$$A_{Lake} = \left(3cV_{lake}\right)^{2/3} \tag{29}$$

with $c$ a form factor determined by average lake length and average lake depth. By keeping track of the volume $V_{lake}$ stored in the lake, changed in lake levels and lake extent can be calculated using Equations (27) and (29).

### 3.6 Surface water temperature and river ice

A somewhat exotic feature of PCR-GLOBWB is that it includes the possibility of solving the surface water energy balance, i.e. lateral latent heat transport. This enables one to calculate surface water temperature which can be useful for ecological and biogeochemical applications. Also, in case water temperatures fall below freezing point, the model has the possibility of calculating river ice formation thereby changing the hydraulic properties (wetted perimeter and Manning's coefficient) of the river. This enables one to model runoff blockage in for instance in arctic rivers.
In line with the general nature of the macro-scale hydrological model the energy stored in open water is modelled in a simple manner as the balance between the heat exchange with the atmosphere and the net advected energy assuming full mixing of the open water:

$$\frac{dE_w}{dt} = -\phi A_w - C_p \rho_w \frac{d(S_w T)}{dt} \tag{30}$$

where $E_w$ [J] is the total energy stored, $t$ is time [T], $\phi$ is the local heat flux at the water surface [W $L^{-2}$], $A_w$ the open water surface [$L^2$] $S_w$ is the total open water storage within the cell, [$L^3$], $T$ is the ambient temperature of the surface water [°C], $\rho_w$ is the density of water [M $L^{-3}$] and $C_p$ is the specific heat of water [4190 J $kg^{-1}$ $°C^{-1}$].
The water body exchanges energy with the atmosphere or with an ice cover if present (Figure 1). This exchange includes the following terms:

$$\phi = \phi_S + \phi_E + \phi_H + \phi_L \tag{31}$$

where $\phi_S$ is energy from incoming shortwave radiation, $\phi_E$ is the latent heat transfer, $\phi_H$ is the sensible heat transfer by convection and $\phi_L$ is the net longwave radiation.
Only the first three terms of Equation (31) are considered, as the energy flux due to long wave radiation heat transfer is very small compared to shortwave radiation [Hicks et al., 1995]. Energy fluxes due to friction and the ground flux across the river bed are of the same order and likewise neglected [Hicks et al, 1995; Ashton, 1986]. In the linear heat transfer approach, the heat transfer rate $\phi_H$ is approximated by:

$$\phi_H = K_H (T_1 - T_0) \tag{32}$$

where $T_1$ and $T_0$ represent respectively the temperatures over and below the surface [°C] and $K_H$ is the (turbulent) heat transfer coefficient [W $L^{-2}$ °$C^{-1}$]. If ice is present, which is assumed to cover the open water body entirely, the surface temperature is that of ice (*TT* or 0°C) and the heat transfer coefficient amounts to 8 W $m^{-2}$ °$C^{-1}$ [Van der Vinne, 1995]. Else, the air temperature is used and the coefficient set to 20 W $m^{-2}$ °$C^{-1}$. The energy flux due to short-wave radiation is given by:

$$\phi_S = (1-\alpha)\phi_{solar} \tag{33}$$

where $\alpha$ is the albedo of the surface [-] and $\phi_{solar}$ is the incoming shortwave radiation [W·$L^{-2}$]. Again, the albedo of water or ice is used dependent on the circumstances. If ice is present, the shortwave radiation heat transfer is directed to the ice cover rather than to the water body itself. The latent heat transfer, $\phi_E$, is neglected if an ice cover is present. For an open water surface it is prescribed by the potential evaporation rate:

$$\phi_E = \rho_w \lambda_v E_p \tag{34}$$

where $E_p$ is potential evaporation rate over the water surface [$LT^{-1}$], $\rho_w$ is the density of water, kept constant at 1000 kg·$m^{-3}$, and $\lambda_v$ is the latent heat of vaporization (2500 J·$kg^{-1}$). The advected energy is closely linked to the changes in the open water storage in the cell. It includes the following contributions (Figure 1):

- rain, ice melt and snow, the latter if no ice cover is present, which enter at respectively the air temperature and the melting point;
- surface runoff, which enters at the air temperature ($Q_1$) and at that of the soil for deeper drainage ($Q_2$ *and* $Q_3$). The default soil temperature is the mean annual temperature, but in case the soil energy balance is solved a calculated soil temperature could be used;
- if so desired, consumptive water use is extracted at the water temperature at the start of the current time step;
- in- and outflow for each segment of the drainage network which enter at the water temperature of the upstream area and leave at the local water temperature.

It should be noted that evaporation and ice growth do not alter the energy budget as they are already incorporated in the surface energy flux, $\phi$. However, they are of influence on the change in water storage, $\Delta S_w$, within the cell.

In the case of cooling, the surface energy flux is limited to 90% of the total available energy, implying that the water temperature cannot drop below freezing point. In the case of warming, the water temperature cannot rise above the ambient air temperature. In practice this will only affect those cells where the open water storage is small. To ensure that no energy is lost or added due to the lateral fluxes, first the surface heat flux and the advected energy within the cell are evaluated and converted in a new water temperature at which the lateral fluxes are routed along the local drainage direction (LDD). The energy budget is then re-evaluated for these new changes and together with the new water temperature passed on to the next time step. To avoid numerical instability, changes in temperature and discharge, are evaluated at a higher temporal resolution that is fixed at 4-6-hourly time steps by default, although this number may be increased if required.

The water energy balance is evaluated together with that of a potential or actual ice cover. Ice will form or grow if the air temperature is below freezing point and is not balanced by

the heat flux coming from the water or by the incoming shortwave radiation ($\phi_i < 0$). The temperature of the ice is kept at 0°C throughout and any additional cooling results in an increase in ice thickness and a reduction in water temperature. The change in ice thickness, $\Delta z_i$ [m], as a result of direct heat inputs to the ice cover over a time step is calculated by:

$$\Delta z_i = -\frac{\phi_i \Delta t}{\rho_w \lambda_f} \tag{35}$$

where $\Delta t$ is the length of the time step [T], $\rho_w$ is the density of water (kept constant at 1000 kg·m$^{-3}$) and $\lambda_f$ is the latent heat of fusion of ice, 333.4 kJ·kg$^{-1}$. The change in ice thickness may be limited by the available stores (ice in the case of thawing, water depth in the case of freezing).

The formation of ice modifies the discharge and is assumed to occur prior to routing in the discretized time space. Discharge is affected by the presence of ice in the following ways: prior to the calculation of the wetted perimeter the change in water height due to ice growth or decay is evaluated. If an ice cover is present, the wetted perimeter $P$ of the rectangular channel itself is modified from ($h$ water height [L], $W$ channel width [L]) $P = 2h + W$ to $P = 2(h + W)$. This ice surface is often rough with typical Manning's $n$ values between 0.01 and 0.10 and a composite Manning's roughness coefficient, $n_c$, is calculated for the ice-covered channel [USACE, 2002]:

$$n_c = \left( \frac{n_i^{3/2} + n_b^{3/2}}{2} \right)^{2/3} \tag{36}$$

where $n_i$ and $n_b$ are respectively the Manning's n values for the ice cover and the bed respectively. Power functions have been used to establish empirical relationships between water height, ice thickness and roughness coefficients [Nezhikhovskiy, 1964]. These show a decreasing roughness with water height and an increasing roughness with ice thickness for different ice types (thick and thin ice jams and frozen-up covers with decreasing roughness). These functions for $n_i$ can be approximated by a general function which is applied here:

$$n_i = 0.0493 h^{-0.23} z_i^{0.57} \tag{37}$$

The influence of ice on river discharge is considerable. If one assumes that the flow area remains constant and that bed and ice roughness are equal, the increase in the wetted perimeter reduces the hydraulic radius by 50% if the width of the channel is large compared to the flow depth, thus resulting in a 37% decrease in flow velocity.

### 3.7 Reservoirs and reservoir operations

*Reservoirs*

Most of the world's major rivers are regulated by artificial reservoirs [Vörösmarty et al., 2004]. Since the total storage capacity of reservoirs (7000 km$^3$ globally) comprises three times the annual average water storage in river channels (1200-2120 km$^3$) and one-sixth of the global annual river discharge [40,000 km$^3$·a$^{-1}$; Baumgartner and Reichel, 1975], the effect of reservoir operations on river discharge is not negligible. In particular if water

availability is to be calculated for global water resources assessments [Arnell, 2004] it is important to include reservoir capacity and possibly also reservoir operations.
The routing scheme of PCR-GLOBWB contains 513 reservoirs selected from the 654 reservoirs contained by the GLWD1 dataset of Lehner and Döll [2004], which incorporates and adds to the information from the WWDR-II and WRD datasets [Vörösmarty et al., 1997, ICOLD, 2003] for the world's largest reservoirs (storage capacity $\geq$ 0.5 km3). The GLWD1 dataset was preferred as it also lists the upstream area for most reservoirs, thus allowing for an accurate positioning of the reservoir on the drainage network. The reservoirs included were the regional reservoirs only, i.e. those reservoirs that are part of the larger drainage network having both a river inflow and a river outflow. Local upstream reservoirs are thus not included. The selected reservoirs contain 95% of the total reservoir capacity of 4446 $km^3$ as specified by the GLWD1 of Lehner and Döll [2004].
Reservoirs can be viewed as a particular kind of lakes, their main characteristic being that the outflow is controlled to meet certain requirements. Similar to lakes, reservoirs therefore occupy a contiguous area of multiple cells, identified by the same identification number, and have a single outlet at which the reservoir connects to the downstream network. Other than lakes, however, reservoirs do not have any substantial inert storage (storage that does not take part in the routing).

*Reservoir operation strategy*
In contrast to existing schemes [e.g. Haddeland et al., 2006; Hanasaki et al., 2006] the reservoir operation scheme within PCR-GLOBWB is prospective rather than retrospective. Retrospective schemes optimize reservoir operation for a certain period on the basis of the simulated discharge and demand, thus ensuring that a reservoir performs near-perfect depending on its purpose. In contrast, a prospective scheme has to work with forecasts of future inputs and demands, a reality that confronts reservoir operators on a daily basis. Also, since the prospective reservoir operation scheme is directly implemented in the routing model the influence of upstream reservoirs on the operation of downstream ones is better resolved than in retrospective schemes that are inherently run as a post-processing step on simulated discharge.
The overall modelling strategy of the prospective reservoir operation scheme is to determine the target storage over a defined period ensuring its proper functioning given the forecasts of inflow and downstream demand. Target storage rather than outflow is used as this has to be updated when actual inflow and demand differ from the forecasts. These updates are carried through at the daily time step on which the routing scheme is run rather than at the monthly scale for which the forecasts in the form of past average values are available. The approach applies to single reservoirs although downstream reservoirs are influenced by the operations at upstream dams. Similar to Haddeland et al. [2006] four reservoir types are distinguished, being water supply, including irrigation, flood control, hydropower generation and navigation (Table 2).

*Forecasts of inflow and demand*
In PCR-GLOBWB reservoir operations are evaluated over the *operational year*, which starts with the month that the inflow falls below the mean annual value [Hanasaki et al., 2006]. The monthly inflow for the coming operational year $\hat{Q}_{m\,y+1}$ is predicted by:

$$\hat{Q}_{m\,y+1} = w\overline{Q}_{m\,y-1} + (1-w)Q_{m\,y} \tag{38}$$

With $Q_{m\,y}$ the inflow into the reservoir [$m^3 \cdot s^{-1}$] of the same month the previous year, $\overline{Q}_{m\,y-1}$ the average inflow of the particular month averaged over a *N*-1 retrospective years, $w = 1/(N\text{-}1)$ is a weight [0-1] determining the size of the averaging window, and the indices *m* and *y* denote respectively the months (1 through 12) and the elapsed model years. Thus, the prospective monthly discharge for the coming operational year approximates a moving-average over the past *N* years. The inflow comprises all incoming discharge, local gains or losses over the reservoir surface and any local freshwater abstractions (currently set to zero). In a similar manner, forecasts of downstream water demand are obtained for each reservoir, where downstream water demand is calculated according to Wada et al. [2009] (See below for details on how downstream demand is allocated to reservoirs).

*Optimization of reservoir operations*

Given the storage at the start of the operational year and the forecasted inflow and demand, the objective is to find the release and corresponding storages that would ensure optimum functioning of the reservoir over the upcoming operational year. Since long-term average values are used it may be postulated that this operation scheme would also result in optimum functioning of the reservoir for all following years. The target reservoir storage at the end of the upcoming operational year is thus defined as optimal for the coming year and the year thereafter, i.e. for the next 24 months.

Reservoir operations are optimized using the following conditions and constraints: For the reservoir purposes, we use the optimization criteria of Haddeland et al., [2006] (Table 2). If a reservoir had multiple purposes, the purposes are weighed proportionally to their ranking (from ICOLD [2003]) with overall precedence being given to water supply. Flood damages are expected if the discharge exceeded bank-full discharge, which is approximated as the mean annual maximum daily flood [Haddeland et al., 2006]. For hydropower generation, we include the option to weigh the objective function by price as proposed by Adam et al. [2007], but kept this constant in absence of information. As for the natural lakes, the relationship between reservoir height and storage is given by a theoretical relationship according to Liebe et al. [2005]:

$$V = \frac{1}{3} A d \tag{39}$$

where $A$ is the surface area of the reservoir [$m^2$], which is given by $l^2/2$, with $l$ being the characteristic length of the reservoir [m], and the maximum depth, $d$ [m]. This results in a change in surface area with changing storage, which improves the estimate of the direct gains and losses (by precipitation and evaporation) over the reservoir, and a reservoir level, $h$, that is required in case of hydropower generation. According to the objective function for hydropower generation, (see Table 2), we estimate the monthly averaged $h$ as the harmonic mean at the start and end of each month.

In addition to the objective functions, there is the practical constraint that the reservoir capacity may never be exceeded and sufficient capacity must be reserved to accommodate excessive discharge. Also, sufficient storage should be kept in reserve to safeguard a certain minimum release. To define these levels, the 7-day averaged maximum $Q_{\max}$ and minimum $Q_{\min}$ discharges were used. This led to the overall formulation for prediction of the next reservoir storage, $S_{m+1}$ [$m^3$]:

$$S_{m+1} = S_m + n\left\{\left(\hat{Q}_m - Q_{rm}\right)\Delta t + \hat{q}_{wm} A_m\right\} \tag{40}$$

where $n$ is the number of days in the given month, $\hat{q}_w$ is the forecasted net gain or loss [m·day$^{-1}$] over the reservoir surface area $A$, which is the average for the values at times $m$ and $m$+1. While the following constraints should be met ($C$ the maximum reservoir capacity):

$$S_{m+1} > \sum_{m+1}^{12} n Q_{\min} \Delta t \tag{41a}$$

$$C - S_{m+1} > 7 Q_{\max} \Delta t \tag{41b}$$

Optimization of reservoir operation requires for each reservoir the optimization of 12 parameters, i.e. the releases for the coming 12 months. To keep this problem tractable, optimization is restricted to the beginning of the operational year (when inflow becomes smaller than release) and the beginning of the recharge period (when inflow is larger than release). For these periods the variation in monthly outflows is assumed to follow the following harmonic function:

$$Q_{rm} = \frac{Q_{r,\max}}{2} - \frac{Q_{r,\max}}{2} \cos\left(\frac{2\pi(m - m_b)}{(m_e - m_b)}\right) \tag{42}$$

where $m_b$ and $m_e$ are respectively the beginning and the end of the operational and recharge and $Q_{r,\max}$ the maximum release during the period. This way, for every year and every reservoir only 2 parameters need to be optimized ($Q_{r,\max}$ for the release and recharge period), instead of 12. Incidentally, the start of the operational year is re-evaluated when the start of the release period is observed (through Equation 38) to change in time.

*Allocation of downstream demand to reservoirs*

Release from a reservoir can only meet the demand in cells that are situated downstream from the reservoir, have an elevation that is less than that of the dam position and that could be reached within 7 days with an average discharge velocity of 1 m·s$^{-1}$ (equivalent with 600 km). As an additional constraint, we limited the supply to cells that were located in the same country as the dam. Cells that receive water from multiple reservoirs have

their demand weighed by reservoir capacity irrespective whether these reservoirs are located on the same river course or on different tributaries.

*Adjusting optimized releases on a monthly and daily basis*

The optimized releases $Q_{rm}$ also lead to monthly target storages $S_m$ through Equation (40) As daily values of inflow and actual storage become available, actual storage may start to deviate from the originally target storages. However, given the imperfect knowledge for the remainder of the operational year, reservoir operation can still be expected to be optimal if the final target storage is met. Thus, the daily water releases from a reservoir are modified in an attempt to meet the specified target storage.

$$Q'_{rm} = \max\left(0, \frac{S_{m,i-1} - S_{m+1}}{(n-i)\Delta t} + \hat{Q}_m\right) \tag{44}$$

where $Q'_{rm}$ is the updated daily release for the remainder of the month. Furthermore daily release is additionally modified as follows: 1) a subsequent reduction in case reservoir storage is below a minimum (10% of maximum storage) or in case actual demand to date is below original monthly forecast; 2) a subsequent increase if actual demand to date is above the original monthly forecast, hoping to recuperate the extra release later in time. Obviously, still the following conditions also apply to the updated daily reservoir release: it may not exceed $Q_{flood}$ to ensure downstream safety and the minimum flow $Q_{min}$ should be warranted during the remainder of the operational year.

*Preliminary results of application of the scheme*

Some preliminary examples of reservoir operation scheme are shown in Figure 11. The results of the reservoir scheme are compared with the scheme of Hanasaki et al. [2006], which is a retrospective reservoir scheme. Figure 12 shows the effect of reservoir operation of discharge range (relative to mean yearly discharge) and the timing of the downstream discharge for North-America. It can be seen that the reservoir operation has indeed a large effect on downstream discharge variability.

## 4. Verification with runoff data

The following sections provide a comparison of simulated discharge data from PCR-GLOBWB with runoff data. PCR-GLOBWB was run for the period 1958 to 2000 using CRU TS 2.1 [New et al., 2000] monthly meteorological time series as input, downscaled to daily values with the ERA-40 re-analysis data [Uppala et al., 2005] (see section 3.1). Simulated daily runoff was aggregated to monthly values and compared to observed monthly discharge for a large number of catchments, as obtained from either the Global Runoff Data Centre [GRDC, 2004] or the Global River Discharge Database were used [Vörösmarty et al., 1998]. We stress that we have performed no model calibration at this time.

Note: this section is under construction. We are in the process of further validating the model, both with and without the reservoir scheme on. Below, lists are presented of the various verification tables and figures of yearly, monthly, daily and multi-year runoff. At this point we present only a validation of yearly runoff and only part of the monthly runoff, both still with an older version of the model without the reservoir scheme. We will add new results as they become available.

Results that are present are given in *italics.*

### 4.1 Yearly average runoff

- *Bar chart of yearly discharge of 99 major rivers from the GRDC and RivDis data set (Figure 13)*
  This shows that the order of magnitude of simulated runoff is correct throughout the globe, except for snow melt dominated rivers where the model shows an under-estimation of discharge (probably attributable to snow under catch of the CRU data set) and the overestimation of discharge of the African rivers (most probably attributable over-estimation of precipitation and under-estimation of open water evaporation). Also, rivers with large extractions and diversions such as the Nile and the Colorado are over-estimated.
- *Table with runoff of continents, compared with previous results (Table 3)*
  This table shows that the continental runoff of PCR-GLOBWB is comparable to the average of runoff produced by other studies, except for Oceania, where higher runoff is simulated. Results are overall very similar to those produced by Döll et al. [2003].
- *Composite map with average yearly discharge and average soil moisture and snow cover on December 1$^{st}$ (Figure 14)*
- *Figure of 19 major rivers, including bar chart of discharge(Figure15, 16)*
  The major rivers are selected to represent different continents and climate zones. Figure 16 shows that average yearly runoff is well reproduced for the selected rivers. Also shown are the bar charts where GRDC discharge is corrected for water use using the downstream water demand as a proxy. For the Murray river this yields a large improvement in predicted discharge.

### 4.2 Monthly Runoff

- *One figure with regime graphs of Runoff of 19 major rivers (Figure 17)*
  Monthly runoff is reproduced reasonably well, with the exception of the arctic rivers (see also Figure 13) where the model shows an under-estimation of peak-discharge and the overestimation of discharge of the African rivers (see also Figure 13). Timing of the runoff peaks are well reproduced for most rivers. Generally, results are better for rivers in Eurasia and North-America, where the meteorological forcing is better constrained by observations.
- *One figure with hydrographs of monthly runoff of 6 major rivers*: *Rhine, Danube*, Ob, *Mississippi*, Amazone, Bamaputhra (Figure 18)
  As can be seen, the monthly runoff of the Rhine and Danube are well reproduced, while the Mississippi shows a good fit, except for the wetter years.

- Figure with histograms of R2 and NS monthly runoff with and without reservoirs for all rivers (Figure 19)

### 4.3 Multiyear runoff

- One Figure with hydrographs of yearly runoff of 6 major rivers: Rhine, Danube, Ob, Mississippi, Amazone, Bamaputhra (Figure 19)
- Figure with histograms of R2 and NS monthly runoff with reservoirs for all rivers (Figure 20)

## 5. Further unverified results

*One figure (Figure 21) with maps of*

- *Soil moisture (January 1)*
- Soil moisture (July 1)
- *Snow depth (January 1)*
- Snow depth (June 1)
- *Yearly average groundwater recharge (mm/year)*
- *Baseflow (mm/year)*
- *Water temperature (January 1)*
- Water temperature (July 1)
- *River ice (January 1)*
- *River Ice (July 1)*

## 6. Conclusions and further research

In this paper we introduced a next generation global-scale hydrological model (PCR-GLOBWB) that is suitable for describing global hydrology at a monthly time-scale. Compared to other global hydrological models PCR-GLOBWB introduces new schemes for subgrid parameterization of surface runoff, interflow and baseflow and adds explicit routing of surface water flow using the kinematic wave approximation, dynamic inundation of floodplains and a reservoir operations scheme. PCR-GLOBWB also contains a routine for lateral water transport of latent heat from which water temperature and river ice thickness can be calculated. A first verification of the uncalibrated model using global runoff data shows that PCR-GLOBWB is able to reproduce yearly runoff for most of the global rivers. Also, monthly runoff is reproduced reasonably well, with the exception of the arctic rivers where the model shows an under-estimation of peak-discharge (probably attributable to snow under catch of the CRU data set) and the overestimation of discharge of the African rivers (most probably attributable over-estimation of precipitation and under-estimation of open water evaporation). Timing of the runoff peaks are well reproduced for most rivers. Generally, results are better for rivers in Eurasia and North-America, where the meteorological forcing is better constrained by observations.

Future work on the model will be focused on model improvement by local tuning of the groundwater residence-time parameters and the snow-melt parameters. Also, additional verification exercises are necessary to further constrain the model for other output variables. This includes comparing simulated total terrestrial water storage with observations from the GRACE satellites [Syed et al., 2008], comparison of soil moisture with remotely sensed soil moisture products [Wagner et al, 1999], lake levels with satellite altimetry [Birkett, 1995], remotely sensed snow cover [Immerzeel et al., 2008], composite remote sensing products providing spatial time series of fraction inundated area [Prigent et al., 2007] and water temperature data [www.gemstat.org].

**References**


Adam, J.C., I. Haddeland, F. Su, and D.P. Lettenmaier (2007), Simulation of reservoir influences on annual and seasonal streamflow changes for the Lena, Yenisei and Ob rivers, *Journal of Geophysical Research*, 112, D24114, doi:10.1029/2007JD008525.

Alcamo, J., Henrichs, T. & Rösch, T. (2000), *World water in 2025—global modeling and scenario analysis for the World Commission on Water for the 21st Century*. Tech. Report, Centre for Environmental Systems Research, University of Kassel, Kassel, Germany.

Allen, P.M., J.C. Arnold and B.W. Byars (1994), Downstream channel geometry for use in planning-level models, *Journal of the American Water Resources Association 30*, 663-671.

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

Arnell, N.W. (2004), Climate change and global water resources: SRES emissions and socio-economic scenarios, *Global Environmental Change*, *14*, 31–52.

Ashton, G.D. (1986), *River and Lake Ice Engineering*, Water Resources Publications, Littleton, Colorado.

Baumgartner, A. and E. Reichel (1975), *The World Water Balance, Mean Annual Global, Continental and Maritime Precipitation, Evaporation and Run-off*, Elsevier, Amsterdam

Bergström, S., 1976, The HBV Model. In: *Computer models of watershed hydrology*. (V. P. Singh ed.),Colorado, Water Resources Publications, 443-476.

Birkett C.M. The contribution of TOPEX/ POSEIDON to the global monitoring of climatically sensitive lakes J. Geophys. Res., 1995, 100(C12), P. 25179–25204.

Bos, M.G. (1989), *Discharge measurement structures, third revised edition*, ILRI Wageningen.

Brankoviç, C. and F. Molteni (2004), Seasonal climate and variability of the ECMWF ERA-40 model, *Climate Dynamics 22*, 139-155.

Chow, V.T., D.R. Maidment and L.W. Mays (1988), *Applied Hydrology*, McGraw-Hill, New York.

Clapp, R.B. and G.M. Hornberger (1978), Empirical equations for some soil hydraulic properties, *Water Resources Research 14*, 601-604.

Dingman, S.L. (1994), *Physical Hydrology*, Prentice-Hall, Upper Saddle River, New Jersey.

Dirmeyer, P. A., A.J., Dolman and N. Sato (1999), The pilot phase of the Global Soil Wetness Project. *Bulletin of the American Meteorological Society 80*, 851–878.

Dirmeyer, P.A., X. Gao, M. Zhao, T. Oki and N. Hanasaki (2006), The second global soil wetness project (GSWP-2), *Bulletin of the American Meteorological Society 87*, 1381-1397.

Döll, P. and B. Lehner (2002), Validation of a new global 30-minute drainage direction map. *Journal of Hydrology 258*, 214-231.
.
Döll, P.A. and S. Siebert (2002), Global modeling of irrigation water requirements, *Water Resources Research 38*, 8.1-8.10.

Döll, P., F. Kaspar and B. Lehner (2003), A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology* 270, 105-134.

Dürr H.H., M. Meybeck and S. Dürr (2005), Lithologic composition of the Earth's continental surfaces derived from a new digital map emphasizing riverine material transfer, *Global Biogeochemical Cycles 19*, GB4S10, doi:10.1029/2005GB002515.

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

Fekete B.M., C.J. Vörösmarty and W. Grabs (2002), High-resolution fields of global runoff combining observed river discharge and simulated water balances, *Global Biogeochemical Cyles 16*, 15.

Fields Development Team (2006). fields: Tools for Spatial Data. National Center for Atmospheric Research, Boulder, CO. URL http://www.cgd.ucar.edu/Software/Fields

FAO (1997), VAMP0, Perennial Inland Water Areas of the World, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy, http://www.fao.org/geonetwork/srv/en/main.search.

FAO (1998) *Digital Soil map of the World*, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.

GRDC (Global Runoff Data Centre) (2004), 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/ (accessed 15.02. 2008).

Gregory, K. J. (1976), Drainage networks and climate, In: *Geomorphology and Climate* (E. Derbyshire, ed.), pp. 289–315, Wiley-Interscience, Chichester, UK.

Güntner, A., J. Stuck, S. Werth, S, P. Döll, K. Verzano, and B. Merz (2007), A global analysis of temporal and spatial variations in continental water storage, *Water Resources Research* 43, W05416, doi: 10.1029/2006WR005247.

Haddeland, I., T. Skaugen and D.P. Lettenmaier (2006), Anthropogenic impacts on continental surface water fluxes, *Geophysical Research Letters 33*, L08406, doi:10.1029/2006GL026047, 2006b.

Hagemann, S., M. Botzet, L. Dümenil and B. Machenhauer (1999), *Derivation of global GCM boundary conditions from 1 km land use satellite data,* Max-Planck-Institute for Meteorology, Report 289, Hamburg, Germany,

Hagemann, S. (2002), *An improved land surface parameter dataset for global and regional climate models*, Max-Planck-Institute for Meteorology, Report 336, Hamburg.

Hagemann, S. and L.D. Gates (2003), Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations, *Climate Dynamics 21,* 349–359.

Hanasaki, N., S. Kanae and T. Oki (2006), A reservoir operation scheme for global river routing models, *Journal of Hydrology 327*, 22– 41.

Hicks, F., W. Cui and D. Andres (1997), Modelling the thermal break up on the Mackenzie River at the outlet of Great Slave Lake, N.W.T. Canadian *Journal of Civil Engineering, 24*, 570-585.

Holdridge, L.R. (1967), *Life Zone Ecology*, Tropical Science Center, San Jose, Costa Rica.

ICOLD (2003), *World Register of Dams*, International Commission on Large Dams (ICOLD), Paris.

Immerzeel, W.W., Droogers, P., de Jong, S.M., Bierkens, M.F.P., 2009, Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing Remote Sensing of Environment 113: 40-49.

Korzun, V.I., A.A. Sokolow, M.I. Budyko, K.P. Voskresensky, G.P. Kalininin, A.A. Konoplyanstev, E.S. Korotkevich, P.S. Kuzin and M.I. Lvovich (Eds.) (1978), World Water Balance and Water Resources of the Earth, *Studies and Reports in Hydrology 25*, UNESCO, Paris.

Kraijenhoff van de Leur, D.A. (1958), A study of non-steady ground-water flow with special reference to the reservoir-coefficient, *De Ingenieur 19*, 87-94.

Leemans, R. (1989) *Global Holdridge Life Zone Classifications, Digital Raster Data on a 0.5-degree Geographic (lat/long) 360x720 grid*, IIASA, Laxenberg, Austria, http://www.ngdc.noaa.gov/ecosys/cdroms/ged_iia/datasets/a06/lh.htm#top

Lehner, B. and P. Döll (2004), Development and validation of a global database of lakes, reservoirs and wetlands, *Journal of Hydrology 296*, 1–22.

Liang, X., D.P. Lettenmaier, E.F. Wood and S.J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for GSMS,. *Journal of Geophysical Research 99*, 14,415-14,428.

Liebe J., N. van de Giesen and M. Andreini (2005), Estimation of small reservoir storage capacities in a semi-arid environment. Case study in the Upper East Region of Ghana, *Physics and Chemistry of the Earth 30*, 448–454.

L'vovich, M.I. (1979), *World Water Resources and Their Future*, AGU, Washington DC (English translation of Russian original published in Moscow in 1974).

Milly, P.C.M. and A.B. Schmakin (2002), Global modeling of land water and energy balances, Part I: the land dynamics (LaD) model. *Journal of Hydrometeorology* 3, 283-299.

Milly, P. C. D., K. A. Dunne and A. V. Vecchia (2005), Global pattern of trends in streamflow and water availability in a changing climate, *Nature 438*, 347-350.

New, M., M. Hulme and P. Jones (1999), Representing twentieth-century space-time climate variability. Part I: Development of a 1961-1990 mean monthly terrestrial climatology, *Journal of Climate 12*, 829-856.

New, M., M. Hulme and P. Jones (2000), Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate, *Journal of Climate 13*, 2217-2238.

New, M., D. Lister, M. Hulme and I. Makin (2002), A high-resolution data set of surface climate over global land areas, *Climate Research 21*, 1-25.

Nezhikovskiy, R.A. (1964), Coefficient of Roughness of Bottom Surfaces of Slush Ice Cover. *Soviet Hydrology, Selected Papers, 2*, 127-150.

Nijssen, B., G.M. O'Donnel, D.P., Lettenmaier, D. Lohmann and E.F.Wood (2001), Predicting the discharge of global rivers, *Journal of Climate 14*: 3307-3323.

Oki, T., Y. Agata, S. Kanae, T. Saruhashi, D. Yang and K. Musiake (2001), Global assessment of current water resources using total runoff integrating pathways. *Hydrological Sciences Journal 46*, 983-995.

Olson, J.S., (1994a), *Global ecosystem framework-definitions*, USGS EROS Data Center Internal Report, Sioux Falls.

Olson, J.S., (1994b), *Global ecosystem framework-translation strategy*, USGS EROS Data Center Internal Report, Sioux Falls.

Prigent, C., F. Papa, F. Aires, W. B. Rossow, and E. Matthews (2007), Global inundation dynamics inferred from multiple satellite observations, 1993–2000, *Journal of Geophysical Research 112*, D12107, doi:10.1029/2006JD007847.

Sheffield, J. and E.F. Wood (2007), Characteristics of global and regional drought, 1950-2000: analysis of soil moisture data from off-line simulation of the terrestrial hydrological cycle. *Journal of Geophysical research 112*, D17115, doi:10.1029/2006JD008288.

Shiklomanov, I.A. (Ed.) (1997), *Assessment of water resources and water availability in the world*, Comprehensive assessment of the freshwater resources of the world. WMO (World Meteorological Organization) and SEI (Stockholm Environmental Institute)

Sloan, P.G. and I.D. Moore (1984), Modeling subsurface stormflow on steeply sloping forested watersheds, *Water Resources Research 20*, 1815–1822.

Syed, T. H., J. S. Famiglietti, M. Rodell, J. Chen, and C. R. Wilson, 2008, Analysis of terrestrial water storage changes from GRACE and GLDAS, Water Resour. Res., 44, W02433, doi:10.1029/2006WR005779.

Troccoli, A.and P. Kållberg (2004), *Precipitation correction in the ERA-40 Reanalysis*, report European Centre for Medium Range Weather Forecasts, Reading UK.

Uppala, S.M. et al. (2005), The ERA-40 re-analysis, *Quarterly Journal of the Royal Meteorological Society 131*, 2961–3012.

USACE (2002), Chapter 4: Hydraulic Computations and Modelling of Ice-Covered Rivers, Engineering and Design, In: *Ice Engineering*, Engineer Manual 1110-2-1612, United States Army Corps of Engineers.

USGS EROS Data Center (2002), *Global land cover characteristics data base version 2.0*, http://edcdaac.usgs.gov/glcc/globedoc2_0.html

USGS EROS Data Center (2006a), *HYDRO1k Elevation Derivative Database*, LP DAAC: http://edcdaac.usgs.gov/gtopo30/hydro/

USGS EROS Data Center (2006b), *Global 30 Arc-Second Elevation Dataset (GTOPO30)*, http://eros.usgs.gov/products/elevation/gtopo30/gtopo30.html

Van Beek, L.P.H. (2008), *Forcing PCR-GLOBWB with CRU data*, Report Department of Physical Geography, Utrecht University, Utrecht, The Netherlands, http://vanbeek.geo.uu.nl/suppinfo/vanbeek2008.pdf

Van Der Vinne, P. G. (1995), *Deterioration of an ice cover on a small pond*, Alberta Cooperative Research Program in Surface Water Engineering Report No. SWE-95/01.

USGS: US Geological Survey, 2000. HYDRO1k Elevation Derivative Database. USGS, EROS Data Center. Available at http://edcdaac.usgs.gov/gtopo30/hydro/.

Vörösmarty, C.J., Sharma, K.P., Fekete, B.M., Copeland, A.H., Holden, J., Marble, J. and Lough, J.A., 1997. The storage and aging of continental runoff in large reservoir systems of the world. Ambio 26 4, pp. 210–219.

Vörösmarty, C.J., B.M. Fekete and B.A. Tucker (1998), *Global River Discharge,1807–1991, Version 1.1 (RivDIS)*, Oak Ridge National Laboratory Distributed Archive Center, Oak Ridge, Tennessee, http://www.daac.ornl.gov]

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

Vörösmarty, C.J., M. Meybeck, B.M. Fekete, K. Sharma, P. Green and J. Syvitski. (2003), Anthropogenic sediment retention: Major global-scale impact from the population of registered impoundments, *Global and Planetary Change 39*, 169-190.

Vörösmarty, C.J., D. Lettenmaier, C. Leveque, M. Meybeck, C. Pahl-Wostl, J. Alcamo, W. Cosgrove, H. Grassl, H. Hoff, P. Kabat, F. Lansigan, R. Lawford and R. Naiman (2004), Humans transforming the global water system, *EOS Transaction of the American Geophysical Union 85*, 509-520.

Vose, R.S., R.L. Schmoyer, P.M. Steurer, T.C. Peterson et al. (1992), The global historical climatology network: Long-term monthly temperature, precipitation, sea level pressure, and station pressure data. *CDIAC Communications 17*, 12.

Wada Y. (2008), *Water Stress over the Year: Quantitative Analysis of Seasonality and Severity on a Global Scale*, MSc-thesis, Faculty of Earth Sciences, Utrecht University.

Wagner, W., G. Lemoine, H. Rott (1999) ,A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data, *Remote Sensing of Environment 70*, 191-207.

Wesseling, C.G., D. Karssenberg, W.P.A. van Deursen and P.A. Burrough (1996), Integrating dynamic environmental models in GIS: the development of a Dynamic Modelling language, *Transactions in GIS* 1, 40-48.

Widén-Nilsson, E., S. Halldin and C. Xu (2007), Global water-balance modelling with WASMOD-M: Parameter estimation and regionalization, *Journal of Hydrology 340*, 105-118.

*Table 1. Mean, minimum and maximum hydraulic conductivity and drainable porosity for the 7 texture classes distinguished in Figure 5.*

| | | **Conductivity ($m\ d^{-1}$)** | | | **Porosity (-)** | |
|---|---|---|---|---|---|---|
| | **Min** | **Mean** | **Max** | **Min** | **Mean** | **Max** |
| **Non- or semi-consolidated sediments** | 0.3 | 2.5 | 20 | 0.1 | 0.23 | 0.35 |
| **Mixed consolidated sediments** | 0.1 | 1.0 | 10 | 0.01 | 0.05 | 0.1 |
| **Siliclastic sedimentary rocks** | 0.01 | 0.1 | 1.0 | 0.01 | 0.03 | 0.06 |
| **Acid volcanic rocks** | 0.1 | 1.0 | 10 | 0.01 | 0.05 | 0.1 |
| **Basic volcanic rocks** | 0.01 | 0.1 | 1.0 | 0.03 | 0.04 | 0.12 |
| **Complex metamorphic and igneous rocks** | 0.001 | 0.01 | 0.1 | 0.001 | 0.01 | 0.05 |
| **Complex lithologies** | 0.01 | 0.1 | 1.0 | 0.01 | 0.04 | 0.09 |

*Table 2. Objective functions used in the reservoir model[a]. After Haddeland et al.[2006] and Adam et al. [2007].*

| Purpose | Objective function |
|---|---|
| Water supply[b] | $\min \sum_{m=1}^{12} \left(Q_{d\,m} - Q_{r\,m}\right),\ Q_d > Q_r$ |
| Flood control | $\min \sum_{m=1}^{12} \left(Q_{r\,m} - Q_{flood}\right)^2,\ Q_r > Q_{flood}$ |
| Hydropower | $\min \sum_{m=1}^{12} \frac{1}{P_m \cdot Q_{r\,m} \cdot \rho \cdot g \cdot h_m}$ |
| Navigation | $\min \sum_{m=1}^{12} \left(Q_{r\,m} - \overline{Q}\right)$ |

[a]: $Q_d$ forecasted water demand, $Q_r$ reservoir release, $Q_{flood}$ bankfull discharge, $\overline{Q}$ mean annual discharge, $P$ variations in the price of hydropower, $\rho$ density of water, $\eta$ efficiency of power generating system, $h_m$ hydrostatic pressure head (water height in reservoir with respect to downstream level) and $g$ acceleration due to gravity.

[b]: Water supply covers net irrigation water demand (blue water) and gross industrial and domestic water demand.

*Fout! Verwijzingsbron niet gevonden. compared to previous estimates; compiled by Wada [2008].*

| Continent | Europe | Asia | Africa | North America | South America | Oceania | Global[1] | Time Period |
|---|---|---|---|---|---|---|---|---|
| Data based estimates | | | | | | | | |
| Baumgartner and Reichel [1975] | 2564 | 12,467 | 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 | 13,091 | 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 | 11,234 | 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* | *12,204* | *3933* | *5730* | *10,536* | *1466* | *36,566* | - |
| This study (PCR-GLOBWB) | 2487 | 11,397 | 4515 | 5040 | 10,558 | 2371 | 36,368 | 1961-1990 |
| This study (PCR-GLOBWB) | 2506 | 11,364 | 4439 | 5028 | 10,505 | 2317 | 36,159 | 1958-2001 |

[1] Excluding Antarctica

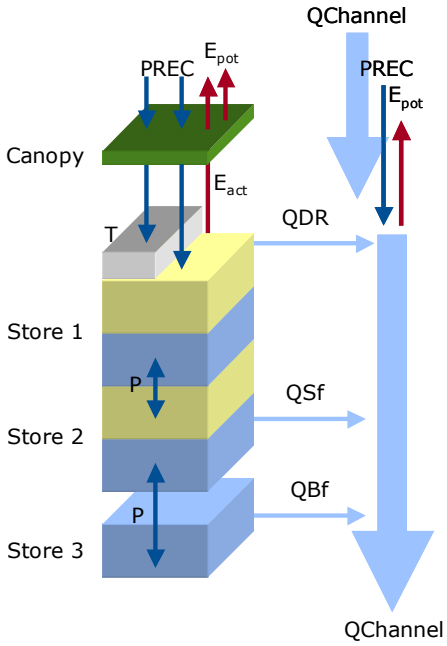



Figure 1. Model concept of PCR-GLOBWB: on the left, the soil compartment, divided in the two upper soil stores and the third groundwater store and their corresponding drainage components of direct runoff (QDR), interflow (QSf) and base flow (QBf). On the right the resulting discharge along the channel (QChannel) with lateral in- and outflow and local gains and losses are depicted.

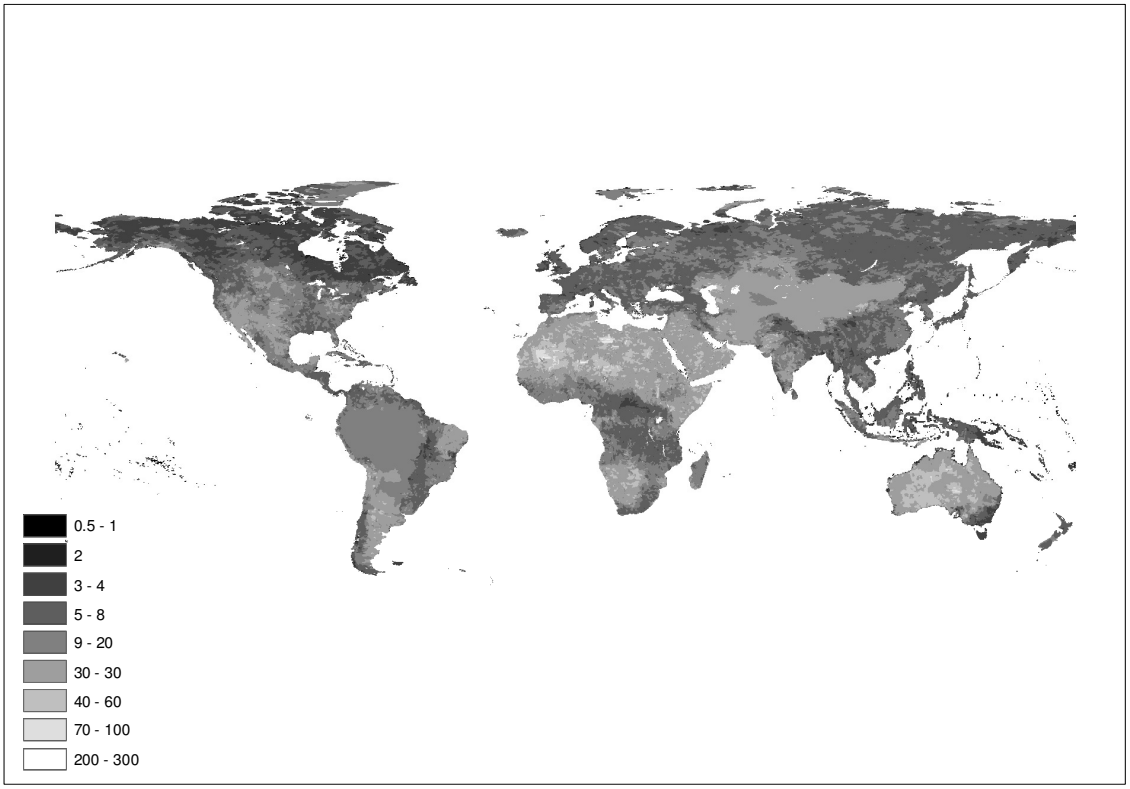



Figure 2. Global map of stream path length or drainage distance*B* (km).

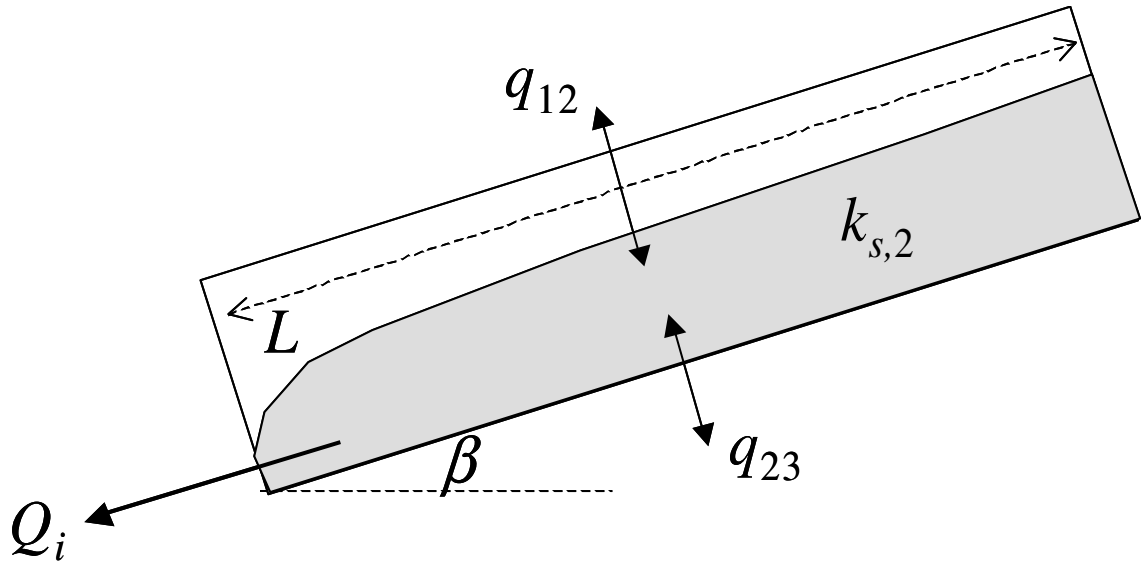



Figure 3. Schematic of perched groundwater body in layer 2 causing interflow; $q_{12}$, $q_{23}$ fluxes from/to first soil layer and groundwater layer ($ms^{-1}$), $\beta$ slope, $L$ drainage distance (m), $k_{s,2}$ saturated hydraulic conductivity of layer 2 ($ms^{-1}$), $Q_i$ interflow per m slope width, ($m^2s^{-1}$).

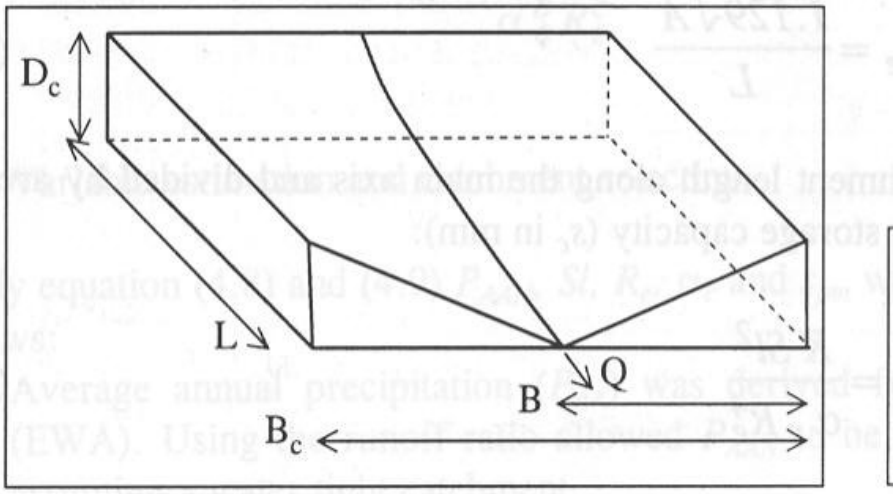



Figure 4. Schematic of a catchment showing catchment width $B_c$, aquifer depth, $D_c$, and slope length/drainage distance $B$.

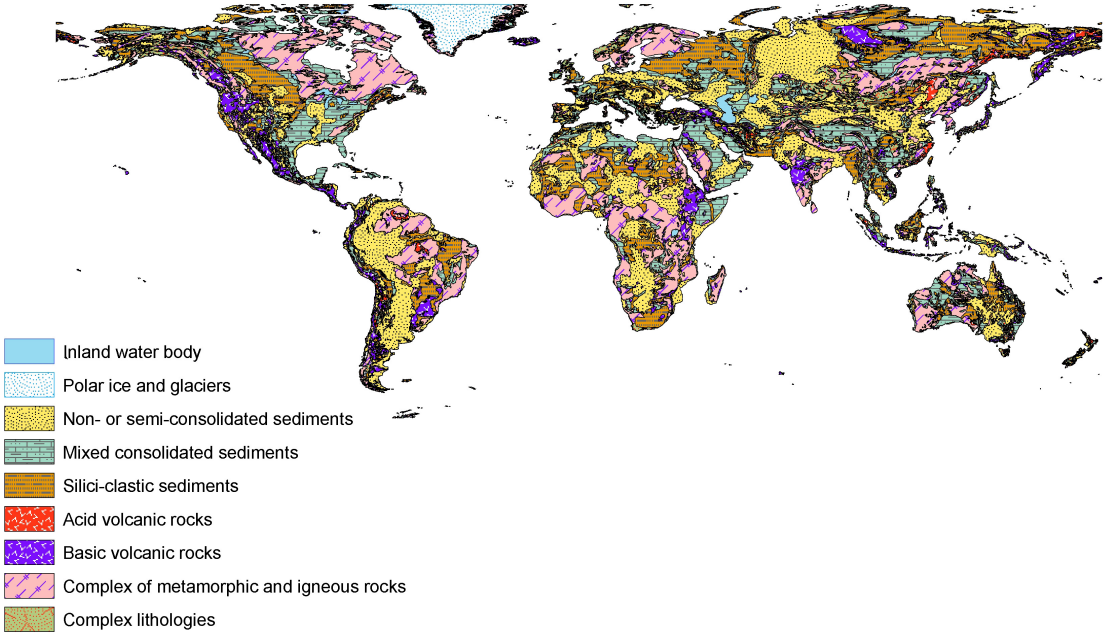



Figure 5: lithological map of the World, simplified from the lithological map of Dürr et al (2005).

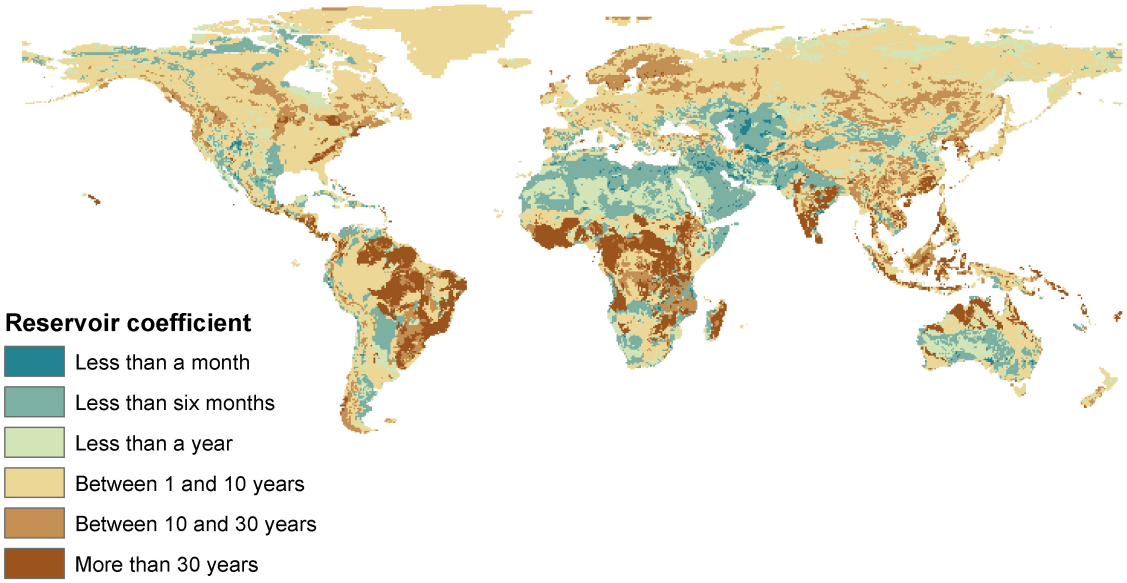



Figure 6. Global map of estimated groundwater residence times.

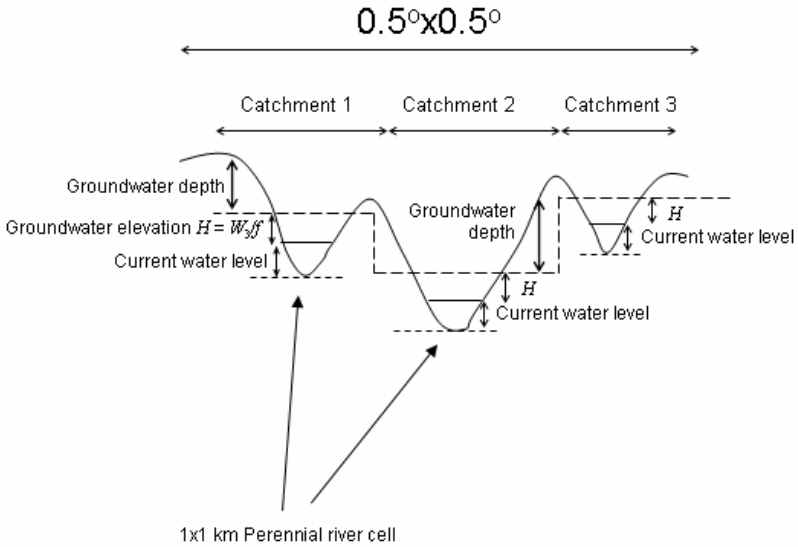



Figure 7. Schematic of determining groundwater depth distribution within a $0.5 \times 0.5^0$ cell based on actual surface water levels, actual groundwater storage and a 1×1 km digital elevation model.

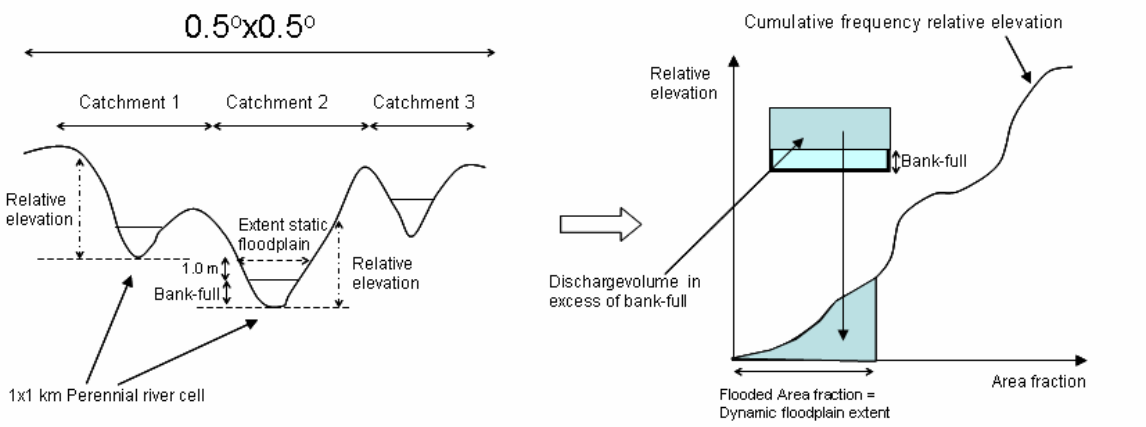



Figure 8. Schematic of determining the static floodplain area and the dynamic floodplain extent (fraction flooded area) within a $0.5^0 \times 0.5^0$ cell based on bank-full discharge surface water level, actual surface water level and a 1×1 km digital elevation model. Left figure determines either the static floodplain extent for the fixed area option or the cumulative distribution of relative elevations for determination of the dynamic floodplain extent as denoted in the right figure.

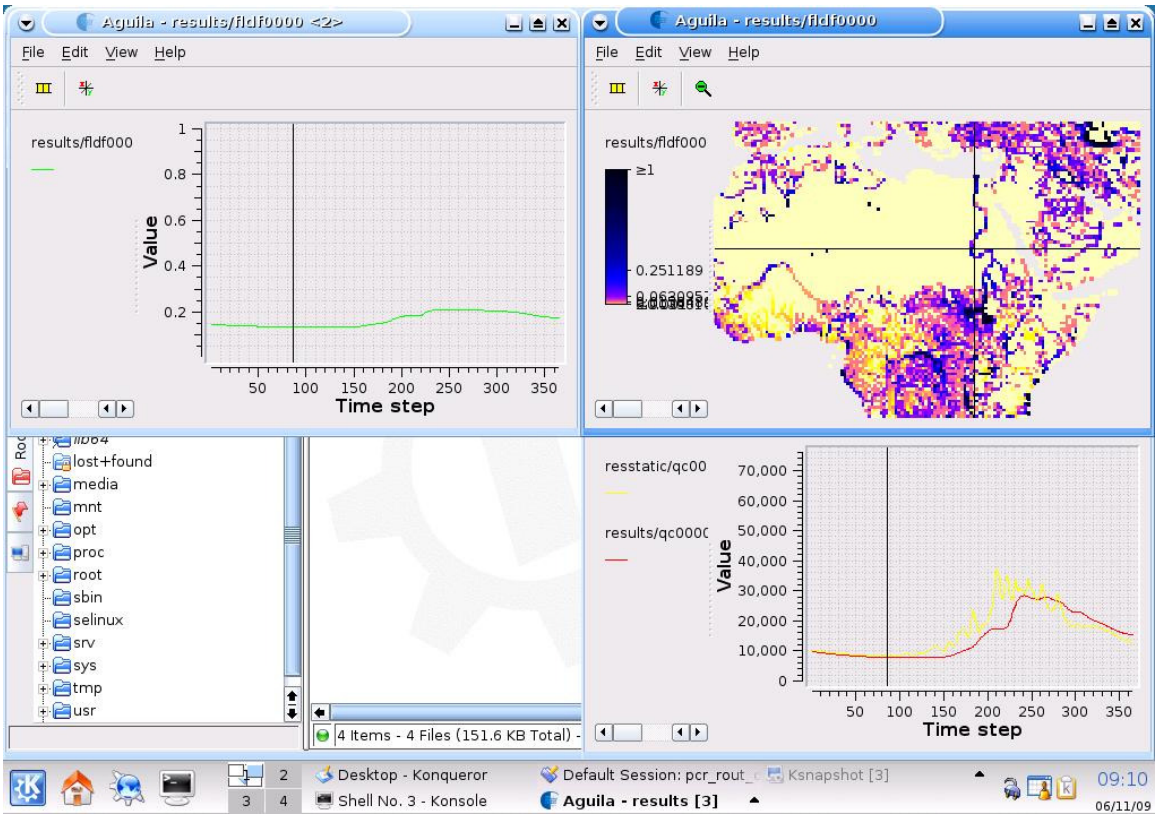



Figure 9 Impression of fraction inundated areas by flooding in Africa calculated with PCR-GLOBWB. Hydrographs are from the Nile, showing a run with the static and dynamic floodplains.

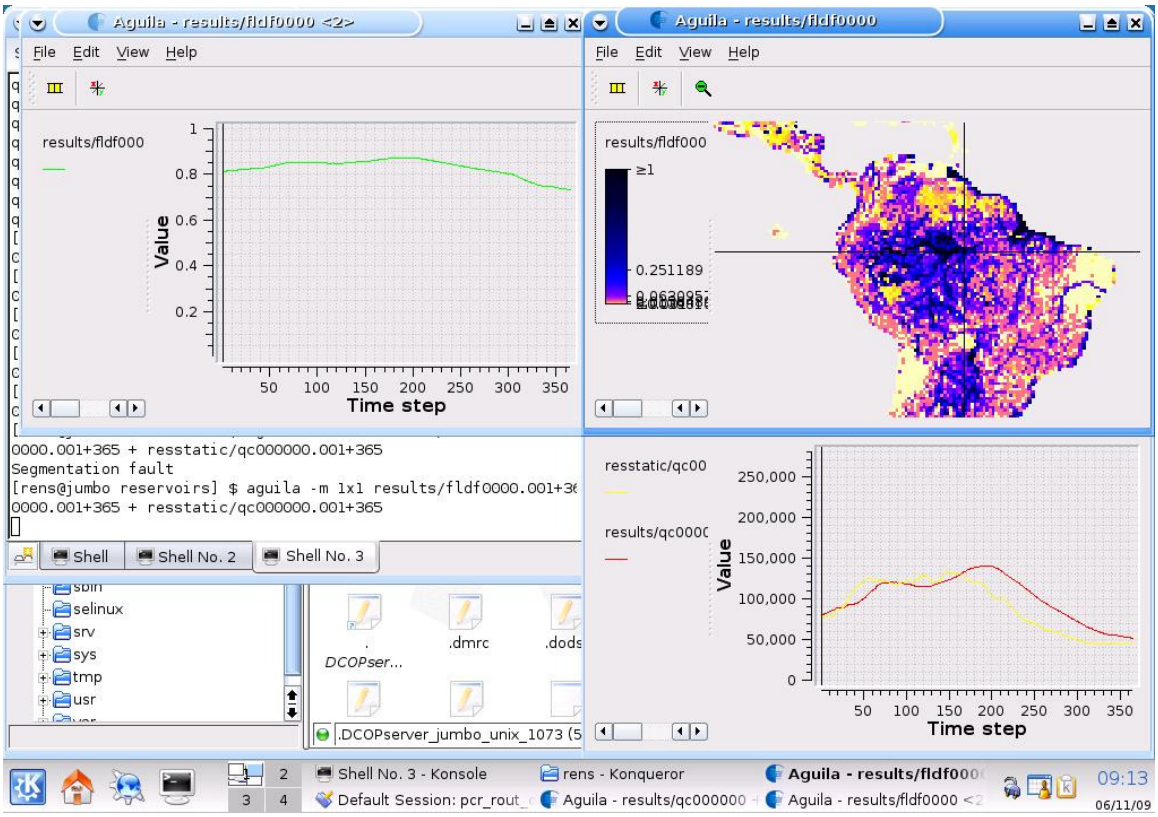



Figure 10 Impression of fraction inundated areas by flooding in Africa calculated with PCR-GLOBWB. Hydrographs are from the Amazone, showing a run with the static and dynamic floodplains.

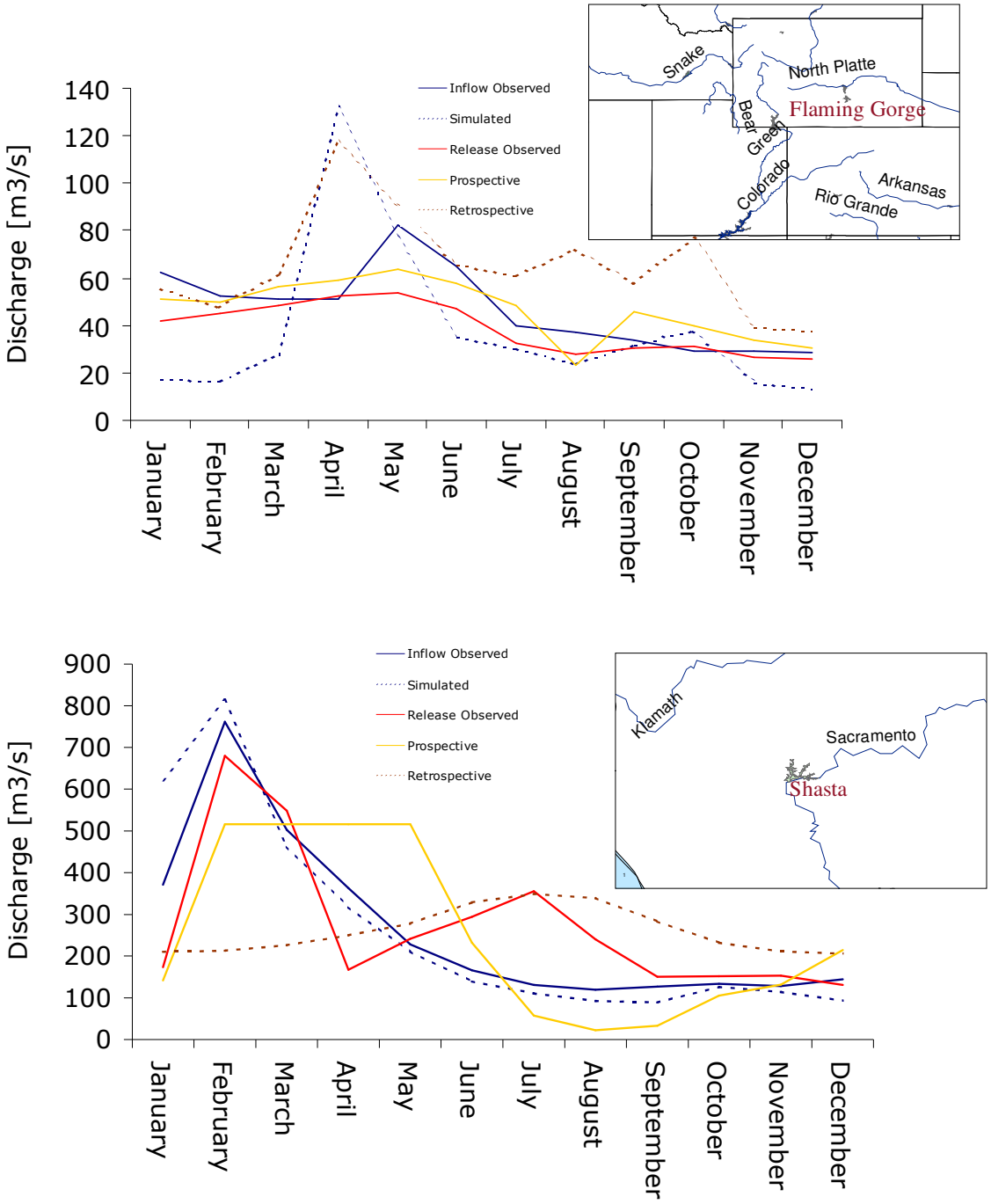



Figure 11. Examples of application of the reservoir scheme for Flaming Gorge on Red River and the Shasta Dam on the Sacramento river for the year 2000. The influx of flaming Gorge is not well predicted by the scheme, but the outflow is well reproduced. For Shasta, the influx is well predicted, but the outflux is to much dominated by the energy production, while in reality downstream irrigation demand has a larger weight.

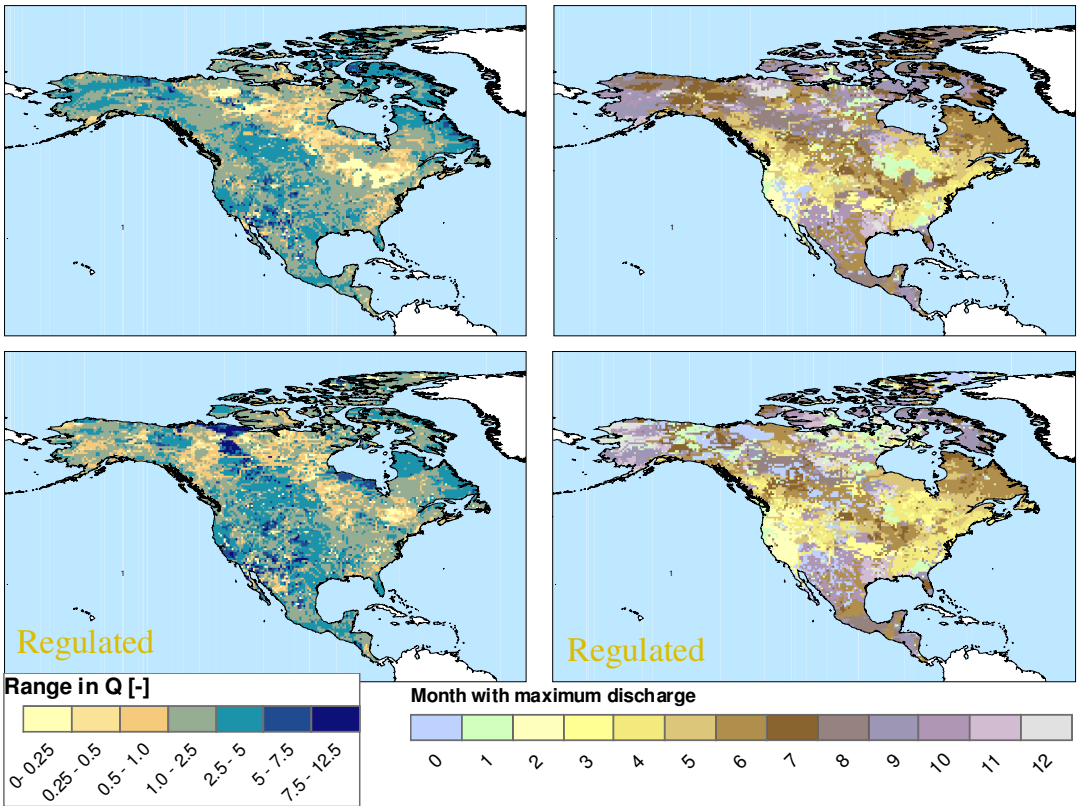



Figure 12. The effect of reservoir operation on river discharge variability of North American rivers as measured by relative discharge range (range divided by runoff mean; left panels) and change in month (modus) with maximum discharge (right panels). Clearly, the effect of reservoir operation on river discharge variability is significant.

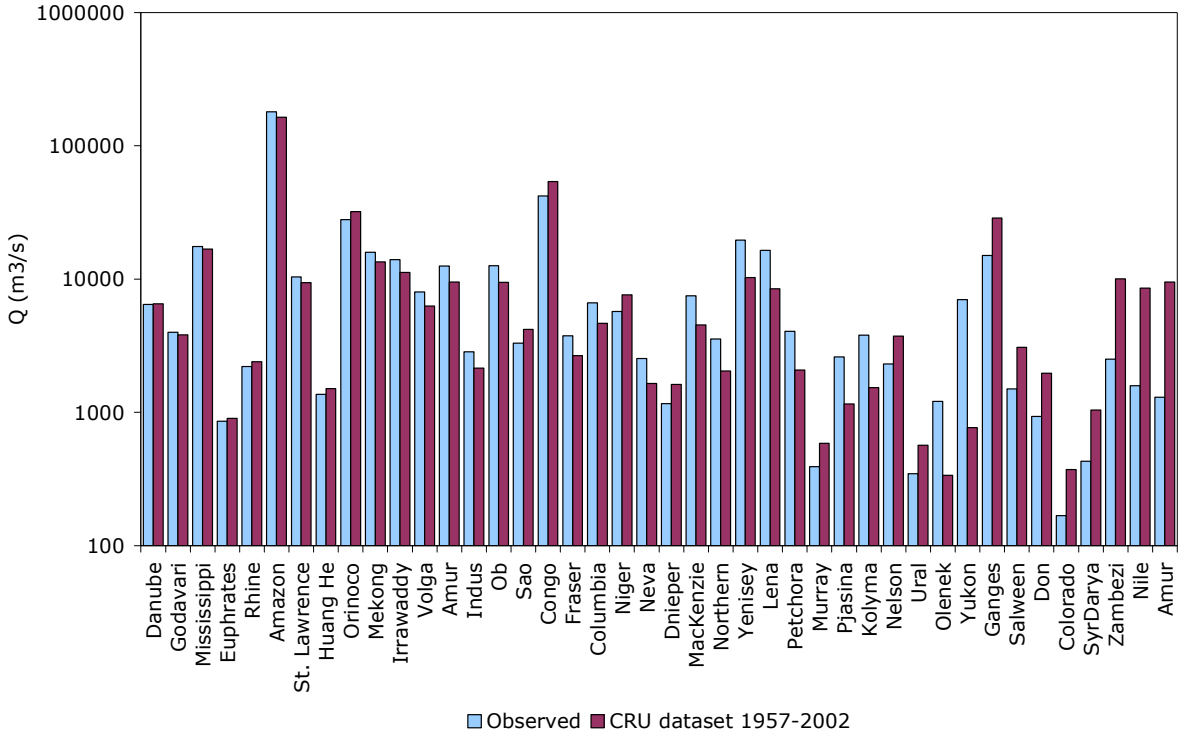



Figure 13 Bar chart comparing simulated mean yearly discharge ($m^3 s^{-1}$) with observed discharge for 99 stations from GRDC and RIVDIS.

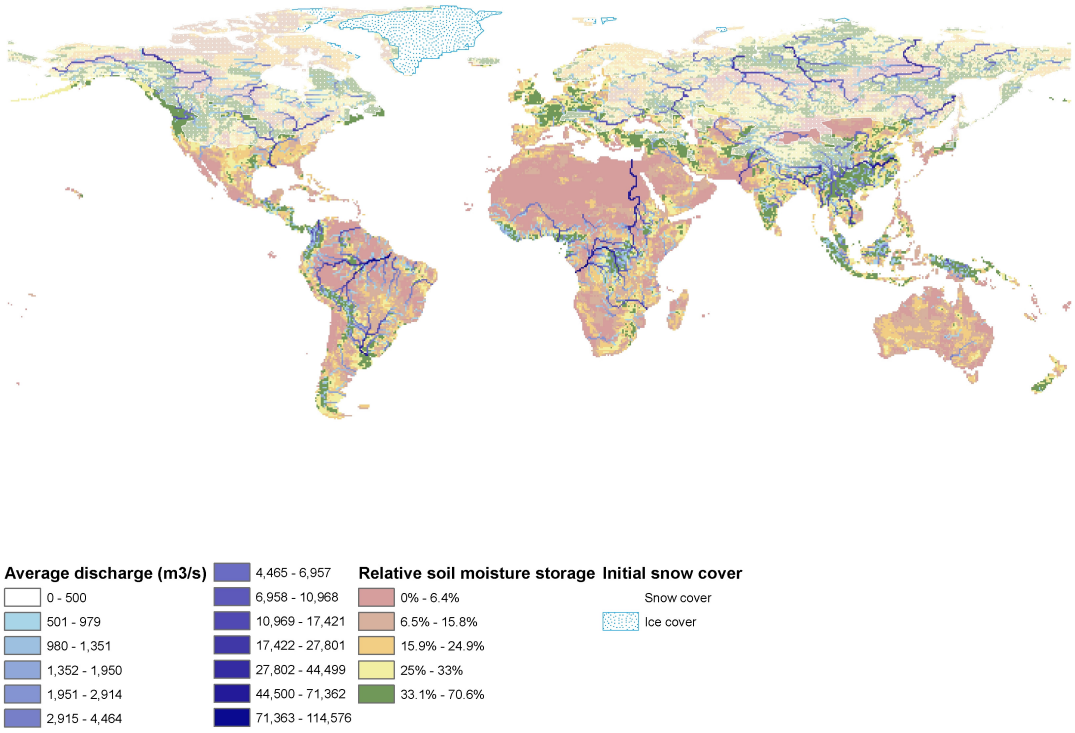
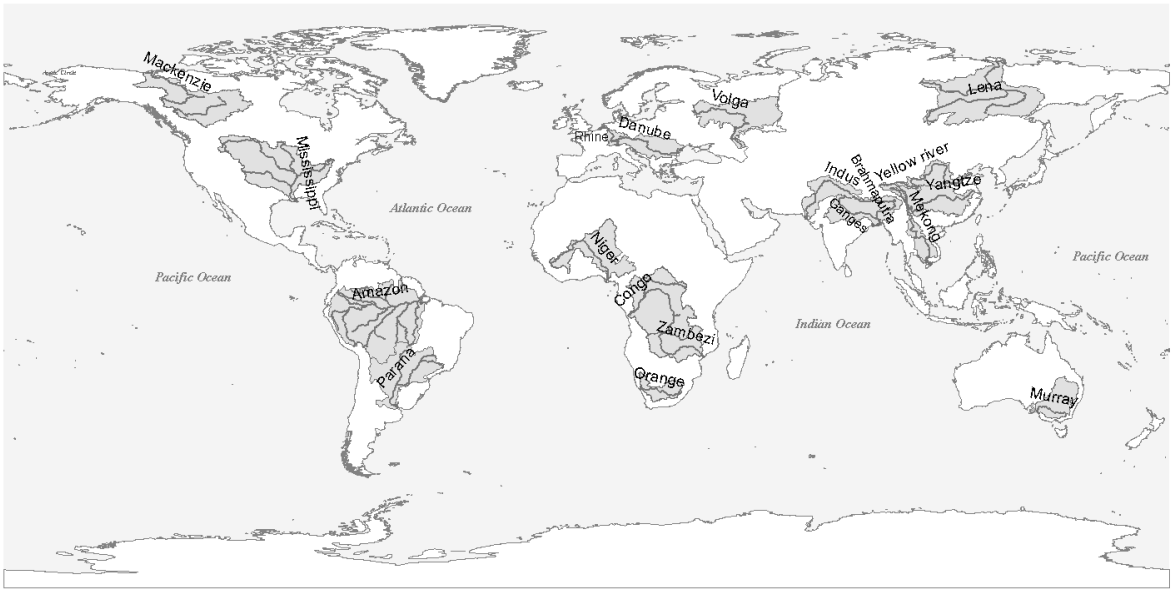



Figure 14. Composite of global hydrology including yearly average river runoff and average soil moisture and average snow cover on December 1.




| 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 |
| Brahmaputra | 930,000 | 48,160 | Bahadurabad | | Niger | 2,117,700 | 6,000 | Dire |
| Congo river | 3,680,000 | 41,800 | Kinshasa | | Nile | 3,400,000 | 2,830 | El ekhshase |
| Danube | 817,000 | 6,400 | Ceatal Izmail | | Orange river | 973,000 | 365 | Aliwal north |
| Ganges | 907,000 | 12,015 | Farakka | | Parana | 2,582,672 | 18,000 | Corientes |
| Indus | 1,165,000 | 6,600 | Attock | | Rhine | 65,638 | 2,200 | Rees |
| Lena | 2,500,000 | 17,000 | Aldan | | Yangtze | 1,800,000 | 31,900 | Datong |
| MacKenzie | 1,805,000 | 10,700 | Norman Wells | | Yellow river | 752,000 | 2,571 | Huayuankou |
| Mississippi | 2,981,076 | 12,743 | Vicksburg | | Volga | 1,380,000 | 8,060 | Volgograd |
| Mekong | 795,000 | 16,000 | Mukdahan | | Zambezi | 1,390,000 | 3,400 | Katima Mulilo |

Figure 15. Catchments for which regime curves are shown in Figure 14. Discharge observed at the catchment's outlet. Catchments are chosen from different climate zones around the world.

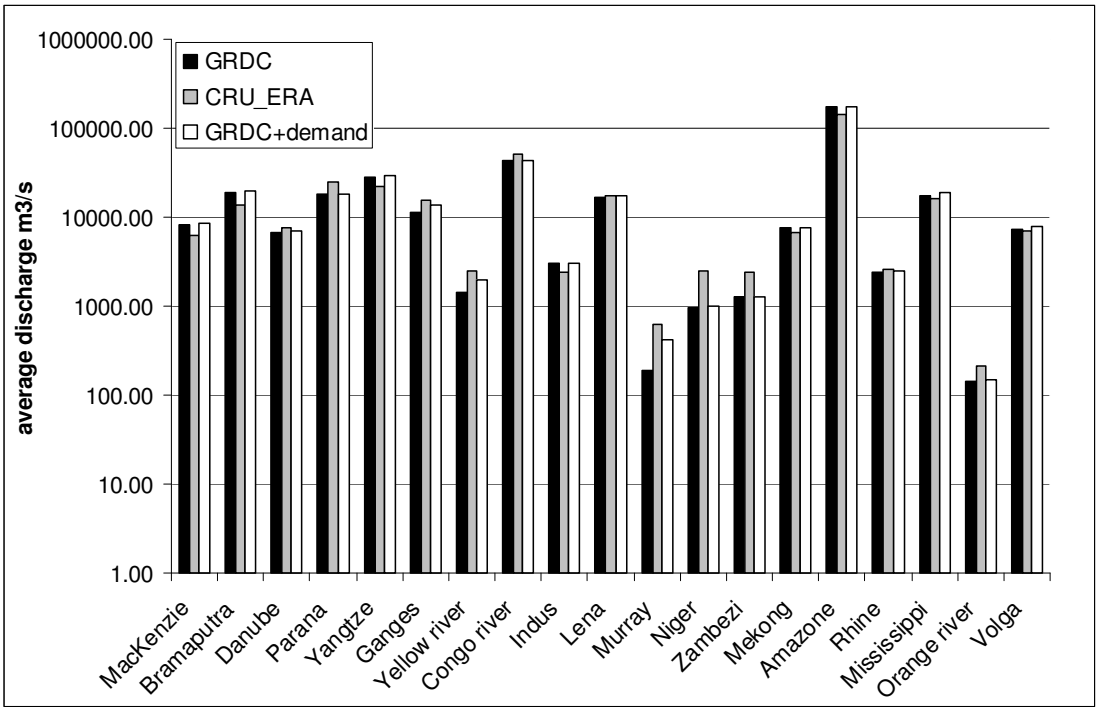



Figure 16. Bar chart comparing simulated mean yearly discharge ($m^3\ s^{-1}$) with observed discharge for the19 catchments shown in Figure 12. As a proxy for natural discharge we have also added a bar chart for observed runoff + water demand from Wada et al. [2008].

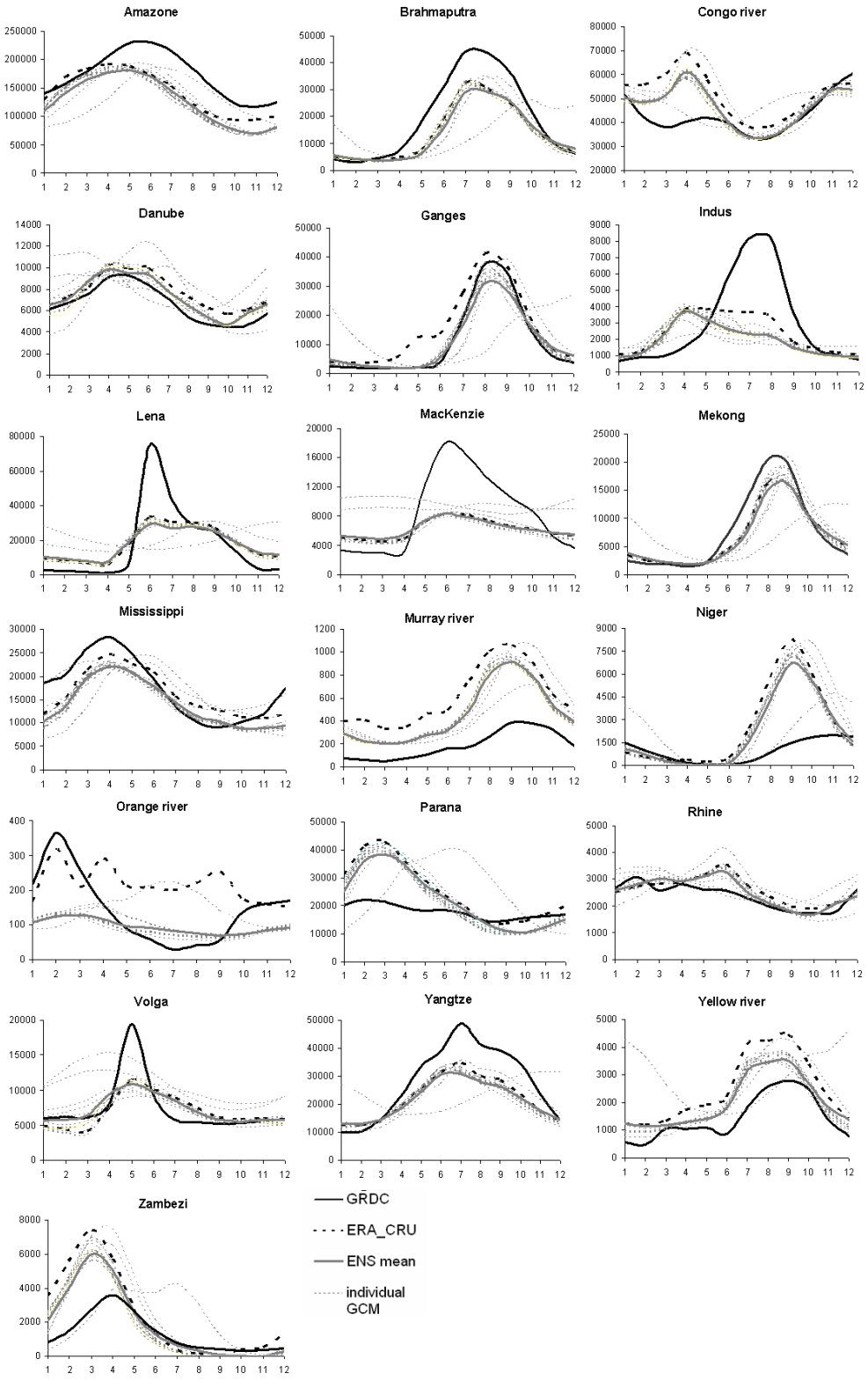



Figure 17. Simulated and observed regime curves of the 20 catchments shown in Figure 12. PCRGLOB-WB was forced with CRU data downscaled to daily data from ERA40 and by 12 GCMS bias-corrected on CRU

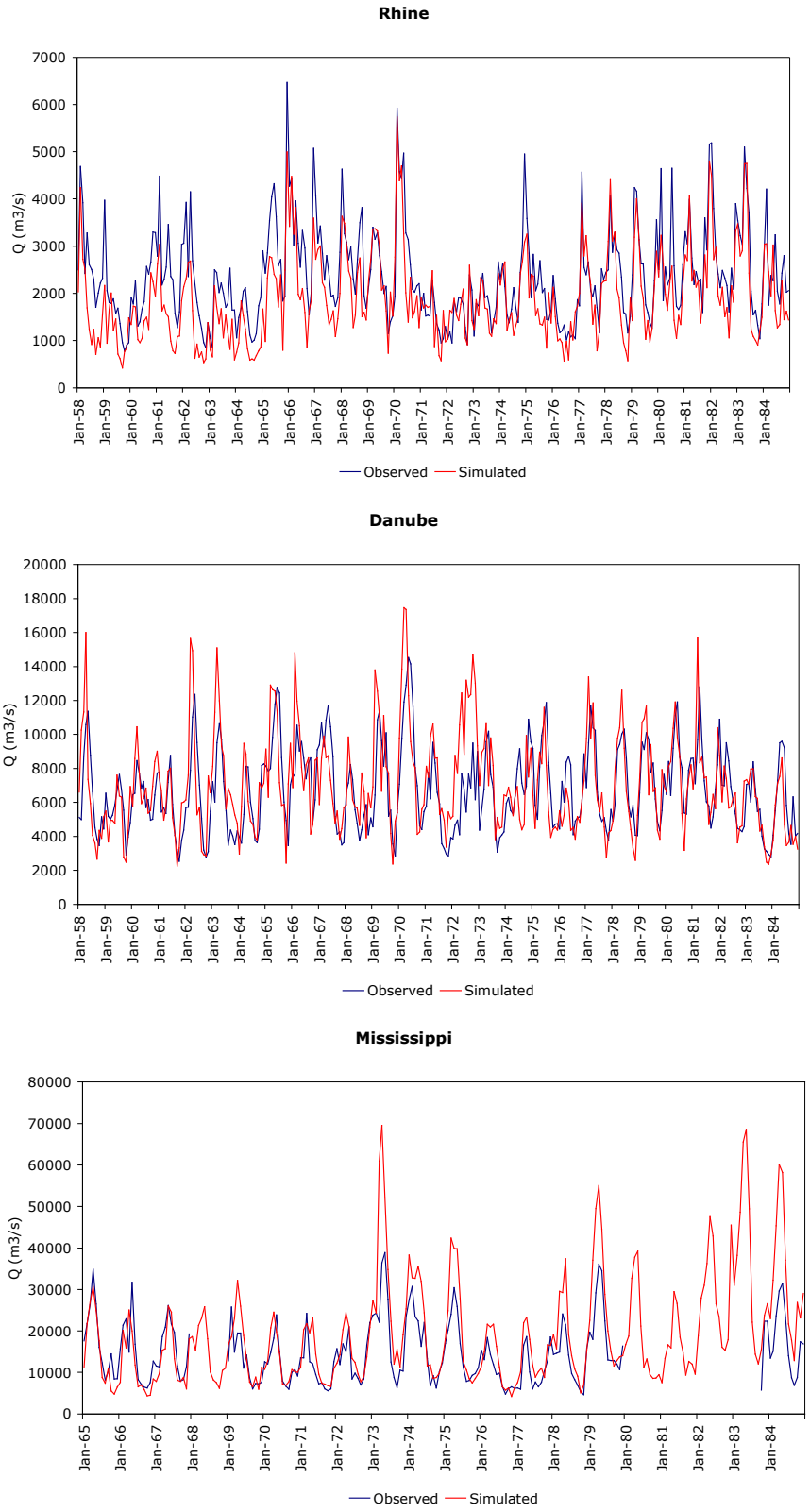



Figure 18 Reproduction of hydrographs of monthly discharge of the Rhine, Danube and Mississippi rivers

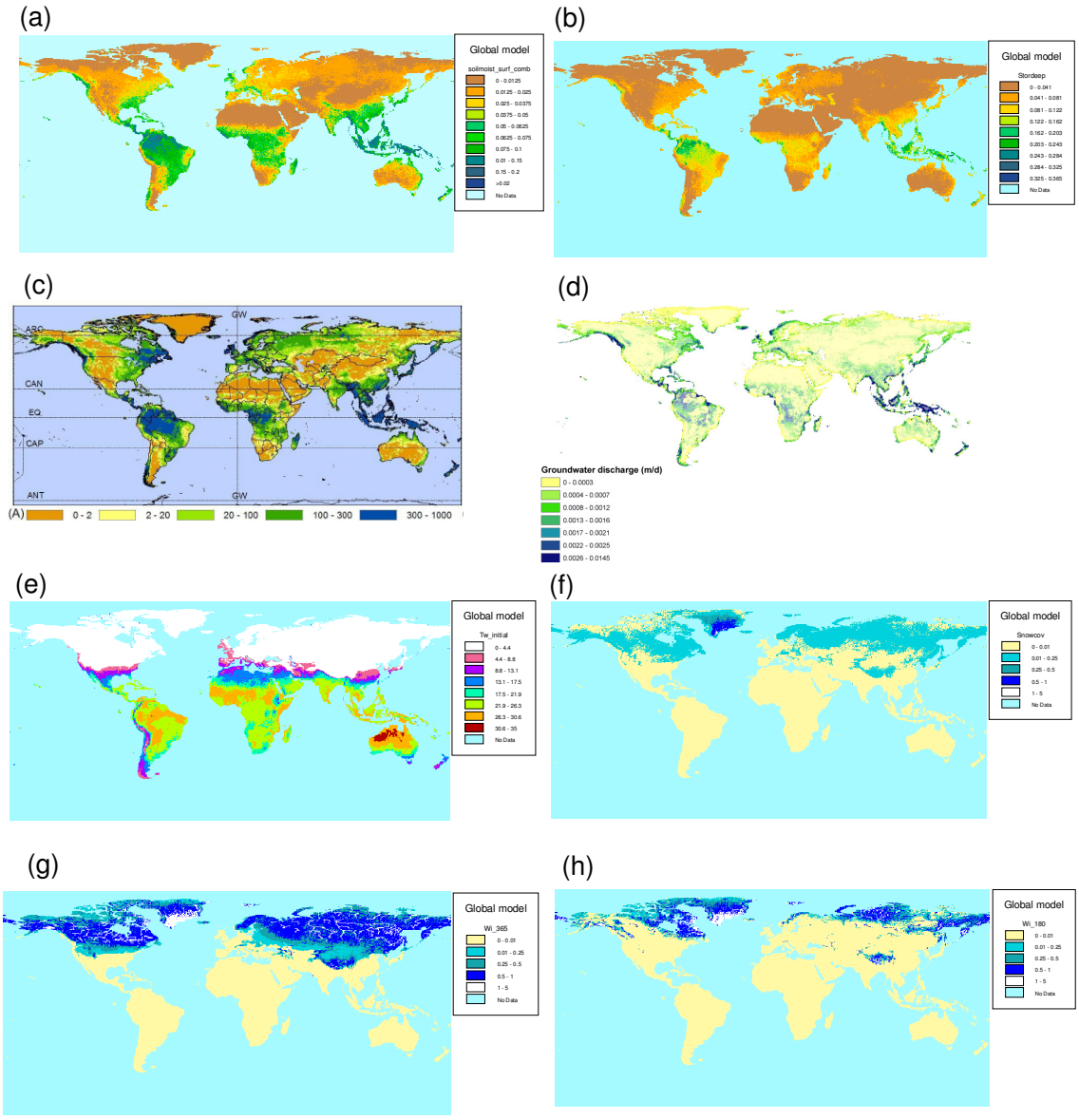



Figure 21 Some collected additional (unverified) output from PCR-GLOBWB;
(a) average soil moisture content on January 1; (b) average soil groundwater storage on January 1; (c) average groundwater recharge (mm/year); (d) average yearly groundwater discharge to surface water (i.e. baseflow; mm/day); (e) average surface water temperature on January 1 ($^{o}$C); (f) average snow cover on January 1 (m water equivalent); (g) average ice thickness (m) on January 1; (g) average ice thickness on July 1.