CrossMark

Modelling Freshwater Resources at the Global Scale: Challenges and Prospects

Petra Döll[1] · Hervé Douville[2] · Andreas Güntner[3] ·
Hannes Müller Schmied[1] · Yoshihide Wada[4,5,6]

Abstract Quantification of spatially and temporally resolved water flows and water storage variations for all land areas of the globe is required to assess water resources, water scarcity and flood hazards, and to understand the Earth system. This quantification is done with the help of global hydrological models (GHMs). What are the challenges and prospects in the development and application of GHMs? Seven important challenges are presented. (1) Data scarcity makes quantification of human water use difficult even though significant progress has been achieved in the last decade. (2) Uncertainty of meteorological input data strongly affects model outputs. (3) The reaction of vegetation to changing climate and CO_2 concentrations is uncertain and not taken into account in most GHMs that serve to estimate climate change impacts. (4) Reasons for discrepant responses of GHMs to changing climate have yet to be identified. (5) More accurate estimates of monthly time

✉ Petra Döll
  p.doell@em.uni-frankfurt.de

  Hervé Douville
  herve.douville@meteo.fr

  Andreas Güntner
  guentner@gfz-potsdam.de

  Yoshihide Wada
  y.wada@uu.nl

[1] Institute of Physical Geography, Goethe University Frankfurt, 60629 Frankfurt, Germany

[2] Centre National de Recherches Météorologiques, Météo-France, 42 av. Coriolis, 31057 Toulouse, France

[3] Helmholtz Centre Potsdam, German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany

[4] Center for Climate Systems Research, Columbia University, 2880 Broadway, New York, NY 10025, USA

[5] NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, USA

[6] Department of Physical Geography, Faculty of Geosciences, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

series of water availability and use are needed to provide good indicators of water scarcity. (6) Integration of gradient-based groundwater modelling into GHMs is necessary for a better simulation of groundwater–surface water interactions and capillary rise. (7) Detection and attribution of human interference with freshwater systems by using GHMs are constrained by data of insufficient quality but also GHM uncertainty itself. Regarding prospects for progress, we propose to decrease the uncertainty of GHM output by making better use of in situ and remotely sensed observations of output variables such as river discharge or total water storage variations by multi-criteria validation, calibration or data assimilation. Finally, we present an initiative that works towards the vision of hyperresolution global hydrological modelling where GHM outputs would be provided at a 1-km resolution with reasonable accuracy.

1 Introduction

Water flows as well as water storage at and below the land surface of the Earth affect water availability for humans and ecosystems, result in hazards such as floods and affect atmospheric processes, sea level and global biogeochemical cycles. They have been increasingly altered by human actions including emissions of greenhouse gases, land use, water abstractions and the construction of dams and dykes (e.g., Vörösmarty and Sahagian 2000; Sterling et al. 2013). Ecosystems suffer from these alterations and, in many regions, human development is constrained by water scarcity. Freshwater systems including their natural and human components need to be characterized regarding both water quantity and quality to support a sustainable land and water management and a better understanding of the Earth system. Global-scale quantification of water flows and storage in freshwater systems under current and future conditions is of particular interest in a globalized world and can support the wise development of global-scale water management and governance (Vörösmarty et al. 2015).

Quantification is best achieved by combining in situ and remote sensing data with physical modelling. Global hydrological modelling serves to estimate water flows on the land areas of the globe such as evapotranspiration, river discharge or groundwater recharge as well as water storage (or only water storage variations) in different compartments, e.g., in soil or in groundwater and surface water bodies. It uses data on precipitation and other climate variables over land areas as input and computes water flows from the land areas to the oceans (or into internal sinks on the continents), thus covering the terrestrial part of the global water cycle. While global hydrological modelling has been refined and extended with respect to modelled processes (in particular regarding human impacts on natural water flows and storages) and computed indicators in the last decade, modelling uncertainties have not become less albeit better known. These uncertainties are generally categorized into uncertainties due to model inputs (e.g., climate variables or soil properties), parameter values and model structure, but uncertainty of observations used for model validation or calibration also has to be considered (Sood and Smakhtin 2015). Different models compute contradictory estimates of, for example, mean annual evapotranspiration and low, mean and high flows in river basins (Gudmundsson et al. 2012) or groundwater depletion (Döll

et al. 2014a). They result in strongly varying projected impacts of climate change on river discharge (Schewe et al. 2014) or irrigation water requirements (Wada et al. 2013). Even global mean annual evaporation estimates as derived from global hydrological modelling (or satellite observations) differ by almost a factor of 2 (Jiménez et al. 2011), which is an important obstacle for the detection and attribution of changes in evapotranspiration due to global warming (Douville et al. 2012). Reasons for the discrepant model output have not been sufficiently analysed.

With this paper, the authors wish to share their perspectives on important challenges of and prospects for modelling continental water flows and storages at the global scale. In the next section, we briefly present existing modelling approaches. In Sect. 3, we discuss seven challenges and illustrate them with results of two global hydrological models (GHMs), WaterGAP (Döll et al. 2003; Müller Schmied et al. 2014) and PCR-GLOBWB (Wada et al. 2014). In Sect. 4, we present three advancements that may help to better characterize freshwater flows and storages at the global scale in the future. Finally, we draw our conclusions.

2 Approaches for Modelling Global Hydrology

To understand and quantify natural and human-induced water flows and storage changes across large scales, a number of models that simulate the continental part of the hydrological cycle on a regional to global scale have been developed in recent decades. Models developed to simulate global hydrology can be roughly classified into GHMs, land surface models (LSMs) and dynamic global vegetation models (DGVMs). Most DGVMs, however, do not include lateral water flows or surface water bodies, and can therefore only be used to assess runoff but not discharge. GHMs focus on simulation of water resources; they have a comprehensive representation of continental hydrological processes and often take into account human water use as well as man-made reservoirs. LSMs serve as a module of global climate models (GCMs) and therefore model both water and energy balances at the land surface. Due to this, they often represent the soil with a higher vertical resolution than GHMs and represent evapotranspiration and snow melt in a less conceptual manner than GHMs. LSMs often lack a groundwater reservoir, lateral routing or consideration of surface water bodies, and in most cases, they do not model the impact of human water use or man-made reservoirs. Finally, some LSMs are able to also model vegetation dynamics, or DGVMs have been extended to simulate global hydrology including not only vertical but lateral water flows as well as human water use and man-made reservoirs. In addition, simulation of irrigation water use is not only done by some GHMs, LSMs and DGVMs but also by global crop models (e.g., Elliott et al. 2014).

In the following, we do not distinguish between GHMs, LSMs and DVGMs but summarily refer to all of them as GHMs (like Schewe et al. 2014 or Hagemann et al. 2013 did) because existing models cannot be strictly classified into the three categories and because we focus on their ability to simulate terrestrial water flows and storages. GHMs typically simulate the dynamics of soil moisture storage due to precipitation and evapotranspiration, the generation of runoff and the discharge through the river network. The majority of these models are based on the water balance concept and track the transfer of water through a number of storage compartments with time steps ranging from a month to less than 1 day. Conceptual models are chosen as they are deemed to be more robust than empirical models and more parsimonious in their data requirements than fully physically based models,

while they maintain the ability to translate the effects of global change on water flows and storages in a consistent manner. Over time process descriptions have become more physically based. Few models simulate human water use that is essential to quantify river discharge, water availability and water stress, and even fewer models represent ground-water including groundwater recharge and abstractions, which is crucial to assess groundwater resources. Sood and Smakhtin (2015) presented an overview over 12 GHMs, and Bierkens et al. (2015) provided a table that describes the main features of GHMs and regional-scale hydrological models.

3 Challenges

Aiming at an improved representation of freshwater systems at the global scale, global hydrological modelling faces diverse challenges. We select some of the most important challenges that have been identified by the scientific community, i.e. constraints that lead to uncertain model output and thus limit the usefulness of global hydrological modelling for understanding freshwater systems.

3.1 Modelling Human Water Use

Human water use leads to anthropogenic water flows in the form of water abstractions from and return flows to surface water or groundwater bodies (Döll et al. 2012; Wada et al. 2011). Quantification of these flows is important for two reasons. On the one hand, water abstractions, consumptive water use (the part of the withdrawn water that evapotranspires during use) or net water abstractions (water abstractions minus return flows) are used in combination with estimates of water availability to compute indicators of water stress (or water scarcity). On the other hand, these anthropogenic water flows alter natural groundwater and surface water flows and storages (Döll et al. 2014a, b; Wada et al. 2012). It was estimated that around the year 2000, mean annual river discharge had been decreased due to water abstractions and man-made reservoirs by more than 10 % on one-sixth of the global land area (excluding Greenland and Antarctica), as compared to natural discharge (Döll et al. 2009). The strongest alterations, with, e.g., both decreases and increases of mean annual water storage (Döll et al. 2012), are found in semi-arid and arid areas of the globe, where irrigation is the dominant water use and alterations of river flow regimes by water abstractions are more important than alterations due to man-made reservoirs (Döll et al. 2009).

In global hydrological modelling, water abstractions and return flows are mostly estimated at a spatial resolution of 0.5° by 0.5° (55 km by 55 km at the equator) or 5′ by 5′ (9 km by 9 km). While water use for domestic and industrial purposes is assumed to vary negligibly throughout the year, monthly estimates of irrigation water use are required due to the often high seasonal variation of irrigation requirements. Modelling of water use for households and manufacturing strongly relies on statistical water abstraction data provided by countries, but data generally exist for a few years only. To derive annual time series by country, abstractions are modelled taking into account structural and technological change (Flörke et al. 2013). In addition, downscaling to the grid cell level is required and is mainly done based on urban and rural population in grid cells (Flörke et al. 2013; Vassolo and Döll 2005). Cooling water requirements for thermoelectric power plants are computed for each

power plant as a function of power plant type and cooling system, as well as values of national thermal electricity production (Flörke et al. 2013).

Unlike domestic and industrial water abstractions, irrigation water abstractions are very rarely measured, and statistical data for countries are generally based on either modelling or rough assumptions on per hectare irrigation water use. Consumptive irrigation water use in a grid cell is modelled as a function of irrigated area, crops and climate (e.g., Döll and Siebert 2002; Rost et al. 2008). Estimation of areas equipped for irrigation and even more estimation of areas actually irrigated are prone to large uncertainties as quality of statistical information is very heterogeneous (comp. information of map quality of the Global Map of Irrigation Areas GMIA v5.0, http://www.fao.org/nr/water/aquastat/irrigationmap/index40.stm). Assumed cropping patterns describing crop-specific growing periods throughout the year strongly impact the seasonality of estimated irrigation water demands (Zhou et al. 2015) and to a lesser degree annual values. Global consumptive irrigation water use as computed by six state-of-the-art global water use models varies between 1000 and 1500 km^3/year (Zhou et al. 2015, Siebert and Döll 2010), but this range may underestimate total uncertainty as all models used the same map of irrigated areas (GMIA). Estimated consumptive use strongly depends on the algorithm to compute potential (or reference crop) evapotranspiration, with global values ranging from 1180 to 1450 km^3/year just due to using three different equally plausible algorithms (FAO Penman–Monteith and two versions of Priestley–Taylor; Siebert and Döll 2010). To estimate water abstractions for irrigation, irrigation water use efficiencies (water consumption-to-abstraction ratios) need to be defined, e.g., for individual world regions (Döll and Siebert 2002), or simulated depending on the crop (in particular paddy rice; Wisser et al. 2008) or type of irrigation system (surface, sprinkler and drip; Jägermeyr et al. 2015). By modelling water flows in the three irrigation systems in a processed-based manner for different crops, Jägermeyr et al. (2015) computed grid cell-specific irrigation water use efficiencies as well as beneficial consumptive use due to transpiration. They estimated that global abstractions for irrigation are twice as high as the global consumptive irrigation water use of 1260 km^3/year, while 52 % of consumptive use is beneficial.

Irrigation water use is normally computed under the assumption that the crops receive enough irrigation water to allow actual evapotranspiration to become equal to the potential evapotranspiration value. However, it is likely that farmers irrigate less in case of water scarcity. Backed by independent information on worldwide groundwater depletion, application of a GHM indicated that in areas with groundwater depletion, farmers irrigated with only 70 % of the optimal water volume (Döll et al. 2014a).

An important uncertainty in human water use modelling relates to the distinction between the groundwater and surface water abstractions, and also to which degree return flows caused by irrigation recharge groundwater. Currently, the distinction of the source of water abstraction appears to more uncertain in the domestic and manufacturing sector than in the irrigation sector (Döll et al. 2012). Table 1 summarizes major uncertainties in modelling sectoral human water use at the global scale.

3.2 Uncertain Climate Input

A number of studies have shown the very strong dependence of computed continental water flows on applied climate input (Biemans et al. 2009; Döll and Fiedler 2008; Guo et al. 2006; Müller Schmied et al. 2014; Nasonova et al. 2011). Not only precipitation, but also radiation data are strong drivers of water flows and storages around the globe, while temperature data have the strongest impact in case of snow and ice. However, both historic

Table 1 Major challenges in modelling sectoral human water use at the global scale (livestock water use is neglected)

Water use sector	Challenges
Irrigation	Areas equipped for irrigation and areas actually irrigated; irrigated crop calendars; degree of deficit irrigation; climate input; equation for computing potential evapotranspiration; fraction of return flows to groundwater and surface water
Domestic	Downscaling to grid cells
Manufacturing	For many countries, only data on sum of water use for manufacturing and thermal power plant cooling available; downscaling to grid cells
Thermal power plant cooling, TP	Temporal development of cooling systems
All sectors	Technological change rates leading to water saving; water consumption-to-abstraction ratio (except TP); differentiation between groundwater and surface water abstractions (except TP)

and future climate information are prone to multiple types of uncertainty which represents a major challenge for providing reliable GHM output.

3.2.1 Uncertainties in Historic Climate Information and Their Impact on Simulating Water Resources

Historic climate information applicable for global-scale studies suffers mainly from insufficient density of high-quality observations but also from observation errors. The effect of incomplete knowledge about the spatio-temporal distribution of climate variables during historic periods on global sums of water flows and storage changes is illustrated in Table 2 where simulation results of the GHM WaterGAP 2.2 (Müller Schmied et al. 2014) are listed. The considered model output encompasses global mean annual river discharge into oceans and inlands sinks (equal to renewable water resources in the case of model runs

Table 2 Mean annual (2000–2009) water flows (precipitation P, discharge into oceans and inland sinks/total water resources Q, actual evapotranspiration AET, consumptive water use WCa) and change in total water storage (dTWS) calculated for different WaterGAP 2.2 model variants, in km^3/year

	1	2	3	4	5	6	7	8
P	111,819	113,215	111,819	113,215	104,768	111,819	111,819	111,819
Q	39,799	42,470	45,743	49,998	41,405	46,908	51,567	44,031
AET	70,948	69,791	64,884	62,154	63,405	64,937	60,282	67,808
WCa	1326	1184	1368	1208	0	0	0	0
dTWS	−246	−220	−167	−135	−34	−18	−22	−12

1: STANDARD, 2: CLIMATE, 3: STANDARD NoCal, 4: CLIMATE NoCal, 5: PRINCETON NoCal NoUse, 6: STANDARD NoCal NoUse, 7: STANDARD NoCal NoUse −20 % net radiation, 8: STANDARD NoCal NoUse +20 % net radiation

NoCal: model version without calibration, NoUse: model version assuming no human water use (for details on model variants, refer to Müller Schmied et al. 2014)

that neglect human water abstractions), actual evapotranspiration, consumptive water use by humans and water storage changes for the time period 2000–2009. Three different climate data sets were used to generate the table, all of which are available at the spatial resolution of WaterGAP (0.5° grids). For the STANDARD model runs, the WFDEI data were used that combine observational data with ERA-Interim reanalysis data for the years 1979–2009, resulting in climate input with a daily resolution (Weedon et al. 2014). In variant CLIMATE, the monthly data set CRU TS 3.2 (Harris et al. 2014) was used, but monthly precipitation totals were replaced by the GPCC v6 precipitation monitoring product (Schneider et al. 2014) because it includes more observation stations. Downscaling to daily values was done within WaterGAP, in case of precipitation based on the number of wet days in each month. Precipitation input used to drive both STANDARD and CLIMATE variants is based on monthly precipitation data from GPCC, but correction of precipitation error and the GPCC product version differ, leading to small differences in global precipitation (Table 2). Further details on climate input for the two model variants can be found in Müller Schmied et al. (2014). The PRINCETON model variant is forced by the "Global Meteorological Forcing Dataset for land surface modelling v2" from Princeton University which is a combination of global observation-based datasets and the NCEP-NCAR reanalysis (Sheffield et al. 2006) and is available at http://hydrology.princeton.edu/data.pgf.php.

WaterGAP 2.2 is calibrated against observed mean annual river discharge at 1319 stations worldwide, by adjusting 1–3 parameters in the basin cells. Table 2 shows that this calibration reduces the impact of uncertain climate variables on global river discharge into oceans and inland sinks Q, actual evapotranspiration AET, actual consumptive water use WCa and change in total water storage dTWS (comp. differences between columns 1 and 2 to differences between columns 3 and 4 in Table 2). While the difference in global precipitation between CLIMATE and STANDARD of 1396 km^3/year is increased to a difference in Q of 4255 km^3/year without calibration, this difference, i.e. sensitivity to climate input, is reduced by calibration to 2671 km^3/year. The difference is still appreciable even in the calibrated model version due to areas without observed river discharge. Global AET reacts much less sensitive to the different climate inputs. This is also seen when comparing the results of two uncalibrated model runs (no human water abstractions assumed) that are driven by either the Princeton or WFDEI climate data (columns 5 and 6). The Princeton climate data, with global precipitation estimated to be 7051 km^3/year less than in case of the WFDEI data, result in global total water resources that are 5503 km^3/year less than with the WFDEI data.

To perform a first sensitivity analysis regarding uncertain radiation data, daily net radiation in the uncalibrated STANDARD variant of WaterGAP 2.2 (without human water abstractions, column 6 in Table 2) was decreased or increased by 20 % in each grid cell (columns 7 and 8 in Table 2). At least at the scale of 0.5° grid cells, an uncertainty of 20 % for net radiation is not exaggerated. In case of −20 % net radiation, global actual evapotranspiration decreases by 7.2 %, while with +20 % net radiation, it increases by only 4.4 %. The asymmetry is due to water availability limiting actual evapotranspiration; radiation-limited areas become water limited when net radiation is increased. With a 20 % lower global net radiation on the continents, renewable water resources (i.e. runoff) would increase by 9.9 %, while with a 20 % higher global net radiation, it would decrease by 6.1 %. Figure 1 shows the spatial distribution of the changes in renewable water resources that are computed by WaterGAP 2.2 to result from a 20 % decrease or increase in net radiation (average for 2000–2009), in addition to a global map of the baseline net radiation computed by WaterGAP 2.2. While renewable water resources decrease in most grid cells
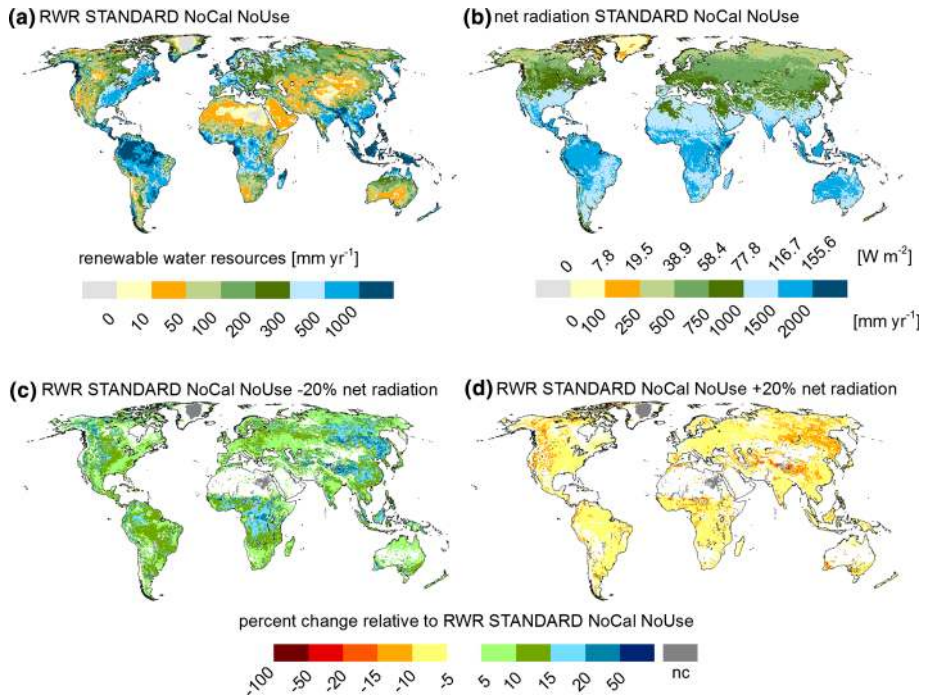
Fig. 1 Sensitivity of renewable water resources RWR during 2000–2009 to a decrease or increase in daily net radiation by 20 % as simulated by WaterGAP 2.2 (Müller Schmied et al. 2014). RWR, i.e. net cell runoff, as computed with uncalibrated model variant STANDARD (**a**) and the related net radiation (**b**), and RWR changes relative to STANDARD in case of decreased (**c**) and increased (**d**) net radiation. nc: Per cent change cannot be meaningfully computed in case of zero or negative values of renewable water resources for STANDARD NoCal NoUse and where the value changes the sign. Negative RWR values occur in cells with lakes or wetlands if evapotranspiration of water flowing in from upstream cells exceeds precipitation

by less than 10 % in case of +20 % net radiation, they increase by more than 10 % in many regions of the globe in case of −20 % net radiation. Changes in per cent of water resources are highest in radiation-limited areas, i.e. in many grid cells with surface water bodies (e.g., in Tibet) but also in tropical and cold regions (even without surface water bodies). In water-limited areas, simulated water resources vary less (Fig. 1).

3.2.2 Uncertainties in Global and Regional Climate Projections and Their Impact on Simulating Future Water Resources

Land and water management as well as assessment of climate change impacts in support of climate change mitigation require quantitative climate projections from seasonal to multi-decadal timescales that can be used as input to GHMs. These projections are subject to uncertainties that are different from the uncertainties of historic climate. The first uncertainty is related to future scenarios of anthropogenic perturbations including emissions of greenhouse gases and sulphate aerosols, land use change and water management itself all of which affect future climate at least on multi-decadal timescales. The second uncertainty, which is related to model deficiencies as illustrated by the different GCM responses to identical radiative forcings, is important at all timescales. Model uncertainty is generally

higher for precipitation than for temperature and arises partly from the limited spatial resolution of GCMs and the need to parameterize unresolved processes within atmosphere and ocean and at the land surface. Higher-resolution regional climate models still show significant biases when driven by atmospheric reanalyses and still require bias correction based on observations for use in impact studies. While there is hope to narrow model uncertainty by further increasing model resolution down to a few kilometres, at least for resolving explicitly atmospheric convective processes, the need of ensemble simulations at both seasonal and multi-decadal timescales will remain a major obstacle for at least a decade and raises the issue of both climate model formulation (structural model uncertainty) and model calibration (parameter uncertainty).

The need of large ensembles is related to the third source of uncertainty: the internal variability of the climate system, that is, the natural fluctuations that arise in the absence of any anthropogenic forcing (and any natural radiative forcing such as solar activity or volcanic eruptions). Appreciation of these fluctuations is an important matter for decision makers because they have the potential to reverse—for a decade or so—the longer-term trends that are associated with anthropogenic climate change (e.g., Douville et al. 2015). While they have been recognized as a fundamental limit to predictability from the early beginning of seasonal forecasting, their relevance in climate scenarios has been emphasized more recently and is probably still underestimated in most impact studies which often use a single realization of a given GCM for driving the impact model.

The relative importance of the three sources of uncertainty varies with prediction lead time and with spatial and temporal averaging scale, but is also variable dependent. Focusing on seasonal precipitation at the regional scale from the CMIP3 archive, Hawkins and Sutton (2011) showed that internal variability contributes 50–90 % of the total uncertainty for all regions for precipitation projections of the next decade and is the most important uncertainty for many regions for lead times up to 30 years. Model uncertainty is generally dominant thereafter. Scenario uncertainty was found to be small or over land areas. This is different for other climate inputs such as surface air temperature or surface radiation, and hydrological impacts of climate change differ appreciably between low and high emissions scenarios in the second half of the twenty-first century (Jiménez Cisneros et al. 2014).

The fourth source of uncertainty is related to the statistical bias correction of GCM outputs (e.g., Hagemann et al. 2011). Such techniques are sometimes included in statistical downscaling tools but, again, must be also implemented on top of dynamical downscaling tools. Statistical bias correction is commonly applied in climate impact modelling to correct GCM output for systematic deviations of the simulated historical data from observations. It has been found that evapotranspiration and river discharge as computed by GHMs that are driven by GCM output differ significantly for historic time periods if climate model output is not bias corrected; bias correction of precipitation and temperature is more important than bias correction of radiation, humidity and wind speed (Haddeland et al. 2012). Bias correction methods are generally based on transfer functions generated to map the distribution of the simulated historical GCM output to that of the observations. Those are subsequently applied to correct the future projections, thus making the assumption that GCM biases are constant over time. While such an assumption is not necessarily true, especially for precipitation biases (Chen et al. 2015), there are other challenging assumptions in most bias correction techniques.

3.3 Quantification of the Role of Active Vegetation Under Changing Climate and CO_2 Concentrations

When climate and atmospheric CO_2 concentrations change due to anthropogenic climate change, vegetation changes too. Therefore, evapotranspiration and runoff are not only affected directly by changes in climatic variables but in addition by the vegetation reaction to changes in climatic variables and CO_2 concentration. Rising CO_2 concentration causes two counteracting effects (Gerten et al. 2014). On the one hand, the physiological effect reduces opening of leaf stomata, as less water is required to assimilate carbon in case of higher CO_2 concentration; this decreases transpiration for the same climatic conditions. On the other hand, CO_2 fertilization may cause increased plant growth leading to increased transpiration per unit area (structural effect). Current quantification of these effects is highly uncertain, due to diverse observational evidence but also algorithms in DGVMs (Gerten et al. 2014). This is true not only for natural vegetation but also for crops (Elliott et al. 2014). In addition, the changing climate itself affects the vegetation, e.g., altering biomass production or vegetation cover or even leading of a biome shift.

It is uncertain how vegetation responses to increasing CO_2 and changing climate will affect water flows, which adds uncertainty to the response of water flows to climate change. Effects may be large, in particular where the type of vegetation changes. A modelling study with a DGVM indicated that until the end of the twenty-first century the active vegetation may cause a relative increase in runoff in response to increased atmospheric CO_2 concentration (physiological effect dominant), except in areas where grassland changes to deep-rooted vegetation in a warmer climate (Murray et al. 2012). In case of 4 °C global warming, the high temperatures lead to decreased vegetation cover, such that runoff as a ratio of precipitation is projected to increase worldwide (Murray et al. 2012). Large uncertainties also surround the response of vegetation to persistent droughts in both present-day and future climates, especially over the Amazon rainforest (Joetzjer et al. 2014).

Most GHMs, like most basin-scale hydrological models, do not model vegetation responses to changes in climate and atmospheric CO_2 concentrations. Thus, if, for example, the physiological effect were dominant, those GHM would underestimate future runoff and thus renewable water resources. Neglecting the reaction of crops to increased CO_2 may lead to an overestimation of future irrigation water demand (Wada et al. 2013; Elliott et al. 2014; Gerten et al. 2014). In the multi-model study of Wada et al. (2013), the only model that considered CO_2 effects on crop photosynthesis and transpiration shows a decreasing trend in future irrigation water demand (about 10 % by the end of this century) and increasing yields, while model runs without CO_2 effect indicated pronounced increases in future irrigation water demand (>20 % by the end of this century).

Unfortunately, models that simulate vegetation responses strongly disagree among each other on the effect of active vegetation on evapotranspiration and runoff. In a multi-model study on projected runoff changes between 1981–2010 and 2070–2099, two DGVMs computed higher runoff, and two computed lower runoff compared to GHMs with passive vegetation (Fig. 2; Davie et al. 2013). Comparing the runs of the four models with elevated CO_2 to runs with constant CO_2, they found that modelling the CO_2 effect on vegetation contributes to an increased spread in runoff projections. The challenge is to (1) improve modelling of climatic and CO_2 effects on vegetation with respect to evapotranspiration and runoff and (2) to include the effect of the active vegetation on evapotranspiration and runoff also in the majority of GHMs that do not model vegetation dynamics.
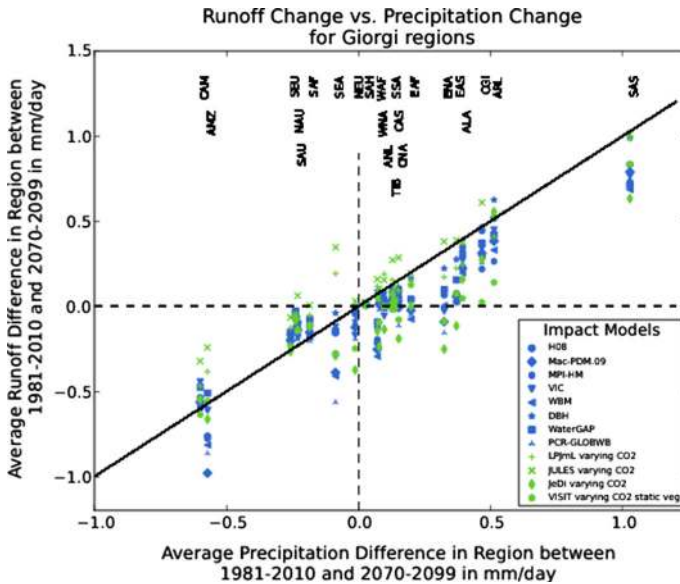
Fig. 2 Uncertain impact of vegetation response on runoff changes under future anthropogenic climate change: scatterplot of runoff change against precipitation change between 1981–2010 and 2070–2099 in mm/day for world regions as computed by global hydrological models not taking into account active vegetation (*blue*) and global vegetation models that do (*green*). All models are forced with HadGEM2-ES RCP8.5 climate. Figure by Davie et al. (2013)

3.4 Understanding of Why GHMs (Including Global Irrigation Models) Respond Differently to Changed Climate Input

Traditionally, GCMs have been considered as a major source of the uncertainty in future hydrological assessments. Therefore, hydrological studies on the impact of climate change applied not only the output of one but multiple GCMs as input to the hydrological model (Jiménez Cisneros et al. 2014). However, recent model intercomparison projects (Water-MIP and ISI-MIP) where various GHMs were driven by either standard historic climate data or the bias-corrected output of multiple GCMs showed that differences among GHMs are also a major source of uncertainty regarding evapotranspiration, runoff and discharge (Dankers et al. 2014; Davie et al. 2013; Gosling et al. 2011; Haddeland et al. 2011, 2014; Hagemann et al. 2011, 2013; Schewe et al. 2014) and irrigation water demand (Wada et al. 2013; Elliott et al. 2014). The uncertainty can be larger than that arising from GCMs, depending on the region and the output variable. Considering eight GHMs but only three GCMs, Hagemann et al. (2013) found that the spread in projected changes of actual evapotranspiration dominantly caused by the different GHMs in most areas of the world, while the spread in projected runoff was dominantly caused either by the GCM or by the greenhouse gas emissions scenario (considering changes until the end of the twenty-first century). However, the small number of applied GCMs and bias correction of GCM output has limited the spread of GCM output. Applying eleven GHMs driven by five bias-corrected GCMs, Schewe et al. (2014) determined that GHMs were responsible for a larger spread in river discharge than the GCMs on most of the global land area. This can be explained by the fact that they compared model results not for a specific time period but for

a specific global warming for which the GCM outputs are more similar. Projections of the impact of climate change on optimal irrigation water abstractions as computed by seven GHMs driven the same GCMs were dominated by GHM uncertainty throughout the twenty-first century (Wada et al. 2014). With these studies, it has become state-of-the-art that an ensemble of model runs that has been generated by driving multiple GHMs with the output of multiple GCMs should be evaluated in water-related climate change impact studies. Analysis of such multi-model ensembles should not be restricted to the ensemble mean but also consider results of individual models that would imply a high risk, i.e. results that may have strong negative impacts due to high vulnerability (Döll et al. 2015).

Why do the responses of GHMs (and hydrological models in general) to climate change projections vary so widely? Possible reasons include (1) different model algorithms for the computation of potential and actual evapotranspiration as well as runoff generation, (2) modelling (or not) of energy balance in addition to water balance, (3) modelling irrigation water requirements based on soil water deficits as derived from soil water balances or as the difference between optimal evapotranspiration and available water, (4) different physiographic input parameters such as soil properties and land use and (5) different simulation of vegetation, including the CO_2 effect on crops and other vegetation. The latter aspect has already been discussed in the previous section.

None of the multi-model studies has been able to analyse in depth the reasons for the discrepant results of the individual GHMs. Hagemann et al. (2013) concluded that large differences in projected changes between the GHMs may be attributed to different model formulations of evapotranspiration but provide no further detail. Different methods for computing potential evapotranspiration (e.g., taking into account only temperature, or also radiation or humidity and wind speed) may explain the disagreement not only in energy-limited regions (Haddeland et al. 2012). Among eleven hydrological models applied for the ISI-MIP project, a few models used temperature-based methods for simulating potential evapotranspiration (Schewe et al. 2014), which has likely contributed to the spread. Haddeland et al. (2011) analysed differences among eleven GHMs driven by the same historic climate input, discussing the impact of energy balance-based snow algorithms as compared to degree day-based snow algorithms (affecting seasonal flows) and the impact of differing parameter values. They found that regarding "the interannual variation in runoff and evapotranspiration, no major differences have been found between the models run at daily or subdaily time steps or between models using different evapotranspiration or runoff schemes" (Haddeland et al. 2011, p. 882).

Model intercomparison studies should go beyond the identification of the spread of the model ensemble but also try to understand reasons for the spread and identify routes towards model improvement. However, the challenge of understanding why a large number of complex GHMs with a high spatial and temporal resolution react differently to a spatially and temporally heterogeneous change in input variables may appear over-whelming. A step towards progress could be to devise feasible intercomparison strategies that aim at understanding the major drivers of the spread. Another step would be to assess the capability of the GHMs to simulate hydrological effects of past climate variability, under the assumption that a model that does a good job in simulating hydrological responses to, e.g., inter-annual climate variability is also able to better simulate impacts of climate change. Hydrologists should study lessons learnt in model intercomparison of climate (e.g., Knutti and Sedlacek 2013) or dynamic vegetation models (Sitch et al. 2008; Warszawski et al. 2013). Regarding GCMs, Knutti and Sedlacek (2013) found that model spread regarding future temperature change has not decreased in the new CMIP5 model ensemble as compared to the CMIP3 ensemble, i.e. after about than 6 years of massive

efforts in model improvements. They argue that GCMs have improved and now represent more processes in greater detail, which implies greater confidence in projections even if model spread, which is often called model uncertainty, has not decreased. A decreased spread might even misleading if caused only by using more similar model input or more similar algorithms that become to be considered "state-of-the-art" within the scientific community even without firm scientific support. Regarding DGVMs, the study of Sitch et al. (2008. p. 2035) showed "the ability of models to satisfy contemporary global carbon cycle constraints, while future projections diverge markedly", as many different parameter combinations allow recreating the historical record but lead to divergent future projections.

3.5 Modelling of Monthly Time Series of River Discharge and Human Water Use to Support More Meaningful Indicators of Water Stress for Both Humans and Ecosystems

Currently, indicators of water stress are mostly defined based on mean annual values of water availability and use (e.g., Kiguchi et al. 2015, Arnell and Lloyd-Hughes 2014). Water stress for humans and freshwater ecosystems as well as environmental flow requirements could be defined more meaningfully at the monthly timescale (Hoekstra and Mekonnen 2011) as use of mean annual values masks differences in seasonality and interannual variability. Consideration of mean annual values only leads to an underestimation of water stress in highly seasonal flow regimes, e.g., in monsoon regions, as well as in regions with high interannual flow variability, e.g., in semi-arid and arid regions.

However, WaterGAP and most likely all other GHMs are not capable of satisfactorily simulating monthly time series or even mean monthly values of river discharge and human water use. Müller Schmied et al. (2014) found that even for most of the 1319 gauging stations used for model calibration (that considers only mean annual river discharge), monthly (their Figs. 6 and 7, and Table 4) or mean monthly river discharge (their Fig. 5) was not well simulated. In case of the best model version (STANDARD), only 28 % of the basins showed a modelling efficiency (Nash–Sutcliffe coefficient) of more than 0.7, while 46 % of the basins had a value of less than 0.5. Simulation of seasonality of irrigation water use is known to be highly uncertain as cropping patterns and calendars are not well known (Portmann et al. 2010). Besides, cropping patterns and calendars would ideally be simulated as a function of climate. Therefore, we believe that it is currently not reasonable to compute, in global-scale studies, water stress indicators based on monthly water availability and use. The challenge is to improve both GHMs and their input data such that reliable monthly time series of river discharge and human water use are computed.

An alternative to indicators based on monthly discharge values is to consider statistical monthly low and high flows, e.g., Q_{90} and Q_{10}, the river discharge that is exceeded in 9 out of 10 months and 1 out of 10 months, respectively. These statistical low and high flows include seasonal and inter-annual variability of monthly river discharge, which is particularly high in the dry regions of the globe, and are ecologically relevant indicators of the flow regime. They can be computed reasonably well by the GHMs WaterGAP and PCR-GLOBWB, at least for the gauging stations whose mean annual discharge was used for calibrating WaterGAP (Fig. 3). Simulation of high flows is better than simulation of low flows. In case of observed Q_{90} values of less than 1 km³/month or 10 mm/month, Q_{90} as simulated by WaterGAP simulations may differ very strongly from the observed value even though WaterGAP was calibrated against mean annual river discharge of these stations (Fig. 3 top). WaterGAP tends to overestimate Q_{90} (with 592 of the simulated 821 values being larger than the observed ones) and to underestimate Q_{10} (with 608 of the
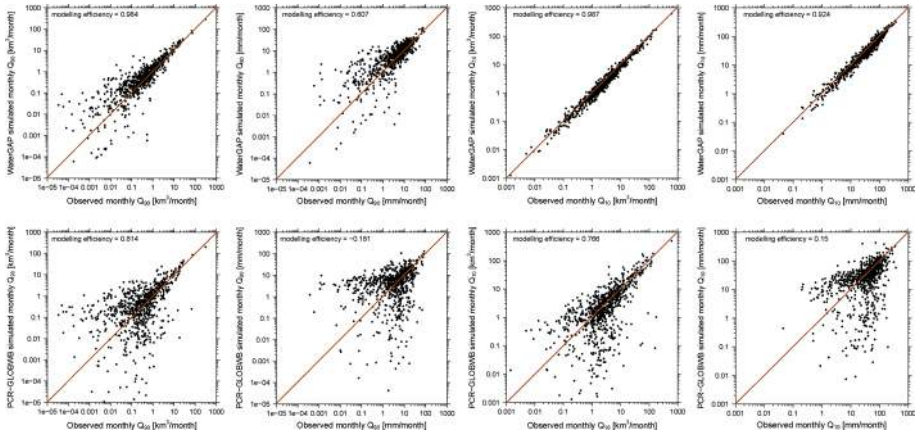
Fig. 3 Validation of monthly low and high flows Q_{90} (*left*) and Q_{10} (*right*) as simulated by the global hydrological models WaterGAP 2.2 (Müller Schmied et al. 2014) (*top*) and PCR-GLOBWB (Wada et al. 2014) (*bottom*) against observations at 821 of the 1319 WaterGAP calibration stations with at least 15 years of data and basin area of at least 20,000 km². Note that different from WaterGAP, PCR-GLOBWB is not calibrated to mean annual observed discharge at the depicted gauging stations

simulated 821 values being smaller than the observed ones). PCR-GLOBWB shows the same behaviour but to a much lesser extent (Q_{90}: 497 of the simulated 821 values are larger than the observed ones; Q_{10}: 449 of the simulated 821 values are smaller than the observed ones). However, model results of PCR-GLOBWB show a lower fit to observed low and high flow values as the model is not calibrated against mean annual discharge at the depicted gauging stations. Water stress indicators based on monthly Q_{90}, where water use is often taken to be mean annual consumptive water use instead of mean annual water abstraction, were applied, e.g., by Alcamo et al. (2007) and Hanasaki et al. (2008).

3.6 Simulation of Groundwater–Surface Water Interaction and Capillary Rise by Gradient-Based Groundwater Modelling

Groundwater is the largest store of freshwater available for human use. It is replenished by precipitation in the form of diffuse groundwater recharge through the soil and sometimes by concentrated recharge from surface water bodies (Taylor et al. 2013). Groundwater flows along gradients of hydraulic head, but this is not represented yet in GHMs. In particular, groundwater flow between grid cells is not simulated. If at all, groundwater is represented in GHMs mainly as a linear storage compartments that discharge groundwater as baseflow into surface water bodies, baseflow being a function of groundwater storage. With a linear groundwater store, the relative temporal changes of groundwater storage can be computed, also as affected by groundwater abstractions (Döll et al. 2014a, b). Groundwater storage changes can be translated to variations of groundwater table elevations, but there is no information of the absolute elevation of the groundwater table or depth to groundwater table. An approach for taking into account groundwater in climate models by Niu et al. (2007) allows to compute depth to groundwater table and capillary rise but does not simulate lateral groundwater flow (nor groundwater recharge from surface water bodies).

Without dynamic modelling of the elevation of the groundwater table, groundwater recharge from surface water bodies, which is particularly important in dry regions, cannot be represented well. Döll et al. (2014a) used a very rough estimate of groundwater recharge from lakes and wetland in semi-arid and arid regions of the globe to avoid underestimation of groundwater recharge and thus overestimation of groundwater depletion. Equally, capillary rise from groundwater to the soil cannot be represented well if the distance of the groundwater table to the land surface is not simulated. Thus, for a simulation of groundwater–surface water interaction and capillary rise, it is necessary to estimate the temporally changing elevation of the groundwater table well; this can only be achieved if lateral groundwater flow driven by gradients of hydraulic head (or groundwater table) is computed (Fan et al. 2007; Jones et al. 2008; Kollet and Maxwell 2008; Krakauer et al. 2014; Maxwell et al. 2007; Maxwell et al. 2015). Lateral groundwater flow is described by a partial differential equation the solution of which is computationally more involved than solving the ordinary differential equation that is used to describe groundwater storage and outflow in typical hydrological models. However, the main reason for not including gradient-based groundwater flow modelling in GHMs may be the extreme lack of information on groundwater that is available for global-scale studies (Taylor et al. 2013; de Graaf et al. 2015).

For North America, Miguez-Macho et al. (2007) linked a land surface scheme with a two-dimensional gradient-based groundwater model and computed, with a daily time step, gradient-based groundwater flow, water table elevation, groundwater–surface water interaction and capillary rise, using a spatial resolution of 12.5 km. One challenge was the determination of the river conductance that affects the degree of groundwater–surface water interaction. Capillary rise was computed using the Richards' equation for a soil column reaching down to the groundwater table by soil layers of variable thickness; the model appears to overestimate capillary rise that is computed to dominate in all flat regions during May–October. Vergnes et al. (2012, 2014) established a gradient-based groundwater model applicable for global-scale modelling and applied it to France with spatial resolutions of 0.5° and 5 arc-min. The transient model simulates two-dimensional groundwater flow dynamics and also accounts for groundwater–river exchange and capillary rise. It is currently being implemented in the GHM of CNMR (Centre National de Recherches Météorologiques, France). Fan et al. (2013) developed a high-resolution (30 arc-s) steady-state global groundwater flow model driven by diffuse groundwater recharge, taking into account land surface elevation. The results indicated that patterns in water table depth explain patterns in wetlands at the global scale and vegetation gradients at regional and local scales. This study was an important step towards simulating groundwater dynamics globally. However, in the chosen approach, neither the important hydraulic connection between rivers, surface water bodies and groundwater nor spatially distributed hydrogeological information was taken into account. A subsequent study by de Graaf et al. (2015) presented an alternative global-scale steady-state groundwater flow model of a shallow aquifer (spatial resolution 6 arc-min), estimating aquifer depths and using a global lithological map (Hartmann and Moosdorf 2012) in combination with estimates of lithology-specific hydraulic conductivity (Gleeson et al. 2014). The results showed the importance of lateral groundwater flows over catchment boundaries as inter-basin flow paths. Both models are likely to overestimate the depth to groundwater. They are not dynamic and are not coupled to a model that dynamically models groundwater recharge and surface water levels. Therefore, the simulation of groundwater–surface water interactions is very limited, and neither capillary rise nor groundwater pumping or groundwater recharge by irrigation return flows is considered.

At a continental scale, Maxwell et al. (2015) showed the possibility of setting up an integrated hydrological model that simulates surface and subsurface flow at a high spatial resolution (1 km). The model solves surface and subsurface flow simultaneously and is constructed entirely of available datasets including topography, soil texture and hydrogeology. However, the steady-state simulation did not take into consideration runoff generation, transient dynamics or human activities such as groundwater pumping that affect the quantity of surface water fluxes and groundwater recharge (Döll et al. 2014a, b). Maxwell et al. (2015) concluded that these limitations can be addressed within the current modelling framework but require additional computational resources. Advanced soil–groundwater–surface water modelling systems such HydroGeoSphere (Brunner and Simmons 2012) that are widely applied at local scales would require not only faster computers and better calibration strategies but also good quality data to be applicable at the global scale.

Given the poor knowledge on the three-dimensional shape and distribution of aquifer bodies as well as restricted computational resources, gradient-based groundwater modelling cannot aim, in the near future, at modelling groundwater flows in three dimensions or at supporting the sustainable management of specific aquifers. The focus is on better representing groundwater–surface water interactions and capillary rise. Here, the major challenge is to achieve a reasonable representation even with relatively large grid sizes such that it is computationally feasible to perform transient groundwater flow simulations that are coupled to soil water and surface water dynamics.

3.7 Detection and Attribution of Observed Changes in Freshwater Systems

Detection is the process of demonstrating that an observed change cannot be explained by internal climate variability only. Attribution of a change to anthropogenic influence requires the additional demonstration that the detected change is consistent with the change simulated in response to a combination of external forcings. While detection is generally done using statistical methods, attribution almost always requires the use of models. Most often in hydrology, detected change, e.g., of river discharge, is attributed to observed changes in climate or CO_2 concentrations (Jiménez Cisneros et al. 2014). Attribution of hydrological changes to human climate-altering activities is seldom attempted because it requires the application of climate models. An exception is the study of Pall et al. (2011) where citizens allowed the researchers to perform a very large ensemble of climate model runs on their computers such that it could be shown that increased greenhouse gas concentrations increased the likelihood of a specific historic flood event in the UK by approximately a factor of 2–3.

Detection of changes of water flows and storages is limited by availability of, e.g., river discharge data or data on groundwater recharge. In the case of river discharge, available data are inhomogeneous as records of many stations end in the 1990s or even earlier, and there are a large number of ungauged basins around the world so that there is no good global coverage of river discharge by gauging stations. In case of the important variable groundwater recharge, there are no measured time series at all. In addition, the strong spatio-temporal variability of hydrological variables makes it hard to detect changes, e.g., changes in flood frequency. Attribution of detected changes of river discharge is challenging because river discharge is affected not only by climatic (and CO_2) changes but also by changes in land use and water abstractions (Jiménez Cisneros et al. 2014).

There are studies that tried to attribute changes in river discharge to changes in climate, CO_2 and land use (Gedney et al. 2006) and, more recently, to the radiative effect of

anthropogenic aerosols (Gedney et al. 2014). Such studies were based on the comparison between annual river discharge derived from offline land surface simulations on the one hand and from observed river discharges on the other hand. The river basins with significant irrigation were ignored so that the role of human water use was considered as negligible. Yet, the conclusions of Gedney et al. (2006) that CO_2 increase played a large role in causing increased global river discharge were challenged. They are highly uncertain not only due to the specific model assumptions on physiological versus structural effects on evapotranspiration but also because the applied precipitation data set is not suitable for evaluating trends as it is based on a temporally varying number of precipitation gauging stations, and because a more recent compilation of observed river discharge resulted in a decrease in global river discharge over time (Gerten et al. 2014). Alkama et al. (2011) were successful in simulating recent river discharge trends without accounting for physiological effects and emphasized the possible relevance of permafrost thawing for capturing the discharge trends of northern high-latitude rivers.

As far as evapotranspiration (ET) is concerned, changes at the continental to global scale are even more difficult to analyse due to lack of direct measurements. Only relatively few monitoring sites operate around the world and the period of record is quite short. Two studies (Jung et al. 2010, Wang et al. 2010) have used such in situ measurements for tuning global empirical ET schemes based on remote sensing and standard meteorological data. They agreed on a global increase in annual mean ET by about 7 mm per year per decade from 1982 to the late 1990s. These results were compared with ET outputs of process-oriented land surface models and were found to be relatively robust (Jung et al. 2010). The 1982–2008 period was, however, too short for a formal detection and attribution. More recently, Douville et al. (2012) used several global ET reconstructions based on two land surface models driven by two precipitation forcings and attributed the reconstructed multi-decadal variations of annual mean ET in three latitudinal belts to anthropogenic climate change. The ET reconstructions neither accounted for direct CO_2 effects on ET nor accounted for changes in land or water use, thereby allowing a fair comparison with GCM outputs and a more robust attribution of the effect of anthropogenic climate change. Yet, this strategy did not allow the authors to assess the possible impacts of other human activities on ET, such as water abstractions.

4 Prospects

The role of global hydrological modelling is to combine large amounts of diverse and mostly spatially and temporally resolved data in order to estimate continental water flows and storages and resulting policy-relevant indicators of the water situation worldwide. To decrease the uncertainty of the computed estimates, it will be fruitful to generate and utilize improved GHM input data such as climate data but also to make better use of observations of GHM output variables such as river discharge or total water storage variations. This can be done by multi-criteria validation, calibration or data assimilation. A higher spatial resolution of GHM models beyond the current 0.5° resolution (55 km × 55 km at the equator) will increase the policy relevance of GHM output, e.g., for supporting integrated water resources management at the scale of river basins.

4.1 Multi-criteria Validation Against River Discharge and Geodetic/Remote Sensing Observations

Observations of river discharge are ideally suited for validating macro-scale hydrological models because the point observation integrates over processes in the whole upstream basin of the gauging station. Besides, river discharge is a flow that can be related quite easily to water availability which is the focus of many assessments. Finally, long time series of observational data exist for many stations around the globe and some of them are compiled by the Global Runoff Data Centre (GRDC). When validating model output against observed river discharge, measurement errors should be taken into account, ideally in a station- and discharge-specific manner. None of the 500 UK gauging stations has a discharge observation uncertainty of less than 10 % (for individual measurements at mean flow conditions) due to uncertain stage–discharge relationships, while 83 % of the stations for which uncertainty could be determined has an uncertainty of less than 40 % (Coxon et al. 2015).

Observational data of other components of the water cycle in addition to river discharge are needed to validate hydrological models due to the well-known equifinality problem (Beven and Freer 2001); more than one parameter combination (or model) can lead to a good fit between discharge observations and simulations, while other water flows or storages would be projected very differently by model variants that result in equally good simulations of the river discharge. Therefore, a multi-criteria validation that considers other observed flows or water storage variations or any other observation that are related to flows and storages can be expected to be highly informative.

As an example, total water storage (TWS) variations as modelled by GHMs can be validated against satellite-based geodetic observations of monthly gravity variations of the GRACE satellites (Tapley et al. 2004), at least if all important storage compartments like the groundwater and large surface water bodies are taken into account in the GHM. Simulated TWS variations can also be validated by continuous GPS observations at more than 200 network stations worldwide because water storage variations cause crustal deformations which lead to displacements of the GPS reference point (Döll et al. 2014b). Comparing simulated TWS variations against both GRACE and GPS, Döll et al. (2014b) identified regions where the WaterGAP GHM underestimates seasonal variability of TWS and found that maximum TWS occurs 1 month too early in WaterGAP for most land areas (based on GRACE only). Validating groundwater depletion as computed by WaterGAP against both in situ well observations and GRACE TWS allowed the conclusion that farmers in groundwater depletion area irrigate with only 70 % of the optimal value (Döll et al. 2014a). In the future, combined validation of GHM model output against river discharge and TWS (e.g., Alkama et al. 2010) should be intensified, and time series of lake or river water tables as measured by radar altimetry should serve as additional observational data sets.

4.2 Multi-criteria Calibration and Data Assimilation

Multi-criteria calibration and data assimilation goes beyond multi-criteria model validation. In model calibration, model parameters are adjusted in a way that simulated water flux or storage (state) variables optimally match historic observations with respect to one or more performance criteria. A primary goal of calibration is to obtain a model (including parameter values) that allows simulations for periods without observation data, such as for

simulations of the global water resources for the time period 1971–2000 or of future climate change impacts. GHMs have rarely been calibrated, with few exceptions such as WaterGAP (Döll et al. 2003; Hunger and Döll 2008) and WASMOD-M (Widén-Nilsson et al. 2007) for which one or more parameters were adjusted by evaluating simulation results against observed river discharge. Without basin-specific calibration, even mean annual simulated river discharge may differ strongly from the observed value (Müller Schmied et al. 2014; Haddeland et al. 2011).

Model calibration is usually hampered by parameter equifinality (see Sect. 4.1), and calibration against more than one observable and performance criterion has long been recognized as an option to allow adjustment of a larger number of model parameters and to constrain the number of plausible model realizations (e.g., Gupta et al. 1999). For continental to global-scale modelling, however, there is not yet much experience with multi-criteria calibration, presumably because adequate satellite-based observation data with sufficient spatial and temporal extent and resolution have become available only recently. While considerable uncertainties of remote sensing data products may still limit their value in a multi-criteria calibration strategy (Livneh and Lettenmaier 2012), large-scale calibration examples demonstrated the benefit of using, in addition to river discharge, satellite-based monitoring data, such as near-surface soil moisture from Envisat (Milzow et al. 2011), MODIS-based evapotranspiration (Livneh and Lettenmaier 2012), altimetry-based water levels (Milzow et al. 2011), MODIS-based snow cover (Parajka and Blöschl 2008) or total water storage (TWS) variations from GRACE (Werth et al. 2009, Xie et al. 2012). GRACE TWS and river discharge were incorporated into a multi-criteria calibration scheme for WaterGAP by Werth and Güntner (2010) by adjusting the most sensitive 6–8 parameters in the 28 largest river basins worldwide. Improved simulations of TWS variations and river discharge were achieved for most basins after calibration, but calibrated mean annual discharge was still poor compared to the observed values in some basins, and a better fit to GRACE TWS did not necessarily lead to a better fit of simulated discharge to observed discharge. In the study of Xie et al. (2012), model parameters appear to be less sensitive to TWS than to river discharge. While large trade-offs in model performance for different objective functions leave the model with considerable uncertainties, they can help to unravel deficiencies of the model structure (e.g., Duethmann et al. 2014).

Direct multi-criteria calibration for the essential terms of the continental water balance (river discharge and TWS changes, and evapotranspiration if monitoring data were available) is particularly appealing if one strives for a closed water balance model of the continental areas. Nevertheless, the particular nature of GRACE TWS data based on the Earth's time-variable gravity field requires specific consideration of the storage compartments considered, data filtering and error terms to make the calibration scheme consistent between model and observations (Güntner 2008). Multi-criteria calibration should comprise more than two observables to further constrain the space of plausible model realizations. Besides the types of satellite-based data on states and fluxes mentioned above, information on water storage in surface water bodies has a high potential as a large-scale calibration constraint, based on currently available multi-sensor combination data (for example, Papa et al. 2013) and future satellite missions such as SWOT. With the development of multi-scale modelling and parameterization concepts (Samaniego et al. 2010), even observation data with a small spatial measurement support but a global coverage such as evapotranspiration from Fluxnet eddy-covariance sites (Jung et al. 2011) or near-surface soil moisture based on GNSS reflectometry (Larson et al. 2008), for instance, may inform parameter adjustment in global models within a multi-criteria calibration approach.

For observing systems that require complex operators to transform the sensor signal into a hydrological variable simulated by a GHM, an inverse strategy can be promising. In this case, the state or flux variable of the hydrological model is forward-transformed to a quantity at the sensor level and, thus, parameter adjustment is done by measuring the performance directly relative to the sensor signal. This avoids the need for using an operator which in turn often is a nonlinear model that is afflicted with uncertainty. An example is to convert simulated water storage variations of a hydrological model into K-Band range rate data which is the key observable of GRACE at the level of the twin satellites, i.e. the inter-satellite distance changes as determined by a K-Band Ranging (KBR) System between the two GRACE satellites (Krogh 2011). Parameters in the hydrological model are then adjusted by minimizing the difference between modelled and observed range rate data.

Data assimilation, i.e. an integration strategy of models and data that primarily adjust state variables of the model, and possibly also parameters, may allow for an optimal quantification of the system status for a period where observations are available, and if the respective errors can be adequately specified. With early developments of data assimilation for large-scale applications being tailored towards appropriate initial conditions in land surface schemes of weather forecasting systems, there are now numerous examples for continental to global-scale assimilation of a variety of data types into LSMs with the general aim of hydrological forecasting and provide land surface hydrological states that are superior to satellite observations or model estimates alone (see, e.g., overviews in Li et al. 2012; Reichle et al. 2014; Lahoz and de Lannoy 2014). However, there are only very few examples of data assimilation for large-scale water cycle modelling and operational water resources assessment so far (Renzullo et al. 2014), presumably because of limited availability of usable data at this scale, but also because of the complexity and computational costs of these techniques (van Dijk et al. 2014). The most widely adopted technique is the ensemble Kalman filter (EnKF) or Smoother, where the otherwise unknown error characteristics of the model are estimated by a Monte Carlo-based ensemble approach to determine the error covariance matrix of the model. First assimilation studies using GRACE TWS show its value for informing simulated subsurface water storage (Zaitchik et al. 2008; Li et al. 2012; Houborg et al. 2012). Eicker et al. (2014) presented an EnKF approach for assimilating GRACE TWS into WaterGAP with combined state and parameter updating and a full error propagation from the monthly GRACE spherical harmonic coefficients. They showed that GRACE data inform the model even at higher spatial resolution than resolved by the GRACE data themselves, with varying gains in time, space and among the different storage compartments. Some of the major challenges for the further development of data assimilation techniques include strategies for conserving mass during the assimilation process, error characterization of model and observation data, and adequate mapping functions between observed and simulated variables.

Data assimilation or fusion techniques can also be applied in an offline mode to provide consistent water balance estimates of continental hydrology. A global water cycle reanalysis product has recently been presented by van Dijk et al. (2014), merging prior estimates of monthly water storage changes based on an ensemble of several LSM outputs and complementary data with GRACE TWS data in a sequential data assimilation framework.

4.3 Hyperresolution Global Hydrological Modelling

There is the vision that one day it may be feasible to perform global-scale hydrological modelling with an acceptable accuracy at a much higher resolution than today, with grid cells of 100 m to 1 km instead of the current 50 km (Wood et al. 2011). Then, global-scale modelling would allow improved global assessment of, e.g., food security and could support river basin management everywhere. This has particular relevance in developing countries where basin models are not yet available or a poorly constrained because of lack of local data; in these cases, information about water resources derived from GHMs that can exploit non-local remote sensing data would be a great asset if it is locally relevant (Bierkens et al. 2015). In addition, highly resolved global-scale information on water flows and storages would be very beneficial for freshwater ecosystem management and for assessing global biogeochemical cycles.

The HyperHydro initiative (www.hyperhydro.org) aims at advancing hyperresolution global hydrological modelling. It is a network of scientists that is open to the broader scientific community and invites anyone who wishes to cooperate. Current efforts include the establishment of testbeds, overcoming of computational challenges and the compilation of input data sets. For further information on motivation, challenges and prospects of hyperresolution global hydrological modelling, please refer to Wood et al. (2011) (including a comment of Beven and Cloke 2012 and the reply of Wood et al. 2012) and Bierkens et al. (2015). Beven et al. (2015) provide valuable critical comments on hyperresolution modelling of water on the land areas of the globe, pointing out that unknown heterogeneities in the subsurface and ignorance about subsurface processes result in a lower gain of accuracy by increased resolution than is the case in atmosphere and ocean modelling. However, a prospect of hyperresolution modelling is that its output can be evaluated more meaningfully than current GHM output by local experts and stakeholders who can help identify model deficiencies.

5 Conclusions

The capabilities and the sheer number of GHMs have increased significantly over the last decade such that global-scale quantification of water resources has improved and uncertainties are better known. We conclude that major challenges remain until GHMs can serve as reliable tools for characterizing current and potential future water resources worldwide. We hope that our presentation of selected challenges informs not only on the state-of-the-art of global hydrological modelling but also indicates fruitful research directions. As outlined in the previous section, we believe that major advancements will be possible if in situ and remotely sensed observational data of model output variables are utilized more efficiently in global hydrological modelling and if spatially more resolved model output can be provided with reasonable accuracy.

References

Alcamo J, Flörke M, Märker M (2007) Future long-term changes in global water resources driven by socio-economic and climatic changes. Hydrol Sci J 52(2):247–275. doi:10.1623/hysj.52.2.247

Alkama R, Decharme B, Douville H, Becker M, Cazenave A, Sheffield J, Voldoire V, Tyteca S, Le Moigne P (2010) Global evaluation of the ISBA-TRIP continental hydrologic system. Part 1: a twofold constraint using GRACE Terrestrial Water Storage estimates and in situ river discharges. J Hydrometeorol 11:583–600. doi:10.1175/2010JHM1211.1

Alkama R, Decharme B, Douville H, Ribes A (2011) Trends in global and basin-scale runoff over the late 20th century? Methodological issues and sources of uncertainty. J Climate 24:2983–2999. doi:10.1175/2010JCLI3921.1

Arnell NW, Lloyd-Hughes B (2014) The global-scale impacts of climate change on water resources and flooding under new climate and socio-economic scenarios. Clim Change 122(1–2):127–140. doi:10.1007/s10584-013-0948-4

Beven K, Cloke H (2012) Comment on "Hyperresolution global land surface modeling: meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al. Water Resour Res 48:W01801. doi:10.1029/2011WR010982

Beven KJ, Freer J (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems. J Hydrol 249:11–29. doi:10.1016/S0022-1694(01)00421-8

Beven K, Cloke H, Pappenberger F, Lamb R, Hunter N (2015) Hyperresolution information and hyper-resolution ignorance in modelling the hydrology of the land surface. Sci China Earth Sci 58:25–35. doi:10.1007/s11430-014-5003-4

Biemans H, Hutjes RWA, Kabat P, Strengers BJ, Gerten D, Rost S (2009) Effects of precipitation uncertainty on discharge calculations for main river basins. J Hydrometeorol 10(4):1011–1025. doi:10.1175/2008JHM1067.1

Bierkens MFP, Bell VA, Burek P, Chaney N, Condon LE, David CH, de Roo A, Döll P, Drost N, Famiglietti JS, Flörke M, Gochis DJ, Houser P, Hut R, Keune J, Kollet S, Maxwell RM, Reager JT, Samaniego L, Sudicky E, Sutanudjaja EH, van de Giesen N, Winsemius H, Wood EF (2015) Hyper-resolution global hydrological modelling: what is next? Hydrol Process 29(2):310–320. doi:10.1002/hyp.10391

Brunner P, Simmons CT (2012) HydroGeoSphere: a fully integrated, physically based hydrological model. Groundwater 50:170–176. doi:10.1111/j.1745-6584.2011.00882.x

Chen J, Brissette FP, Lucas-Picher P (2015) Assessing the limits of bias-correcting climate model outputs for climate change impact studies. J Geophys Res Atmos 120:1123–1136. doi:10.1002/2014JD022635

Coxon G, Freer J, Westerberg IK, Wagener T, Woods R, Smith PJ (2015) A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. Water Resour Res 51:5531–5546. doi:10.1002/2014WR016532

Dankers R, Arnell NW, Clark DB, Falloon PD, Fekete BM, Gosling SN, Heinke J, Kim H, Masaki Y, Satoh Y, Stacke T, Wada Y, Wisser D (2014) First look at changes in flood hazard in the Inter-Sectoral Impact Model Intercomparison Project ensemble. Proc Natl Acad Sci USA 111(9):3257–3261. doi:10.1073/pnas.1302078110

Davie JCS, Falloon PD, Kahana R, Dankers R, Betts R, Portmann FT, Wisser D, Clark DB, Ito A, Masaki Y, Nishina K, Fekete B, Tessler Z, Wada Y, Liu X, Tang Q, Hagemann S, Stacke T, Pavlick R, Schaphoff S, Gosling SN, Franssen W, Arnell NW (2013) Comparing projections of future changes in runoff from hydrological and biome models. Earth Syst Dynam 4:359–374. doi:10.5194/esd-4-359-2013

de Graaf IEM, Sutanudjaja EH, van Beek LPH, Bierkens MFP (2015) A high-resolution global-scale groundwater model. Hydrol Earth Syst Sci 19:823–837. doi:10.5194/hess-19-823-2015

Döll P, Fiedler K (2008) Global-scale modeling of groundwater recharge. Hydrol Earth Syst Sci 12(3):863–885. doi:10.5194/hess-12-863-2008

Döll P, Siebert S (2002) Global modeling of irrigation water requirements. Water Resour Res 38(4):8-1–8-10. doi:10.1029/2001WR000355

Döll P, Kaspar F, Lehner B (2003) A global hydrological model for deriving water availability indicators: model tuning and validation. J Hydrol 270(1–2):105–134. doi:10.1016/S0022-1694(02)00283-4

Döll P, Fiedler K, Zhang J (2009) Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. Hydrol Earth Syst Sci 13:2413–2432. doi:10.5194/hess-13-2413-2009

Döll P, Hoffmann-Dobrev H, Portmann FT, Siebert S, Eicker A, Rodell M, Strassberg G, Scanlon B (2012) Impact of water withdrawals from groundwater and surface water on continental water storage variations. J Geodyn 59–60:143–156. doi:10.1016/j.jog.2011.05.001

Döll P, Fritsche M, Eicker A, Müller Schmied H (2014a) Seasonal water storage variations as impacted by water abstractions: comparing the output of a global hydrological model with GRACE and GPS observations. Surv Geophys 35(6):1311–1331. doi:10.1007/s10712-014-9282-2

Döll P, Müller Schmied H, Schuh C, Portmann F, Eicker A (2014b) Global-scale assessment of groundwater depletion and related groundwater abstractions: combining hydrological modeling with information from well observations and GRACE satellites. Water Resour Res 50:5698–5720. doi:10.1002/2014WR015595

Döll P, Jiménez Cisneros B, Oki T, Arnell NW, Benito G, Cogley JG, Jiang T, Kundzewicz ZW, Mwakalila S, Nishijima A (2015) Integrating risks of climate change into water management. J Hydrol Sci 60(1):3–14. doi:10.1080/02626667.2014.967250

Douville H, Decharme B, Ribes A, Alkama R, Sheffield J (2012) Anthropogenic influence on multi-decadal changes in reconstructed global evapotranspiration. Nat Clim Change. doi:10.1038/NCLIMATE1632

Douville H, Voldoire A, Geoffroy O (2015) The recent global warming hiatus: What is the role of Pacific variability? Geophys Res Lett. doi:10.1002/2014GL062775

Duethmann D, Peters J, Blume T, Vorogushyn S, Güntner A (2014) The value of satellite-derived snow cover images for calibrating a hydrological model in snow-dominated catchments in Central Asia. Water Resour Res 50(3):2002–2021. doi:10.1002/2013WR014382

Eicker A, Schumacher M, Kusche J, Döll P, Müller Schmied H (2014) Calibration/data assimilation approach for integrating GRACE data into the WaterGAP Global Hydrology Model (WGHM) using an ensemble Kalman filter: first results. Surv Geophys 35(6):1285–1309. doi:10.1007/s10712-014-9309-8

Elliott J, Deryng D, Müller C, Frieler K, Konzmann M, Gerten D, Glotter M, Flörke M, Wada Y, Eisner S, Folberth C, Foster I, Gosling SN, Haddeland I, Khabarov N, Ludwig F, Masaki Y, Olin S, Rosenzweig C, Ruane AC, Satoh Y, Schmid E, Stacke T, Tang Q, Wisser D (2014) Constraints and potentials of future irrigation water availability on agricultural production under climate change. Proc Natl Acad Sci USA 111(9):3239–3244. doi:10.1073/pnas.1222474110

Fan Y, Miguez-Macho G, Weaver CP, Walko R, Robock A (2007) Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations. J Geophys Res Atmos 112:D10125. doi:10.1029/2006JD008111

Fan Y, Li H, Miguez-Macho G (2013) Global patterns of groundwater table depth. Science 339:940–943. doi:10.1126/science.1229881

Flörke M, Kynast E, Bärlund I, Eisner S, Wimmer F, Alcamo J (2013) Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: a global simulation study. Global Environ Change 23:144–156. doi:10.1016/j.gloenvcha.2012.10.018

Gedney N, Cox PM, Betts RA, Boucher O, Huntingford C, Stott PA (2006) Detection of a direct carbon dioxide effect in continental river runoff records. Nature 439:835–838. doi:10.1038/nature04504

Gedney N, Huntingford C, Weedon GP, Bellouin N, Boucher O, Cox PM (2014) Detection of solar dimming and brightening effects on Northern Hemisphere river flow. Nature Geosci 7:796–800. doi:10.1038/ngeo2263

Gerten D, Betts R, Döll P (2014) Cross-chapter box on the active role of vegetation in altering water flows under climate change. In: Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL (eds) Climate Change 2014: impacts, adaptation, and vulnerability. Part A: global and sectoral aspects. Contribution of working group II to the 5th assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, pp 157–161

Gleeson T, Moosdorf N, Hartmann J, van Beek LPH (2014) A glimpse beneath earth's surface: gLobal HYdrogeological MaPS (GLHYMPS) of permeability and porosity. Geophys Res Lett 41:3891–3898. doi:10.1002/2014GL059856

Gosling SN, Taylor RG, Arnell NW, Todd MC (2011) A comparative analysis of projected impacts of climate change on river runoff from global and catchment-scale hydrological models. Hydrol Earth Syst Sci 15:279–294. doi:10.5194/hess-15-279-2011

Gudmundsson L, Tallaksen LM, Stahl K, Clark DB, Dumont E, Hagemann S, Bertrand N, Gerten D, Heinke J, Hanasaki N, Voss F, Koirala S (2012) Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe. J Hydrometeorol 13(2):604–620. doi:10.1175/JHM-D-11-083.1

Güntner A (2008) Improvement of global hydrological models using GRACE data. Surv Geophys 29(4):375–397. doi:10.1007/s10712-008-9038-y

Guo Z, Dirmeyer PA, Hu Z-Z, Gao X, Zhao M (2006) Evaluation of the Second Global Soil Wetness Project soil moisture simulations: 2. Sensitivity to external meteorological forcing. J Geophys Res 111(D22):D22S03. doi:10.1029/2006JD007845

Gupta HV, Bastidas LA, Sorooshian S, Shuttleworth WJ, Yang ZL (1999) Parameter estimation of a land surface scheme using multicriteria methods. J Geophys Res 104(D16):491–503. doi:10.1029/1999JD900154

Haddeland I, Clark DB, Franssen W, Ludwig F, Voß F, Arnell NW, Bertrand N, Best M, Folwell S, Gerten D, Gomes S, Gosling SN, Hagemann S, Hanasaki N, Harding R, Heinke J, Kabat P, Koirala S, Oki T, Polcher J, Stacke T, Viterbo P, Weedon GP, Yeh P (2011) Multi-model estimate of the global terrestrial water balance: setup and first results. J Hydrometeorol 12:869–884. doi:10.1175/2011JHM1324.1

Haddeland I, Heinke J, Voß F, Eisner S, Chen C, Hagemann S, Ludwig F (2012) Effects of climate model radiation, humidity and wind estimates on hydrological simulations. Hydrol Earth Syst Sci 16:305–318. doi:10.5194/hess-16-305-2012

Haddeland I, Heinke J, Biemans H, Eisner S, Flörke M, Hanasaki N, Konzmann M, Ludwig F, Masaki Y, Schewe J, Stacke T, Tessler Z, Wada Y, Wisser D (2014) Global water resources affected by human interventions and climate change. Proc Natl Acad Sci USA 111(9):3251–3256. doi:10.1073/pnas.1302078110

Hagemann S, Chen C, Haerter JO, Heinke J, Gerten D, Piani C (2011) Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. J Hydrometeorol 12(4):556–578. doi:10.1175/2011JHM1336.1

Hagemann S, Chen C, Clark DB, Folwell S, Gosling SN, Haddeland I, Hanasaki N, Heinke J, Ludwig F, Voss F, Wiltshire AJ (2013) Climate change impact on available water resources obtained using multiple global climate and hydrology models. Earth Syst Dynam 4:129–144. doi:10.5194/esd-4-129-2013

Hanasaki N, Kanae S, Oki T, Masuda K, Motoya K, Shirakawa N, Shen Y, Tanaka K (2008) An integrated model for the assessment of global water resources—part 2: applications and assessments. Hydrol Earth Syst Sci 12(4):1027–1037. doi:10.5194/hess-12-1027-2008

Harris I, Jones PD, Osborn TJ, Lister DH (2014) Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset. Int J Climatol 34(3):623–642. doi:10.1002/joc.3711

Hartmann J, Moosdorf N (2012) The new global lithological map database GLiM: a representation of rock properties at the Earth surface. Geochem Geophys Geosyst 13:Q12004. doi:10.1029/2012GC004370

Hawkins E, Sutton R (2011) The potential to narrow uncertainty in projections of regional precipitation change. Clim Dyn 37:407–418. doi:10.1007/s00382-010-0810-6

Hoekstra AY, Mekonnen MM (2011) Global water scarcity: monthly blue water footprint compared to blue water availability for the world's major river basins. Value of Water Research Report Series No. 53. UNESCO-IHE Institute for Water Education. Delft, The Netherlands

Houborg R, Rodell M, Li BL, Reichle R, Zaitchik BF (2012) Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. Water Resour Res 48:W07525. doi:10.1029/2011WR011291

Hunger M, Döll P (2008) Value of river discharge data for global-scale hydrological modeling. Hydrol Earth Syst Sci 12(3):841–861. doi:10.5194/hess-12-841-2008

Jägermeyr J, Gerten D, Heinke J, Schaphoff S, Kummu M, Lucht W (2015) Water savings potentials of irrigation systems: global simulation of processes and linkages. Hydrol Earth Syst Sci 19:3073–3091. doi:10.5194/hess-19-3073-2015

Jiménez C, Prigent C, Mueller B, Seneviratne SI, McCabe MF, Wood EF, Rossow WB, Balsamo G, Betts AK, Dirmeyer PA, Fisher JB, Jung M, Kanamitsu M, Reichle RH, Reichstein M, Rodell M, Sheffield J, Tu K, Wang K (2011) Global intercomparison of 12 land surface heat flux estimates. J Geophys Res. doi:10.1029/2010JD014545

Jiménez Cisneros BE, Oki T, Arnell NW, Benito G, Cogley JG, Döll P, Jiang T, Mwakalila SS (2014) Freshwater resources. In: Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL (eds). Climate Change 2014: impacts, adaptation, and vulnerability. Part a: global and sectoral aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, pp 229–269

Joetzjer E, Delire C, Douville H, Ciais P, Decharme B, Fisher R, Christoffersen B, Calvet JC, da Costa ACL, Ferreira LV, Meir P (2014) Predicting the response of the Amazon rainforest to persistent drought conditions under current and future climates: a major challenge for global land surface models. Geosci Model Dev 7:2933–2950. doi:10.5194/gmd-7-2933-2014

Jones JP, Sudicky EA, McLaren RG (2008) Application of a fully-integrated surface-subsurface flow model at the watershed-scale: a case study. Water Resour Res 44:W03407. doi:10.1029/2006wr005603

Jung M, Reichstein M, Ciais P, Seneviratne SI, Sheffield J, Goulden ML, Bonan G, Cescatti A, Chen J, de Jeu R, Dolman AJ, Eugster W, Gerten D, Gianelle D, Gobron N, Heinke J, Kimball J, Law BE, Montagnani L, Mu Q, Mueller B, Oleson K, Papale D, Richardson AD, Roupsard O, Running S, Tomelleri E, Viovy N, Weber U, Williams C, Wood E, Zaehle S, Zhang K (2010) Recent decline in the global land evapotranspiration trend due to limited moisture supply. Nature 467:951–954. doi:10.1038/nature09396

Jung M, Reichstein M, Margolis HA, Cescatti A, Richardson AD, Arain MA, Arneth A, Bernhofer C, Bonal D, Chen J, Gianelle D, Gobron N, Kiely G, Kutsch W, Lasslop G, Law BE, Lindroth A, Merbold L, Montagnani L, Moors EJ, Papale D, Sottocornola M, Vaccari F, Williams C (2011) Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. J Geophys Res Biogeosci 116(G3):2156–2202. doi:10.1029/2010JG001566

Kiguchi M, Shen Y, Kanae S, Oki T (2015) Re-evaluation of future water stress due to socio-economic and climate factors under a warming climate. Hydrol Sci J 60(1–2):14–29. doi:10.1080/02626667.2014.888067

Knutti R, Sedlacek J (2013) Robustness and uncertainties in the new CMIP5 climate model projections. Nat Clim Change 3:369–373. doi:10.1038/nclimate1716

Kollet SJ, Maxwell RM (2008) Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. Water Resour Res 44:W02402. doi:10.1029/2007WR006004

Krakauer NY, Li H, Fan Y (2014) Groundwater flow across spatial scales: importance for climate modeling. Environ Res Lett 9:034003. doi:10.1088/1748-9326/9/3/034003

Krogh PE (2011) Large scale hydrological model calibration with remote sensing data from GRACE. PhD thesis, DTU space, National Space Institute, Technical University of Denmark, Copenhagen

Lahoz WA, De Lannoy GJM (2014) Closing the gaps in our knowledge of the hydrological cycle over land: conceptual problems. Surv Geophys 35(3):623–660. doi:10.1007/s10712-013-9221-7

Larson KM, Small EE, Gutmann ED, Bilich AL, Braun JJ, Zavorotny VU (2008) Use of GPS receivers as a soil moisture network for water cycle studies. Geophys Res Lett 35(24):L24405. doi:10.1029/2008GL036013

Li BL, Rodell M, Zaitchik BF, Reichle RH, Koster RD, van Dam TM (2012) Assimilation of GRACE terrestrial water storage into a land surface model: evaluation and potential value for drought monitoring in western and central Europe. J Hydrol 446:103–115. doi:10.1175/2007JHM951.1

Livneh B, Lettenmaier DP (2012) Multi-criteria parameter estimation for the Unified Land Model. Hydrol Earth Syst Sci 16(8):3029–3048. doi:10.5194/hess-16-3029-2012

Maxwell RM, Chow FK, Kollet SJ (2007) The groundwater-land-surface-atmosphere connection: soil moisture effects on the atmospheric boundary layer in fully-coupled simulations. Adv Water Resour 30:2447–2466. doi:10.1016/j.advwatres.2007.05.018

Maxwell RM, Condon LE, Kollet SJ (2015) A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. Geosci Model Dev 8:923–937. doi:10.5194/gmd-8-923-2015

Miguez-Macho G, Fan Y, Weaver CP, Walko R, Robock A (2007) Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation. J Geophys Res Atmos 112:D13108. doi:10.1029/2006JD008112

Milzow C, Krogh PE, Bauer-Gottwein P (2011) Combining satellite radar altimetry, SAR surface soil moisture and GRACE total storage changes for hydrological model calibration in a large poorly gauged catchment. Hydrol Earth Syst Sci 15(6):1729–1743. doi:10.5194/hess-15-1729-2011

Müller Schmied H, Eisner S, Franz D, Wattenbach M, Portmann FT, Flörke M, Döll P (2014) Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration. Hydrol Earth Syst Sci 18:3511–3538. doi:10.5194/hess-18-3511-2014

Murray SJ, Foster PN, Prentice IC (2012) Future global water resources with respect to climate change and water withdrawals as estimated by a dynamic global vegetation model. J Hydrol 448–449:14–29. doi:10.1016/j.jhydrol.2012.02.044

Nasonova ON, Gusev YM, Kovalev YE (2011) Impact of uncertainties in meteorological forcing data and land surface parameters on global estimates of terrestrial water balance components. Hydrol Process 25(7):1074–1090. doi:10.1002/hyp.7651

Niu GY, Yang ZL, Dickinson RE, Gulden LE (2007) Su H (2007) Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. J Geophys Res 112:D07103. doi:10.1029/2006JD007522

Pall P, Aina T, Stone DA, Stott PA, Nozawa T, Hilberts AGJ, Lohmann D, Allen MR (2011) Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. Nature 470(7334):382–385. doi:10.1038/nature09762

Papa F, Frappart F, Güntner A, Prigent C, Aires F, Getirana ACV, Maurer R (2013) Surface freshwater storage and variability in the Amazon basin from multi-satellite observations, 1993–2007. J Geophys Res Atmos 118(21):11951–11965. doi:10.1002/2013JD020500

Parajka J, Blöschl G (2008) The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models. J Hydrol 358(3–4):240–258. doi:10.1016/j.jhydrol.2008.06.006

Portmann F, Siebert S, Döll P (2010) MIRCA2000—global monthly irrigated and rainfed crop areas around the year 2000: a new high resolution data set for agricultural and hydrological modelling. Global Biogeochem Cy 24:GB1011. doi:10.1029/2008GB003435

Reichle RH, De Lannoy GJM, Forman BA, Draper CS, Liu Q (2014) Connecting satellite observations with water cycle variables through land data assimilation: examples using the NASA GEOS-5 LDAS. Surv Geophys 35(3):577–606. doi:10.1007/s10712-013-9220-8

Renzullo LJ, van Dijk AIJM, Perraud JM, Collins D, Henderson B, Jin H, Smith AB, McJannet DL (2014) Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment. J Hydrol 519:2747–2762. doi:10.1016/j.jhydrol.2014.08.008

Rost S, Gerten D, Bondeau A, Lucht W, Rohwer J, Schaphoff S (2008) Agricultural green and blue water consumption and its influence on the global water system. Water Resour Res 44:W09405. doi:10.1029/2007WR006331

Samaniego L, Kumar R, Attinger S (2010) Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. Water Resour Res 46:W05523. doi:10.1029/2008WR007327

Schewe J, Heinke J, Gerten D, Haddeland I, Arnell NW, Clark DB, Dankers R, Eisner S, Fekete BM, Colon-Gonzalez FJ, Gosling SN, Kim H, Liu X, Masaki Y, Portmann FT, Satoh Y, Stacke T, Tang Q, Wada Y, Wisser D, Albrecht T, Frieler K, Piontek F, Warszawski L, Kabat P (2014) Multimodel assessment of water scarcity under climate change. Proc Natl Acad Sci USA 111(9):3245–3250. doi:10.1073/pnas.1222460110

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M, Rudolf B (2014) GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theor Appl Climatol 115:15–40. doi:10.1007/s00704-013-0860-x

Sheffield J, Goteti G, Wood EF (2006) Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. J Clim 19:3088–3111. doi:10.1175/JCLI3790.1

Siebert S, Döll P (2010) Quantifying blue and green water uses and virtual water contents in global crop production as well as potential production losses without irrigation. J Hydrol 384:198–217. doi:10.1016/j.jhydrol.2009.07.031

Sitch S, Huntingford C, Gedney N, Levy PE, Lomas M, Piao SL, Betts R, Ciais P, Cox P, Friedlingstein P, Jones CD, Prentice IC, Woodward FI (2008) Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five dynamic global vegetation models (DGVMs). Global Change Biol 14:2015–2039. doi:10.1111/j.1365-2486.2008.01626.x

Sood A, Smakhtin V (2015) Global hydrological models: a review. Hydrol Sci J. doi:10.1080/02626667.2014.950580

Sterling SM, Ducharne A, Polcher J (2013) The impact of global land-cover change on the terrestrial water cycle. Nat Clim Change 3:385–390. doi:10.1038/nclimate1690

Tapley BD, Bettadpur S, Watkins M, Reigber C (2004) The gravity recovery and climate experiment: mission overview and early results. Geophys Res Lett 31:L09607. doi:10.1029/2004GL019920

Taylor RG, Scanlon B, Döll P, Rodell M, van Beek R, Wada Y, Longuevergne L, Leblanc M, Famiglietti JS, Edmunds M, Konikow L, Green TR, Chen J, Taniguchi M, Bierkens MFP, MacDonald A, Fan Y, Maxwell RM, Yechieli Y, Gurdak JJ, Allen DM, Shamsudduha M, Hiscock K, Yeh PJF, Holman I, Treidel H (2013) Ground water and climate change. Nat Clim Change 3:322–329. doi:10.1038/nclimate1744

van Dijk AIJM, Renzullo LJ, Wada Y, Tregoning P (2014) A global water cycle reanalysis (2003–2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble. Hydrol Earth Syst Sci 18:2955–2973. doi:10.5194/hess-18-2955-2014

Vassolo S, Döll P (2005) Global-scale gridded estimates of thermoelectric power and manufacturing water use. Wat Resour Res 41:W04010. doi:10.1029/2004WR003360

Vergnes J-P, Decharme B, Alkama R, Martin E, Habets F, Douville H (2012) A simple groundwater scheme for hydrological and climate applications: description and off-line evaluation over France. J Hydrometeorol 13:1149–1171. doi:10.1175/JHM-D-11-0149

Vergnes J-P, Decharme B, Habets F (2014) Introduction of groundwater capillary rises using subgrid spatial variability of topography into the ISBA land surface model. J Geophys Res Atmos 119:11065–11086. doi:10.1002/2014JD021573

Vörösmarty CJ, Sahagian D (2000) Anthropogenic disturbance of the terrestrial water cycle. Bioscience 50(9):753–765. doi:10.1641/0006-3568(2000)050[0753:ADOTTW]2.0.CO;2

Vörösmarty CJ, Hoekstra AY, Bunn SE, Conway D, Gupta J (2015) Fresh water goes global. Science 349(6247):478–479. doi:10.1126/science.aac6009

Wada Y, van Beek LPH, Bierkens MFP (2011) Modelling global water stress of the recent past: on the relative importance of trends in water demand and climate variability. Hydrol Earth Syst Sci 15:3785–3808. doi:10.5194/hess-15-3785-2011

Wada Y, van Beek LPH, Bierkens MFP (2012) Nonsustainable groundwater sustaining irrigation: a global assessment. Water Resour Res 48:W00L06. doi:10.1029/2011WR010562

Wada Y, Wisser D, Eisner S, Flörke M, Gerten D, Haddeland I, Hanasaki N, Masaki Y, Portmann FT, Stacke T, Tessler Z, Schewe J (2013) Multi-model projections and uncertainties of irrigation water demand under climate change. Geophys Res Lett 40:4626–4632. doi:10.1002/grl.50686

Wada Y, Wisser D, Bierkens MFP (2014) Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. Earth Syst Dynam 5:15–40. doi:10.5194/esd-5-15-2014

Wang K, Dickinson ED, Wild M, Liang S (2010) Evidence for decadal variation in global terrestrial evapotranspiration between 1982 and 2002: 2. Results. J Geophys Res 115:D20113. doi:10.1029/2009JD013671

Warszawski L, Friend A, Ostberg S, Frieler K, Lucht W, Schaphoff S, Beerling D, Cadule P, Ciais P, Clark DB, Kahana R, Ito A, Keribin R, Kleidon A, Lomas M, Nishina K, Pavlick R, Rademacher TT, Buechner M, Piontek F, Schewe J, Serdeczny O, Schellnhuber HJ (2013) A multi-model analysis of risk of ecosystem shifts under climate change. Environ Res Lett 8:044018. doi:10.1088/1748-9326/8/4/044018

Weedon GP, Balsamo G, Bellouin N, Gomes S, Best M, Viterbo P (2014) The WFDEI meteorological forcing data set: WATCH forcing data methodology applied to ERA-interim reanalysis data. Water Resour Res 50(9):7505–7514. doi:10.1002/2014WR015638

Werth S, Güntner A (2010) Calibration analysis for water storage variability of the global hydrological model WGHM. Hydrol Earth Syst Sci 14(1):59–78. doi:10.5194/hess-14-59-2010

Werth S, Güntner A, Schmidt R, Petrovic S (2009) Integration of GRACE mass variations into a global hydrological model. Earth Planet Sci Lett 277:166–173. doi:10.1016/j.epsl.2008.10.021

Widén-Nilsson E, Halldin S, Xu C (2007) Global water-balance modelling with WASMOD-M: parameter estimation and regionalization. J Hydrol 340:105–118. doi:10.1016/j.jhydrol.2007.04.002

Wisser D, Frolking S, Douglas EM, Fekete BM, Vörösmarty CJ, Schumann AH (2008) Global irrigation water demand: variability and uncertainties arising from agricultural and climate data sets. Geophys Res Lett 35(24):L24408. doi:10.1029/2008GL035296

Wood E, Roundy JK, Troy TJ, van Beek R, Bierkens M, Blyth E, de Roo A, Döll P, Ek M, Famiglietti J, Gochis D, van de Giesen N, Houser P, Jaffe P, Kollet S, Lehner B, Lettenmaier DP, Peters-Lidard C, Sivapalan M, Sheffield J, Wade A, Whitehead P (2011) Hyper-resolution global land surface modeling: meeting a grand challenge for monitoring Earth's terrestrial water. Water Resour Res 47:W05301. doi:10.1029/2010WR010090

Wood E, Roundy JK, Troy TJ, van Beek R, Bierkens M, Blyth E, de Roo A, Döll P, Ek M, Famiglietti J, Gochis D, van de Giesen N, Houser P, Jaffe P, Kollet S, Lehner B, Lettenmaier DP, Peters-Lidard C., Sivapalan M, Sheffield J, Wade A, Whitehead P (2012) Reply to comment by Keith J. Beven and Hannah L. Cloke on "Hyperresolution global land surface modeling: meeting a grand challenge for monitoring Earth's terrestrial water". Water Resour Res 48:W01802. doi:10.1029/2011WR011202

Xie H, Longuevergne L, Ringler C, Scanlon BR (2012) Calibration and evaluation of a semi-distributed watershed model of Sub-Saharan Africa using GRACE data. Hydrol Earth Syst Sci 16(9):3083–3099. doi:10.5194/hess-16-3083-2012

Zaitchik BF, Rodell M, Reichle RH (2008) Assimilation of GRACE terrestrial water storage data into a land surface model: results for the Mississippi River basin. J Hydrometeorol 9(3):535–548. doi:10.1175/2007JHM951.1

Zhou T, Haddeland I, Nijssen B, Lettenmaier DP (2015) Human induced changes in the global water cycle. AGU Geophysical Monograph Series, Submitted