Beyond the bucket – Developing a global gradient-based groundwater model (G³M v1.0) for a global hydrological model from scratch

Robert Reinecke[1], Laura Foglia[3], Steffen Mehl[4], Tim Trautmann[1], Denise Cáceres[1], Petra Döll[1,2]

[1]Institute of Physical Geography, Goethe University Frankfurt, Frankfurt am Main, Germany
[2]Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Frankfurt am Main, Germany
[3]Department of Land, Air and Water Resources, University of California, Davis, USA
[4]Department of Civil Engineering, California State University, Chico, USA

Correspondence to: Robert Reinecke (reinecke@em.uni-frankfurt.de)

Abstract. To quantify water flows between groundwater (GW) and surface water (SW) as well as the impact of capillary rise on evapotranspiration by global hydrological models (GHMs), it is necessary to replace the bucket-like linear GW reservoir model typical for hydrological models with a fully integrated gradient-based GW flow model. Linear reservoir models can only simulate GW discharge to SW bodies, provide no information on the location of the GW table and assume that there is no GW flow among grid cells. A gradient-based GW model simulates not only GW storage but also hydraulic head, which together with information on SW table elevation enables the quantification of water flows from GW to SW and vice versa. In addition, hydraulic heads are the basis for calculating lateral GW flow among grid cells and capillary rise.

G³M is a new global gradient-based GW model with a spatial resolution of 5' that will replace the current linear GW reservoir in the 0.5° WaterGAP Global Hydrology Model (WGHM). The newly developed model framework enables in-memory coupling to WGHM while keeping overall runtime relatively low, allowing sensitivity analyses and data assimilation. This paper presents the G³M concept and specific model design decisions together with results under steady-state naturalized conditions, i.e. neglecting GW abstractions. Cell-specific conductances of river beds, which govern GW-SW interaction, were determined based on the 30" steady-state water table computed by Fan et al. (2013). Together with an appropriate choice for the effective elevation of the SW table within each grid cell, this enables a reasonable simulation of drainage from GW to SW such that, in contrast to the GW model of de Graaf et al. (2015, 2017), no additional drainage based on externally provided values for GW storage above the floodplain is required in G³M. Comparison of simulated hydraulic heads to observations around the world shows better agreement than de Graaf et al. (2015). In addition, G³M output is compared to the output of two established macro-scale models for the Central Valley, California, and the continental United States, respectively. As expected, depth to GW table is highest in mountainous and lowest in flat regions. A first analysis of losing and gaining rivers and lakes/wetlands indicates that GW discharge to rivers is by far the dominant flow, draining diffuse GW recharge, such that lateral flows only become a large fraction of total diffuse and focused recharge in case of losing rivers and some areas with very low GW recharge. G³M does not represent losing rivers in some dry regions. This study presents the first steps towards replacing the linear GW reservoir model in a GHM while improving on recent efforts, demonstrating the feasibility of the approach and the robustness of the newly developed framework.

1 Introduction

Groundwater (GW) is the source of about 40% of all human water abstractions (Döll et al., 2014). It is also an essential source of water for freshwater biota in rivers, lakes and wetlands, which are in most cases recharged by GW. GW strongly affects river flow regimes and supplies the majority of river water during ecologically and economically critical periods with little precipitation. GW may receive recharge not only from the soil but also from surface water (SW) bodies. In case of small distances between GW table and land surface, GW enhances evapotranspiration via capillary rise. GW storage and flow dynamics have been altered by human GW abstractions as well as climate change and will continue to change in the future

(Taylor et al., 2012). Around the globe, GW abstractions have led to lowered water tables and, in some regions, even GW depletion (Konikow, 2011; Scanlon et al., 2012; Wada et al., 2012; Döll et al., 2014). This has resulted in reduced base flows to rivers and wetlands (with negative impacts on water quality and freshwater ecosystems), land subsidence and increased pumping costs (Gleeson et al., 2012; Döll et al., 2014; Gleeson et al., 2016; Wada, 2016). The strategic importance of GW for

5   global water and food security will probably intensify under climate change as more frequent and intense climate extremes increase variability of SW flows (Taylor et al., 2012). International efforts have been made to promote sustainable GW management and knowledge exchange among countries, e.g., UNESCO's program on International Shared Aquifer Resources Management (ISARM) (http://isarm.org) and the ongoing GW component of the Transboundary Waters Assessment Programme (TWAP) (http://www.geftwap.org). To support priority setting for investment among transboundary aquifers as

10   well as identification of strategies for sustainable GW management, information on current conditions and possible trends of the GW systems is required (UNESCO-IHP, IGRAC, WWAP, 2012). In a globalized world, an improved understanding of GW systems and their interaction with SW and soil is needed not only at the local and regional but also at the global scale.

To assess GW at the global scale, global hydrological models (GHMs) are used (e.g., (Wada et al., 2012; Döll et al., 2012, 2014; Wada, 2016)). In particular, they serve to quantify GW recharge (Döll and Fiedler, 2008). Like typical

15   hydrological models at any scale, GHMs simulate GW dynamics by a linear reservoir model. In such a model, the temporal change of GW storage in each grid cell is computed from the balance of prescribed inflows and an outflow that is a linear function of GW storage. Linear reservoir models can only simulate GW discharge to SW bodies but not a reversal of this flow, even though losing streams may provide focused GW recharge that allows the aquifer to support ecosystems alongside the GW flow path (Stonestrom et al., 2007) as well as human GW abstractions. This flow direction typically occurs in semi-arid

20   and arid but seasonally also in humid regions. In addition, such linear reservoir models provide no information on the location of the GW table, and assume that GW flow among grid cells is negligible. To simulate the dynamics of water flow between SW bodies and GW in both directions as well as the effect of capillary rise on evapotranspiration, it is necessary to compute lateral GW flows among grid cells as function of hydraulic head gradients and thus the dynamic location of the GW table. To achieve an improved understanding of GW systems at the global scale, and in particular of the interactions of GW with SW

25   and soil, it is therefore necessary to replace the linear GW reservoir model in GHMs by a hydraulic gradient-based GW flow model.

Macro-scale gradient-based GW flow models are still rare and mainly available for data-rich regions, e.g. for the Death Valley (Belcher and Sweetkind, 2010) and the Central Valley (Faunt, 2009; Dogrul et al., 2016) in the USA, but also for large fossil groundwater bodies in arid regions (e.g. the Nubian Aquifer System in North Africa, Gossel et al., 2004). However, they are in

30   most cases not integrated within hydrological models that quantify GW recharge based on climate data and provide information on the condition of SW. For North America, Fan et al. (2007) and Miguez-Macho et al. (2007) linked a land surface model with a two-dimensional gradient-based GW model and computed, with a daily time step, gradient-based GW flow, water table elevation, GW–SW interaction and capillary rise, using a spatial resolution of 1.25 km. One challenge was the determination of the river conductance that affects the degree of GW–SW interaction. A computationally very expensive integrated simulation of dynamic SW, soil and GW

35   flow using Richards' equation for variably saturated flow was achieved at a spatial resolution of 1 km for the continental US by applying the ParFlow model (Maxwell et al., 2015). In both studies, GW abstractions were not taken into account.

A first simulation of the steady-state GW table for the whole globe at the very high resolution of 30" was presented by Fan et al. (2013) and compared to an extensive compilation of observed hydraulic heads. However, there was no head-based interactions with SW; GW above the land surface was simply discarded. Global GW flow modeling is strongly hampered

40   by data availability, including the geometry of aquifers and aquitards as well as parameters like hydraulic conductivity (de Graaf et al., 2017), and by computational restrictions on spatial resolution leading to conceptual problems, e.g., regarding SW-GW interactions (Morel-Seytoux et al., 2017). In the last years, some GW flow models that are in principle applicable for the global scale were developed but were applied only regionally in data-rich regions (Rhine basin: Sutanudjaja et al., 2011; France: Vergnes et al., 2012; Vergnes et al., 2014). The first global gradient-based GW model that was run for both steady-

state (de Graaf et al., 2015) and transient conditions (de Graaf et al., 2017) was driven by GW recharge and SW data of the GHM PCR-GLOBWB (van Beek et al., 2011). However, there is not yet a two-way coupling of a GW flow model and a GHM. This may be due to the way de Graaf et al. (2015, 2017) modelled river-GW interaction. To achieve plausible hydraulic head results, they found it necessary to add an additional drainage flux to GW drainage driven by the hydraulic head difference

5   between GW and river. This additional drainage, which accounts for about 50% of global GW drainage, is simulated as a function of GW storage above the floodplain, the values of which are computed externally by the linear GW reservoir model of PCR-GLOBWB (Equation 3 of de Graaf et al. (2017) – the model component that the gradient-based model was intended to replace. This prevents a full integration of the global GW flow model of de Graaf et al. (2017) into a GHM, as then, the linear GW reservoir model would be replaced by the GW flow model.

10   In this study, we present the Global Gradient-based Groundwater Model (G³M) that is to be integrated into the GHM WaterGAP 2. With a spatial resolution of 0.5° by 0.5° (approximately 55 km by 55 km at the equator), the WaterGAP 2 model (Alcamo et al., 2003) computes human water use in five sectors and the resulting net abstractions from GW and SW for all land areas of the globe excluding Antarctica. These net abstractions are then taken from the respective water storages in the WaterGAP Global Hydrology Model (WGHM) (Döll et al., 2003; Döll et al., 2012; Döll et al., 2014; Müller Schmied et al.,

15 2014). With daily time steps, WGHM simulates flows among the water storage compartments canopy, snow, soil, GW, lakes, man-made reservoirs, wetlands and rivers. As in other GHMs, the dynamic of GW storage (GWS) is represented in WGHM by a linear GW reservoir model, i.e.

$$\frac{dGWS}{dt} = R_g + R_{g\_swb} - NA_g - k_g\, GWS \tag{1}$$

where R_g $[L^3T^{-1}]$ is diffuse GW recharge from soil, R_{g_swb} $[L^3T^{-1}]$ GW recharge from lakes, reservoirs and wetlands (only in arid and semiarid regions, with the value per SW body area globally constant), NA_g $[L^3T^{-1}]$ net GW abstraction. The

20 product $k_g\, GWS$ quantifies GW discharge to SW bodies as a function of GWS and the GW discharge coefficient k_g (Döll et al., 2014). G³M is to replace this linear reservoir model in WGHM.

The G³M framework (Reinecke, 2018) was developed to provide full control over the involved processes and allow an optimal in-memory coupling to WGHM. Our model development approach was to learn from existing large-scale regional models (Faunt, 2009; Vergnes et al., 2014; Maxwell et al., 2015; Dogrul et al., 2016) to generate robust information at the

25 global scale, i.e. also in data scarce regions without available regional models. In the next section, the model concept (including the concept for coupling with WGHM) and equations as well as applied data and parameter values are presented. In section 3, we show steady-state results of G³M driven by WGHM data, without any two-way coupling. The stand-alone steady-state simulations have been performed to limit complexity and degrees of freedom without losing important model behaviour, thus providing insights into dominant processes and uncovering possible model-inherent characteristics impossible to observe in a

30 fully coupled transient model. In addition, they served to generate a steady-state solution that can be used as initial condition for future fully coupled transient runs. Simulated hydraulic heads are compared to observations world-wide and to the output of established regional models. We also discuss the influence of scale and modeling decisions and finally draw conclusions.

2 Model description

2.1 G³M model concept

35 Although G³M is based on principles of the well-known GW flow modelling software MODFLOW (Harbaugh, 2005), G³M differs from traditional local and regional GW models. These models are generally based on rather detailed information on hydrogeology (including aquifer geometry and properties such as hydraulic conductivity derived from pumping tests), topography, pumping wells, location and shape of SW bodies as well as on observations of hydraulic head in GW and SW. Local observations guide the developer in constructing the model such that local conditions and processes can be properly

represented. The lateral extent of individual grid cells of such GW flow models is generally smaller or similar to the depth of the aquifer(s) and the size of the SW bodies that interact with the GW. The global GW flow model G³M, however, covers all continents of the Earth except Greenland and Antarctica. At this scale, information listed above is poor or non-existing, and the lateral extent of grid cells needs to be relatively large due to computational (and data) constraints. We selected a grid cell

5   size of 5' by 5' (approx. 9 km by 9 km at the equator), as this size fits well to WaterGAP. WaterGAP 3 (Eisner, 2016) has the same cell size, and 36 of such cells fit to into one 0.5° WaterGAP 2 cell. Global climate data are only available for 0.5° grid cells.

Instead of selecting model layers according to sub-surface information on e.g. texture, we chose to use, in the current version of G³M, two GW layers with a depth of 100 m each. G³M focuses on a plausible simulation of water flows between

10   GW and SW bodies and on the simulation of capillary rise, and we deemed it suitable to have an upper GW layer that interacts with SW and soil and a lower one in which GW may flow laterally without such interactions. As land surface elevation within each 5' grid cell may vary by more than 200 m (Fig. S1), neighbouring cells in G³M may not be adjacent anymore (Fig. 1), in contrast to (regional) GW models with smaller grid cells. This makes G³M a rather conceptual model in which water exchange between groundwater boxes is driven by hydraulic head gradients but flow can no longer be conceptualized as occuring through

15   continuous pore space.

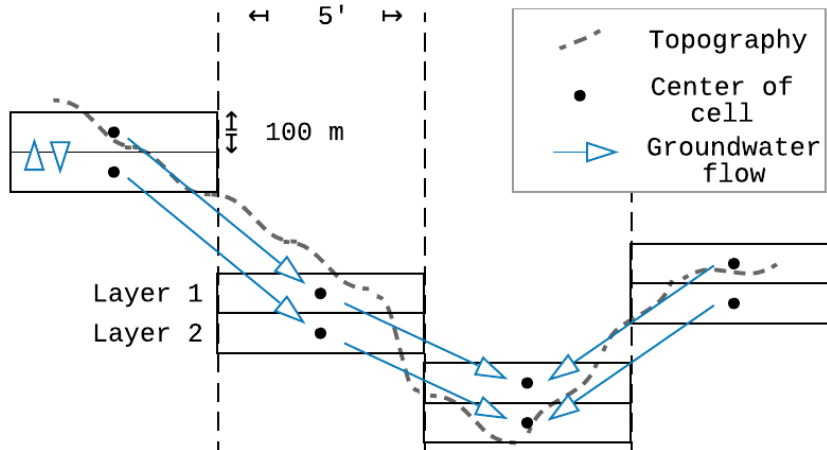



Figure 1 Schematic of G³M spatial structure, with grid cells and GW flow directions, in relation to actual higher-resolution topography.

Three-dimensional groundwater flow is described by a partial differential equation as a function of hydraulic head gradients

$$\frac{dGWS}{dt} = \left(\frac{\partial}{\partial x}\left(K_x \frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_z \frac{\partial h}{\partial z}\right) + W\right) \Delta x \Delta y \Delta z = S_s \frac{\partial h}{\partial t} \Delta x \Delta y \Delta z \tag{2}$$

where $K_{x,y,z}$ is the hydraulic conductivity $[LT^{-1}]$ along the x, y, and z axis between the cells (harmonic mean of grid cell

20   conductivity values), S_s the specific storage $[L^{-1}]$, $\Delta x \Delta y \Delta z$ $[L^3]$ the volume of the cell, and h the hydraulic head $[L]$. Inflows in the groundwater are accounted for as

$$W \Delta x \Delta y \Delta z = R_g + Q_{swb} - NA_g - Q_{cr} + Q_{ocean} \tag{3}$$

where Q_{swb} is flow between the SW bodies (rivers, lakes, reservoirs and wetlands) and GW $[L^3 T^{-1}]$, Q_{cr} is capillary rise, i.e. the flow from GW to the soil, and Q_{ocean} is the flow between ocean and GW $[L^3 T^{-1}]$, representing the boundary condition. In case of Q_{swb} and Q_{ocean}, a positive value represents a flow into the groundwater.

25   The flux across the model domain boundary Q_{ocean} is modeled as a head-dependent flow based on a static head boundary.

$$Q_{ocean} = c_{ocean}(h_{ocean} - h_{aq}) \tag{4}$$

Here h_{ocean} is the elevation of the ocean water table $[L]$, h_{aq} the hydraulic head in the aquifer $[L]$ and c_{ocean} the conductance of the boundary condition $[L^2T^{-1}]$ (Table 1). Q_{cr} is not yet implemented in G³M.

Q_{swb} in Eq. (3) replaces $k_g\ GWS$ and R_{g_swb} in Eq. (1), such that loosing conditions of all types of SW bodies can be simulated dynamically. It is calculated as a function of the difference between the elevation of the water table in the SW bodies h_{swb} $[L]$ and h_{aq} as

$$Q_{swb} = \begin{cases} c_{swb}(h_{swb} - h_{aq}) & h_{aq} > B_{swb} \\ c_{swb}(h_{swb} - B_{swb}) & h_{aq} \le B_{swb} \end{cases} \tag{5}$$

where c_{swb} is the conductance $[L^2T^{-1}]$ of the SW body bed (river, lake, reservoir or wetland) and B_{swb} the SW body bottom elevation $[L]$.

Conductance of SW bodies is often a calibration parameter in traditional GW models (Morel-Seytoux et al., 2017). Following Harbaugh (2005), it can be estimated by

$$c_{swb} = \frac{K\ L\ W}{h_{swb} - B_{swb}} \tag{6}$$

where K is hydraulic conductivity, L is length and W is width of the SW body. For lakes (including reservoirs) and wetlands, c_{swb} is estimated based on hydraulic conductivity of the aquifer K_{aq} and SW body area (Table 1). For gaining rivers, conductance is quantified individually for each grid cell following an approach proposed by Miguez-Macho et al. (2007). The value of river conductance c_{riv}, according to Miguez-Macho et al. (2007), in a GW flow model needs to be set to such a values that, for steady-state conditions, the river is the sink for all the inflow to the grid cell (GW recharge and inflow from neighbouring cells) that is not transported laterally to neighbouring cells such that

$$c_{riv} = \frac{R_g + Q_{eq_{lateral}}}{h_{eq} - h_{riv}} \qquad h_{aq} > h_{riv} \tag{7}$$

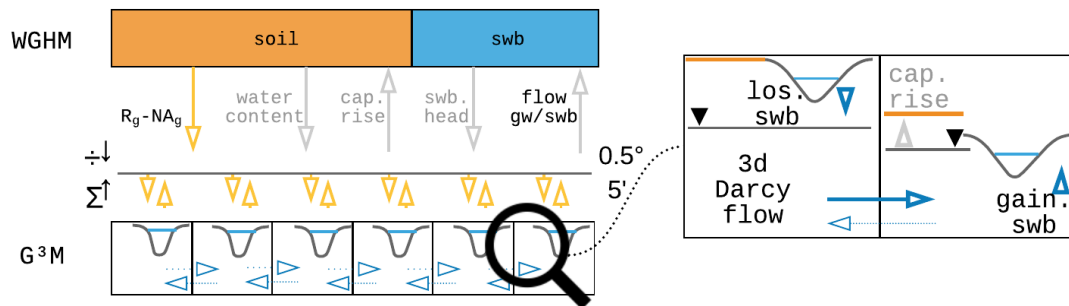
For G³M, we computed the equilibrium head h_{eq} as the 5' average of the 30" steady-state heads calculated by Fan et al. (2013). Using WGHM diffuse GW recharge lateral equilibrium flow $Q_{eq_{lateral}}$ $[L^3T^{-1}]$ is net lateral inflow into the cell computed based on the h_{eq} distribution as well as G³M K_{aq} and cell thickness (Table 1). Elevation of the river water table h_{riv} [L] is to be provided by WGHM. Using a fully dynamic approach, i.e. utilizing the hydraulic head and lateral flows from the current iteration to re-calculate c_{riv} in each iteration towards the steady-state solution, has proven to be too unstable due to its non-linearity affecting convergence. We limit c_{riv} to a maximum of $10^7\ m^2day^{-1}$; this would be approximately the value for a 10 km long and 1 km wide river with a head difference between GW and river of 1 m and hydraulic conductivity of the river bed of 10^{-5} m/s.

If the river recharges the GW (losing river), Eq. (7) cannot be used as the Fan et al. (2015) high-resolution equilibrium model only models groundwater outflows but not inflows from SW bodies. If h_{aq} drops below h_{riv}, Eq. (5) is used to compute c_{riv}, with K equals to K_{aq}.

2.2 Coupling to WGHM

We intend to couple G³M to WaterGAP 2, i.e. the 0.5° version of WGHM. It will not be coupled to WaterGAP 3 (Eisner, 2016), which has the same spatial resolution as G³M, in order to keep computation time low enough for performing sensitivity analyses and ensemble-based data assimilation and calibration. However, data from WaterGAP 3 were used to set up G³M. Location and area of the 5' grid cells of G³M are the same as in the landmask of WaterGAP 3. In addition, the percentage of the 5' grid cell area that is covered by lakes (including reservoirs) and by wetlands, based on Lehner and Döll (2004) is taken from WaterGAP 3, as well as the length and width of the main river within each 5' grid cell as estimated by WaterGAP 3 (Table 1).

G³M will be integrated into WGHM by exchanging information on (1) R_{g_swb} and NA_g, (2) soil water content, (3) Q_{cr}, (4) h_{swb}, and (5) Q_{swb}. Figure 2 indicates the direction of the information flows. Water flows from the 0.5° cells of WGHM are distributed equally to all 5' G³M grid cells inside a 0.5° cell. Flows transferred from the 5' cells of G³M to WGHM are aggregated. GW recharge and net abstraction from GW together with SW tables are the main drivers of the GW model that

5   will be provided dynamically by WGHM. GW-SW flow volumes computed by G³M will be aggregated and added or subtracted from the SW body volumes in WGHM, and SW body heads will be recalculated. WGHM soil water content together with G³M depth to GW will be used to calculate capillary rise and thus a change of soil water content. Capillary rise is not yet implemented in G³M, and SW heads are currently based on land surface elevation.

10   **Figure 2** Conceptual view of the coupling between WGHM and G³M. WGHM provides calculated GW recharge (R_g) (Döll and Fiedler, 2008) and if the human impact is considered, net abstraction from GW (NA_g) (Döll et al., 2012). G³M spreads this input equally to all 5' grid cells inside a 0.5° cell and calculates hydraulic head and interactions with SW bodies (swb) as well as capillary rise (cap. rise) at the 5' resolution. Grey arrows show information flow that is not yet implemented.

2.3 The steady-state uncoupled model version

15   In a first implementation stage, G³M was developed as a steady-state (right-hand side of Eq. (2) is zero) standalone model that represents naturalized conditions (i.e. without taking into account human water use) during 1901-2013. Input data and parameters used are listed in Table 1 and described below.

The landmask of G³M, i.e. location and size of 5' grid cells, is that of WaterGAP 3. The current model version has only two confined layers with an aquifer thickness of 100 m each. Choosing a confined condition instead of an unconfined

20   upper layer provides a stable solution and reduces convergence time, which is favorable for a preliminary analysis of the model behavior. We performed a sensitivity analysis that confirmed the findings of other studies (de Graaf et al., 2015) that the aquifer thickness has a relatively small impact on the model results. Therefore, we assume a uniform thickness of 100 m (motivated by the assumed valid depth of the lithology data) worldwide even though this is an oversimplification as aquifers can reach many times that depth. Elevation of the G³M grid cells is determined by the 5' average of the 30" DEM (Digital Elevation

25   Model) used by Fan et al. (2013) (Fig. S2).

Gleeson et al. (2014) provided a global subsurface permeability data set from which K_{aq} was computed. The data set was derived by relating permeabilities from a large number of local to regional GW models to the type of hydrolithological units (e.g., "unconsolidated" or "crystalline"). The geometric mean permeability values of nine hydrolithological units were mapped to the high-resolution global lithology map GLiM (Hartmann and Moosdorf, 2012). In continuous permafrost areas,

30   a very low permeability value was assumed by Gleeson et al. (2014). The estimated values represent the shallow surface on the scale of 100 m depth. The unique dataset has three inherent problems when used as input for a GW model: (1) At this scale, important heterogeneities such as discrete fractures or connected zones of high hydraulic conductivity controlling the GW flow are not visible. (2) Jurisdictional boundaries due to different data sources in the global lithological map lead to artifacts. (3) The differentiation between coarse and fine-grained unconsolidated deposits is only available in some regions resulting in

35   $10^{-4}\ m\ s^{-1}$ as hydraulic conductivity for coarse-grained unconsolidated deposits. If the distinction is not available, a rather

low value of $10^{-6}\ m\ s^{-1}$ is set for unconsolidated porous media (Fig. S3). The original data was gridded to 5' by using an area-weighted average and used as hydraulic conductivity of the upper model layer. For the second layer, hydraulic conductivity of the first layer is reduced assuming that conductivity decreases exponentially with depth. Based on the e-folding factor f used by Fan et al. (2013) (a calibrated parameter based on terrain slope), conductivity of the lower layer is calculated

5 by multiplying the upper layer value by $\exp(-50\ m\ f^{-1})^{-1}$ (Fan et al., 2007, Eq. 7).

Mean annual GW recharge computed by WaterGAP 2.2c for the period 1901-2013 is used as input (Fig. S4), while no net abstraction from GW was taken into account. It would not be meaningful to try to derive a steady-state solution under existing net groundwater abstractions that in some regions cause GW depletion with continuously dropping water tables. Regarding the ocean boundary condition, h_{ocean} is set to 0 m and c_{ocean} to $10\ m^2\ day^{-1}$ (Table 1).

10 It is assumed that there is exchange of water between GW and one river stretch in each 5' grid cell, and in addition where lakes and wetlands exist according to WaterGAP 3, which provides, for each grid cell, the area of "local" and "global" lakes and wetlands. In WaterGAP, "local" SW bodies are only recharged by runoff produced within the grid cell, while "global" SW bodies also obtain inflow from the upstream cell. In an uncoupled model, it is difficult to prescribe the area of lakes and wetlands that affect the flow exchange between SW body and GW. Maps generally show the maximum spatial extent

15 of SW bodies. This maximum extent is seldom reached, in particular in case of wetlands in dry areas. For global wetlands (wetlands greater than one 5' cell), it is therefore assumed in this model version that only eighty percent of their maximum extent is reached. A further difficulty in an uncoupled model run is that the water table elevation of SW bodies does not react to the GW-SW exchange flows Q_{swb} and that water supply from SW is not limited by availability. A loosing river may in reality dry out due to loss to GW and therefore cease to lose any more water. In case of rivers, B_{swb} is equal to $h_{riv} - 0.349 \times$

20 $Q_{bankfull}^{0.341}$ (Allen et al., 1994), where $Q_{bankfull}$ is the bankfull river discharge in the 5' grid cell (Verzano et al., 2012). Globally constant but different values are used for B_{swb} in case of wetlands, local lakes and global lakes (Table 1).

For the steady-state model, river elevation h_{riv} is set in each grid cell to the same elevation as all other SW bodies, h_{swb}. We found that for both gaining and loosing conditions, Q_{swb} and thus computed hydraulic heads are highly sensitive to h_{swb}. The overall best agreement with the hydraulic head observations of Fan et al. (2013) was achieved if h_{swb} (Eq. 5, 6 and

25 7) was set to the 30th percentile (P_{30}) of the 30" land surface elevation values of Fan et al. (2013) per 5' cell, e.g. the 30" elevation that is exceeded by 70% of the thousand 30" elevation values within one 5' cell. To decrease convergence time we used h_{eq} derived from the high-resolution steady-state hydraulic head distribution of Fan et al. (2013) as initial guess.

2.4 Model implementation

G³M is implemented using a newly developed open-source model framework G³M-f (Reinecke, 2018). Written in C++, the

30 framework allows the implementation of global and regional groundwater models alike while providing an extensible purely object-oriented model environment. It is primarily targeted as extension to WGHM but allows an in-memory coupling to any GHM or a standalone groundwater model. It provides a unit-tested environment offering different modules that can couple results in-memory to a different model or write out data flows to different file formats. As internal numerical library it uses Eigen3 (eigen.tuxfamily.org). Different from Vergnes et al. (2014), G³M's computations are not based on spherical coordinates

35 directly but on an irregular grid of rectangular cells. Cell sizes are provided by WaterGAP3 and are derived from their spherical coordinates maintaining their correct area and centre location. The model code will be adapted in the future to account for the different length in x and y direction per cell correctly.

Eq. (2) is reformulated as finite-difference equation and solved using a conjugate gradient approach and an Incomplete LUT preconditioner (Saad, 1994). In order to keep the memory footprint low, the conjugate gradient method makes use of the

40 sparse matrix. Furthermore, it solves the equations in parallel (preconditioner currently non-parallel). G³M can compute the presented steady-state solution (with the right-hand side of Eq. (2) being zero and the heads of Fan et al. (2013) as initial guess,

Table 1) on a commodity computer with four computational cores and a standard SSD in about 30 minutes while occupying 6 GB of RAM.

Similar to MODFLOW, G³M-f solves Eq. (2) in two nested loops: (1) the outer iteration checks the head and residual convergence criterion and adjusts whether external flows have changed into a different state e.g. from gaining to losing conditions and optimizes the matrix if flows are no longer head dependant. (2) The inner loop primarily consists of the conjugate gradient solver, which runs for a number of iterations defined by the user or until the residual convergence criterion is reached (Table 1), solving the current matrix equation.

Because the switch between Eq. (6) and Eq. (7) that occurs if e.g. h_{aq} drops below h_{riv} from one iteration to the next causes an abrupt change of c_{riv} inducing a nonlinearity that affects convergence we introduced an $\epsilon = 1$ m interval around h_{riv} and interpolate c_{riv} by a cubic hermite spline polynomial over that interval. This allows for a smoother transition between both states, reducing the changes in the solution if a river is in a gaining condition in one iteration and in a losing condition in the next and vice versa.

3 Results

3.1 Global hydraulic head and depth to GW distribution under natural steady-state conditions

As expected, the computed global distribution of steady-state hydraulic head under natural conditions (Fig. S5) follows largely the land surface elevation (Fig. S2), albeit with a lower range and locally different ratios between the hydraulic head and land surface gradients (Fig. S7). Deep GW, i.e. a large depth to GW, occurs mainly in mountainous regions (Fig. 10). Due to the steep topography (Fig. S8), large hydraulic gradients persist there that easily drain GW recharge even if the GW table drops. Deep GW also occurs in some flat desert area with very low GW recharge. Shallow GW occurs at the ocean boundary and in humid regions like the Amazon and in low elevation like the Caspian Sea and the Netherlands. In 2.1 % of all cells, GW head is simulated to be above the land surface elevation, by more than 1 m in 0.3 % and by more than 100 m in 0.004 % of the cells. The resulting depth to GW clearly resembles the difference between surface elevation and P_{30}, the assumed water level in the SW bodies h_{swb}, shown in Fig. S1, which indicates that simulated depth to GW is strongly governed by the SW body elevation.

The shallow water table in large parts of the Sahara is caused by losing rivers (and some wetlands) that cannot run dry in the model, causing an overestimation of the GW table (section 2.3). Please note that the computed steady-state depth to GW certainly underestimates the current depth to GW in GW depletion areas such as the High Plains Aquifer and the Central Valley in the USA (section 3.4), Northwestern India, North China Plain and parts of Saudi Arabia and Iran (Döll et al., 2014).

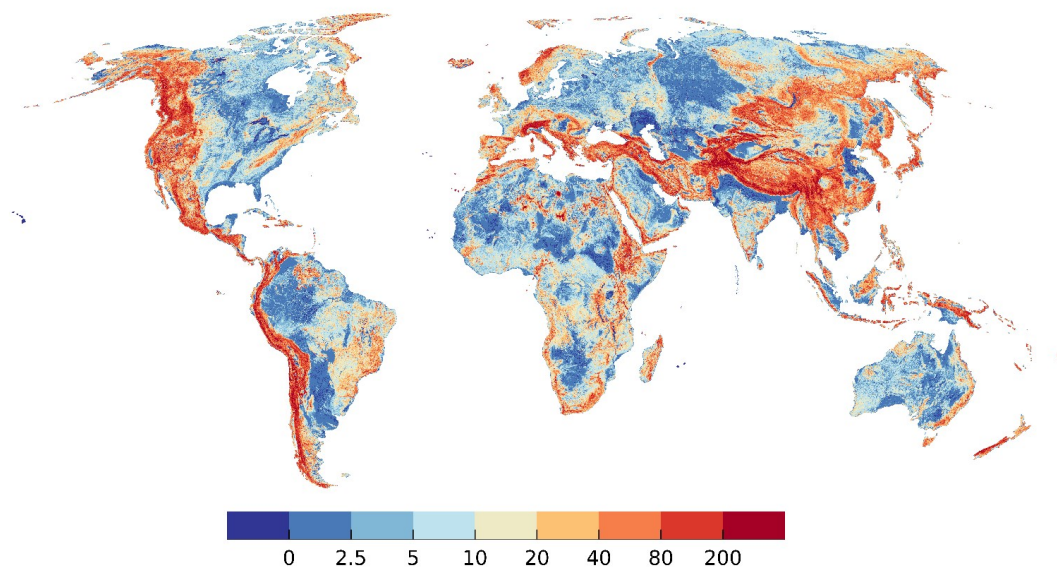



Figure 3 Simulated equilibrium depth to GW [m]. Maximum value 2070 m, minimum value -414 m (Extremes included in dark blue and dark red)**.**

3.2 Global water budget

5   Inflows to and outflows from GW of all G³M grid cells were aggregated according to the compartments ocean, river, lake, wetland, and diffuse GW recharge from soil (Fig. 4). The difference between the global sum of inflows and outflows in less than 10^{-6} %. This small volume balance error indicates the correctness of the numerical solution.

Total diffuse GW recharge from soil is $3.9\ 10^{10}\ m^3\ day^{-1}$ and approximately equal to the drainage of GW to rivers. Rivers are the ubiquitous drainage component of the model), followed by wetlands, lakes and the ocean boundary. According

10 to G³M, the amount of river water that recharges GW is only about a 40th of the drainage to GW, and the relative inflow to GW from lakes and wetlands is even smaller (Fig. 4). Possibly, flow from SW to GW is even overestimated, as outflow from SW is not limited by water availability in the SW, and depending on the hydraulic conductivity, Eqs. (5) and (6) can lead to rather large flows. Inflow from the ocean, which is more than two magnitudes smaller than outflow to ocean, occurs in regions where $h_{swb} = P_{30}$ is below h_{ocean}.

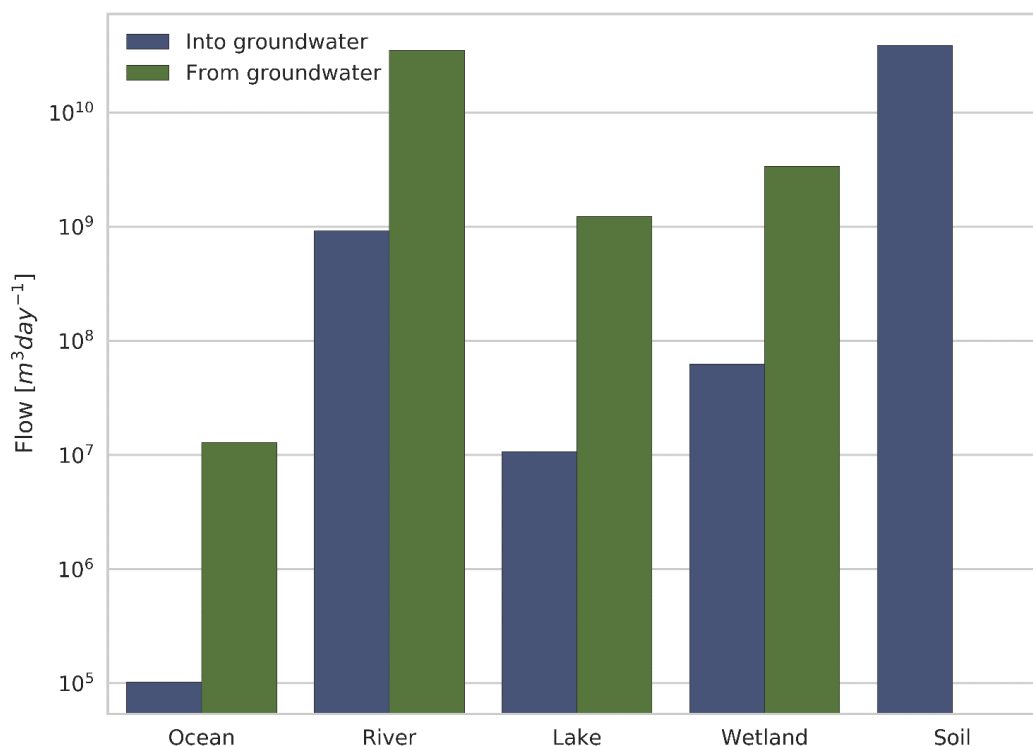



Figure 4 Global sums of flows from different compartments into or from GW at steady state. Flows into the GW are denoted by the color blue, flows out of the GW into the different compartments by green. The compartment soil is the diffuse GW recharge from soil calculated by WaterGAP.

3.3 GW-SW interactions

Figure 5 plots the spatial distribution of simulated flows from and to lakes and wetlands (Fig. 5a) as well as from and to rivers (Fig. 5b). It reveals strong interaction between GW and SW bodies that is dominated by GW discharging to SW bodies. Parallel to the overall budget (Fig. 4), the map reveals the globally large but locally strongly varying influence of lakes and wetlands

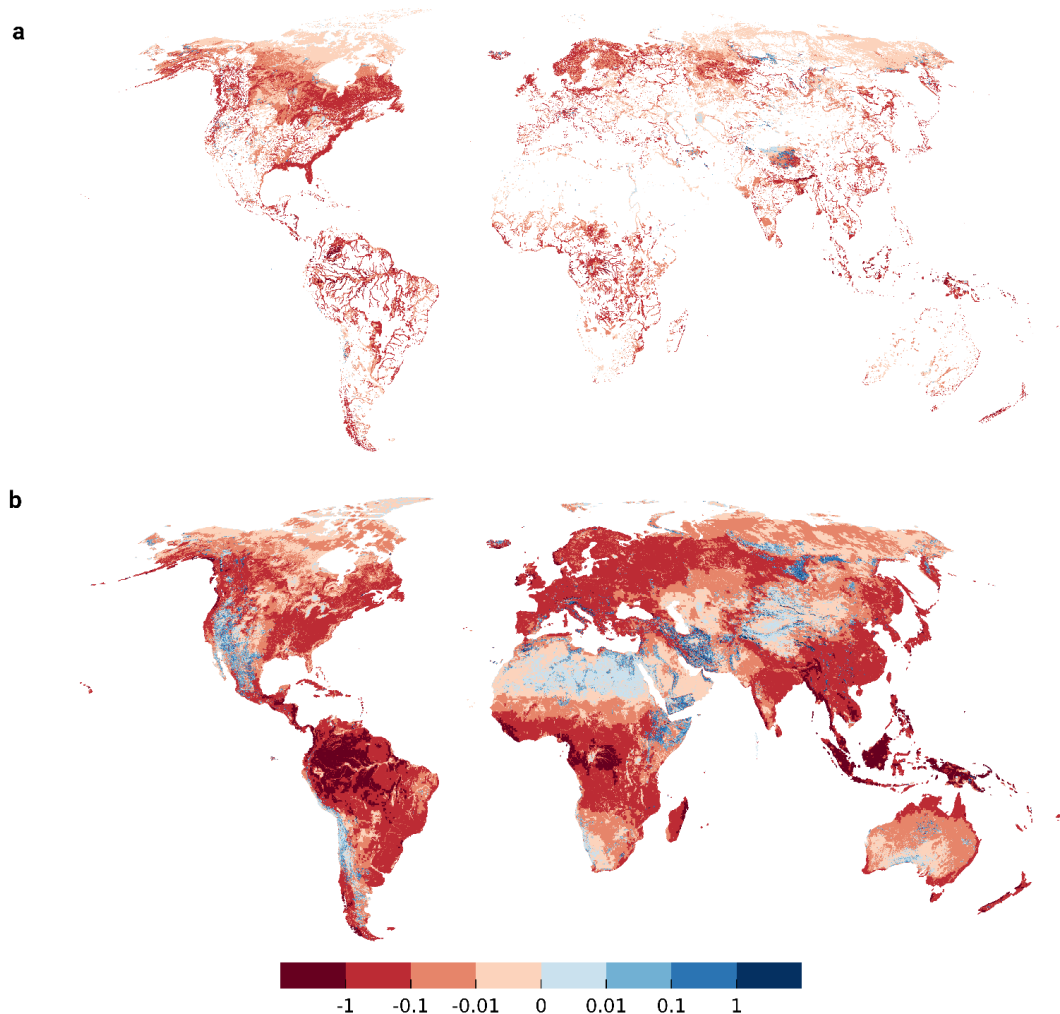
10 (Fig. 5a). Rivers with riparian wetlands such as the Amazon River receive comparably small amounts of GW as most of the GW is drained by the wetland (compare Figs. 5a and 5b). Similarly, areas dominated by wetlands and lakes (e.g. parts of Canada and Scandinavia) show less inflow for rivers (Fig. 5b). 93 % of all grid cells contain gaining rivers, and only 7% loosing rivers. Gaining lakes and wetlands are found in 12 % and 11 % of the cells, respectively, whereas only 0.2 % contain a losing lake or wetland. In G3M, all SW bodies (rivers, lakes and wetlands) in a grid cell either loose or gain water.

15   Gaining rivers, lakes and wetlands with very high absolute Q_{swb} values over 1 mm day^{-1} (averaged over the grid cell area of approximately 80 km^2) can be found in the Amazon, Congo, Bangladesh, and Indonesia, where GW recharge in very high (Fig. S4). Values below 0.01 mm day^{-1} are present in dry and in permafrost areas where groundwater recharge is small.

      Losing SW bodies occur in the model under two conditions, in mountainous regions where depth to GW is high and

20 in arid and semi-arid climates with low diffuse groundwater recharge. Without focused GW recharge, the GW table would drop to even further in the mountains and is necessary to counteract the large hydraulic gradients caused by the large topographic gradients. Rivers lose more than 1 mm day^{-1} in Ethiopia and Somalia, West Asia, Northern Russia, the Rocky

Mountains and the Andes whereas lower values can be observed in Australia and in the Sahara. High values of outflow from wetlands and lakes are found in Tibet, the Andes and northern Russia, lower values in the Sahara and Kazakhstan. The river Nile in the Northern Sudan and Egypt is correctly simulated to be a losing river (Fig. 5b), being an allogenic river that is mainly sourced from the upstream humid areas, including the man-made Lake Nasser (Elsawwaf et al., 2014) (Fig. 5a). Furthermore,

5   the following lakes and riparian wetlands are simulated to recharge GW: parts of the Congo River, Lake Victoria, the Ijsselmeer, Lake Ladoga, the Aral Sea, parts of the Mekong Delta, the Great Lakes of North America. On the other hand, no losing stretches are visible at the Niger River and its wetlands and almost none Northeastern Brazil even though that losing conditions are known to occur there (FAO, 1997; Costa et al., 2013).

10   **Figure 5** Flow Q_{swb} [$mm\ day^{-1}$] from/to wetlands, lakes (a) and losing/gaining streams (b) with respect to the 5´grid cell area. Gaining rivers are shown in red, rivers recharging the aquifer in blue. Focused recharge occurs in arid regions, e.g. alongside the river Nile, and in mountainous regions where the average water table is well below the land surface elevation.

Simulated flows between GW and SW depend on assumed conductances for both rivers and lakes/wetlands (Eqs. 5, 6, 7) shown in Fig. 6. Q_{swb} (Fig. 5) correlates positively with conductance. Conductance for gaining rivers correlates positively

15   with GW recharge (Eq. (7) and Fig. S4). High conductance values are reached in the tropical zone due to a higher GW recharge but are capped at a plausible maximum value of $10^7\ m^2 day^{-1}$s in case of river (section 2.1) (Fig. 6b), while lakes and wetlands, with a larger area, can reach larger values, e.g. in Canada or Florida.

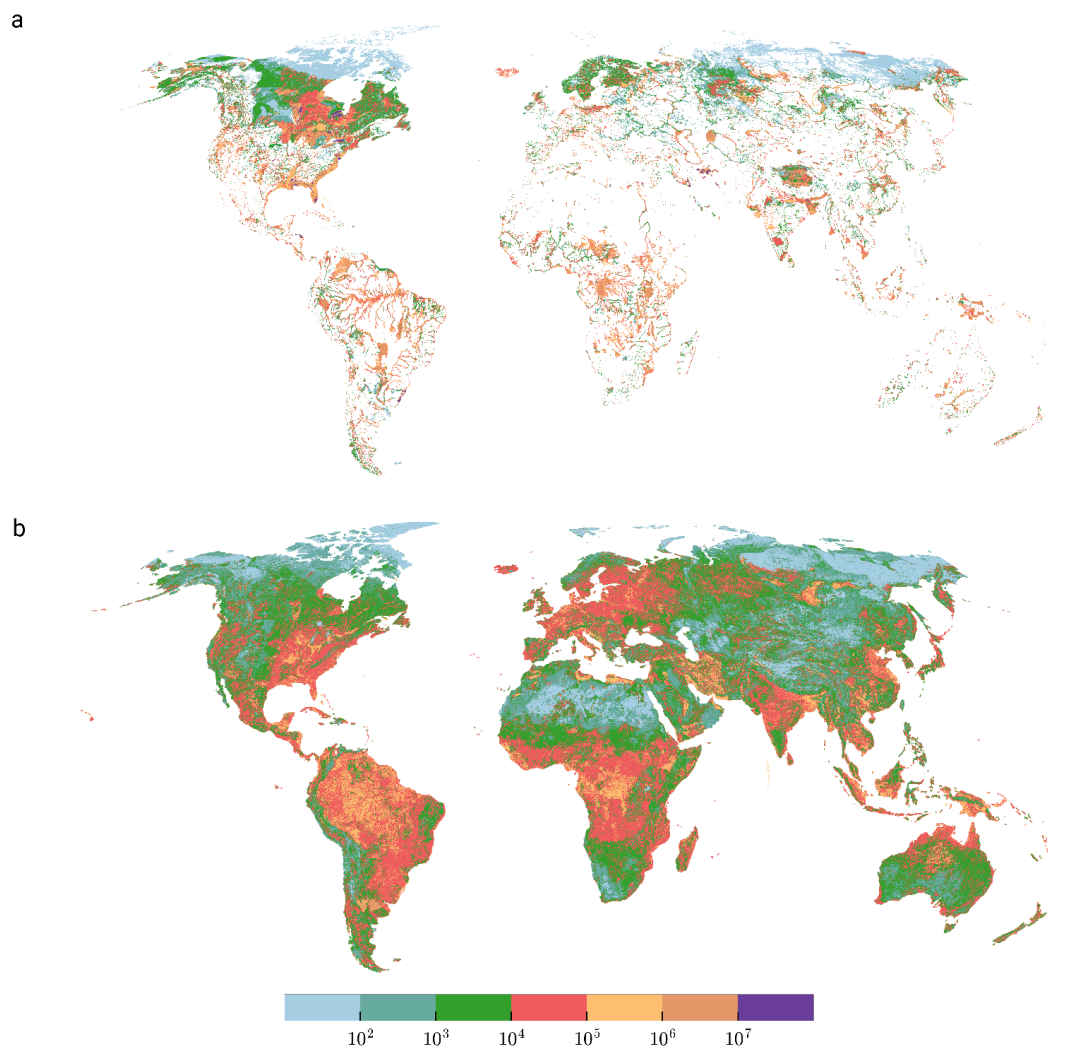



Figure 6 Conductance [m^2day^{-1}] of lakes and wetlands (a) and rivers (b). In regions close to the pole conductance is in general lower due to the influence of the low aquifer conductivity (losing conditions), and relatively small GW recharge due to permafrost conditions (only applies for gaining conditions). Max conductance of wetlands is 10^8.

3.4 Lateral flows

Figure 7 shows lateral outflow from both model layers in percent of the sum of diffuse GW recharge from soil and GW recharge from SW bodies. The percentage of recharge that is transported through lateral flow to neighbouring cells depends on 5 main factors: (1) hydraulic conductivity (Fig. S3), (2) diffuse GW recharge (Fig. S4), (3) losing or gaining SW bodies (Fig. 5), (4) their conductance (Fig. 6) and (5) the head gradients (Fig. S5).

On large areas of the globe, where GW discharges to SW, the lateral flow percentage is less than 0.5% of the total GW recharge as GW recharge in a grid cells is simulated to leave the grid cell by discharge to SW. For example, in the permafrost regions, very low hydraulic conductivity limits the outflow to neighbouring cells of the occurring recharge, leading to these very low percent values. Such values also occur in regions with high SW conductances and rather low hydraulic conductivity, e.g. in the Amazon Basin. Values of more than 5% occur where hydraulic conductivity is high even if the terrain

in rather flat, such as in Denmark. Higher values may occur for in case of gaining rivers in dry areas like Australia or in mountainous regions where large hydraulic gradients can develop. In mountains with gaining surface water bodies, lateral outflows may even exceed GW recharge of the cell. In grid cells where SW bodies recharge the GW, outflow tends to be a large percentage of total GW recharge as there is no outflow from GW other than in lateral direction, and values often exceed

5   100% (Fig. 7).

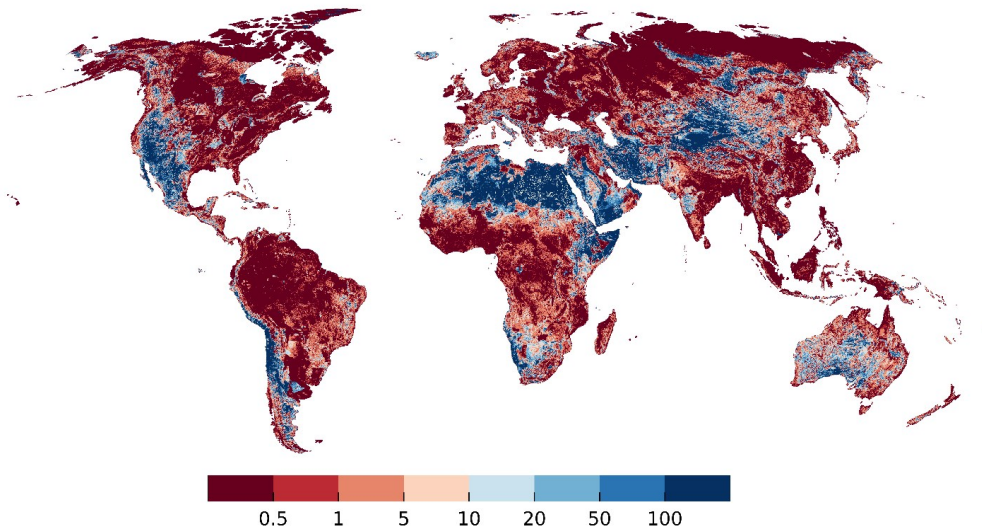



Figure 7 Percentage of GW recharge from soil and surface water inflow that is transferred to neighboring cells through lateral out flow (sum of both layers). Grid cells with zero total GW recharge are shown in white (a few cells in the Sahara and the Andes).

3.5 Comparison to groundwater well observations

10   Global observations of depth to GW were assembled by Fan et al. (2007, 2013). We selected only observations with known land surface elevation and removed observations where a comparison to local studies suggested a unit conversion error. This left total of 1,070,402 depth to GW observations. An "observed head" per 5' model cell was then calculated by first computing hydraulic head of each observation by subtracting depth to GW from the 5' land surface elevation used in G³M and then calculating the arithmetic mean of all observations within the 5' model cell. Multiple obstacles limit the comparability of

15   observations to simulated values. (1) Observations were recorded at a certain moment in time influenced by seasonal effects and abstraction from GW, whereas the simulated heads represent a natural steady-state condition. (2) Observation locations are biased towards river valleys and productive aquifers. (3) Observations may be located in valleys with shallow local water tables too small to be captured by a coarse resolution of 5'.

Simulated steady-state hydraulic heads in the upper model layer are compared to observations in Fig. 8. Shallow GW

20   is generally better represented by the model than deeper GW. Especially the water table in mountainous areas is underestimated, which may be related to observations in perched aquifers caused by low permeability layers (Fan et al. 2013) that are not represented in G³M due to lacking information. Because the steady-state model cannot take into account the impact of GW abstraction, the computed depth to GW values are considerably smaller than currently observed values in GW depletion areas like the Central Valley in California (where once wetlands existed before excessive GW use depleted the aquifer) and

25   the High Plains Aquifer in the Midwest of the USA. Still, the elevation of the GW table in the non-depleted Rhine valley in Germany is overestimated, too. Figure 9a shows the hydraulic head comparison as scatter plot. Overall, the simulation results tend to underestimate observed hydraulic head but much less than the steady-state model presented by de Graaf et al. (2015).

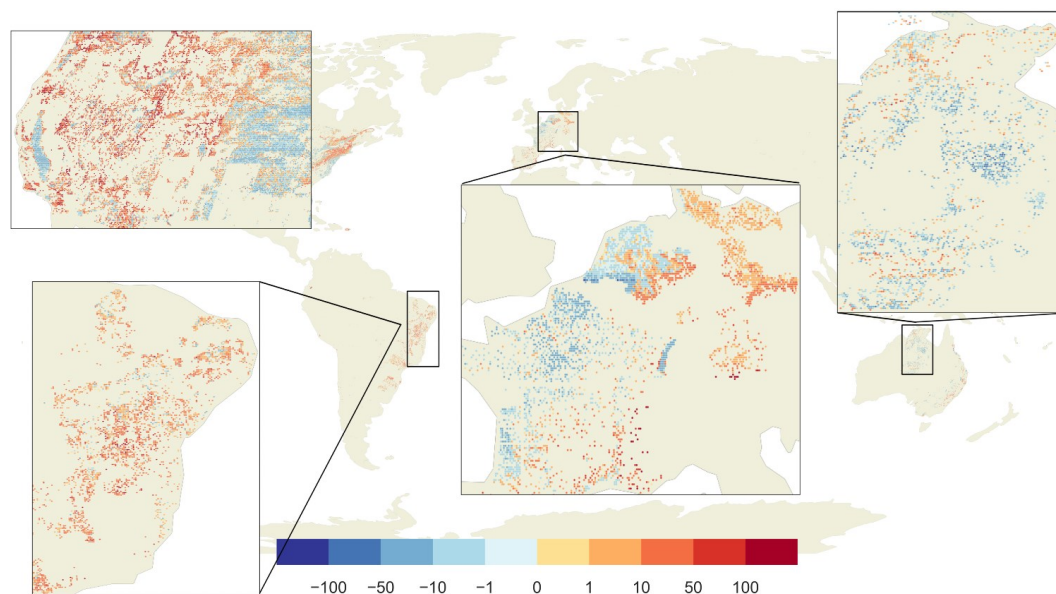



Figure 8 Differences between observed and simulated hydraulic head [*m*]. Red dots show areas where the model simulated deeper GW as observed, blue shallower GW. In the grey areas, no observations are available.

To compare performance of G³M to the steady-state results of two high-resolution model of Fan et al. (2013) and
5   ParFlow (Maxwell et al., 2015), heads in 30" (Fan et al., 2013) and 1 km (ParFlow) grid cells were averaged to the G³M 5'
grid cells. The comparison of 5' observations to the 5' average of ParFlow seem to be consistent with the 1 km model
comparison in Maxwell et al., 2015, their Fig. 5, even though over/under -estimates in the original resolution seemed to be
smoothed out by averaging to 5' (not shown). The heads of Fan et al. (2013) fit better to observations than G³M heads, with
less underestimation (Fig. 9b). The comparison of G³M heads to Fan et al. (2013) values for all 5' grid cells, which are also
10   the initial heads of G³M and the basis to compute river conductances, show that heads computed with the G³M are mostly
much lower except in regions with a shallow GW (Fig. 9c). This cannot be attributed to the 100 times lower spatial resolution
per se but to the selection of the 30th percentile of the 30" as the SW drainage level. Outliers in the upper half of the scatter
plot, with much larger heads than the initial values, are mainly occurring in steep mountain areas like the Himalayas where the
5' model is not representing smaller valleys with a lower head.

15   For the continental US, the computationally expensive 1-km integrated hydrological model ParFlow (Maxwell et al.,
2015), fits much better to observations than G³M (Figs. 9d, e). G³M produces a generally lower water table (Fig. 9f), a main
reason being that ParFlow assumes an impermeable bedrock at a depth of more than 100 m below the land surface elevation.
Plotting hydraulic head instead of depth to GW has the disadvantage that the goodness of fit is dominated by the topography
as the observed heads are calculated based on the surface elevation of the model. Even though hydraulic heads are a direct
20   result of the model and are forcing lateral GW flows, depth to GW is more relevant for processes like capillary rise. For G³M,
there is almost no correlation between depth to GW observations and simulated values. To our knowledge, no publication on
large-scale GW modeling presents correlations of simulated with observed depth to GW.

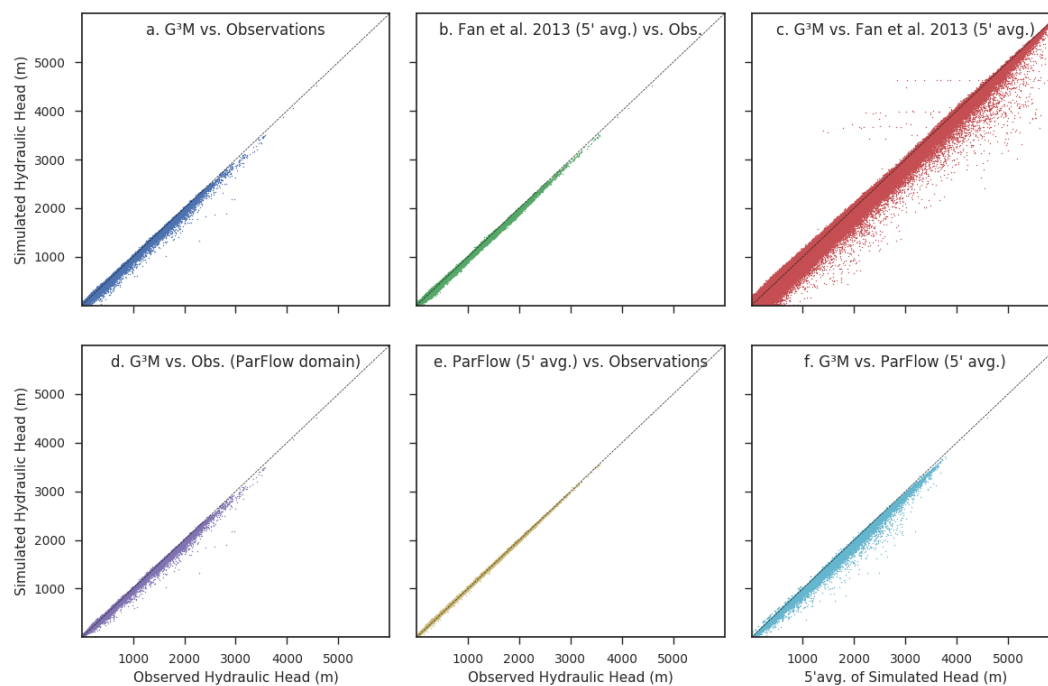



Figure 9 Scatterplots of simulated vs. observed hydraulic head and inter-model comparison of heads. (Upper panel) The steady-state run of G³M vs. observations (a), the 5' average of the equilibrium head of Fan et al. (2013) vs. observations (b) and the avg. equilibrium vs. G³M (c). (Lower panel) The steady-state run of G³M vs. observations only for the ParFlow domain (d), the 5' average of the ParFlow average annual GW table (Maxwell et al., 2015) vs. observations (e) and the steady-state run of G³M vs. 5' average of the ParFlow average annual GW table (f).

3.6 Case study Central Valley

To evaluate G³M further, its results were analysed for to a well-studied region, the Central Valley in California, USA. The Central Valley is one of the most productive agricultural regions of the world and heavily relies on GW pumpage to meet irrigation demands (Faunt et al., 2016). GW pumping in the valley increased rapidly in the 1960s (Faunt, 2009). Figure 10a shows simulated depth to GW for the Central Valley, the coast and the neighboring Sierra Nevada mountainside as well as parts of the Great Basin. The depth to GW table represents natural conditions without any pumping and is rather small. It roughly resembles the depth to GW assumed in the Central Valley Hydrological Model (CVHM) as initial condition, representing a natural state (Faunt, 2009) (Fig. 10b). Towards the Sierra Nevada east of the valley, depth to GW increases in both G³M and CVHM. Beyond the CVHM domain, depth to GW in mountainous regions is probably overestimated by G³M. The elevation of neighboring cells may differ up to a 1000 meter resulting in a large gradient (Fig. S6b and S6e).

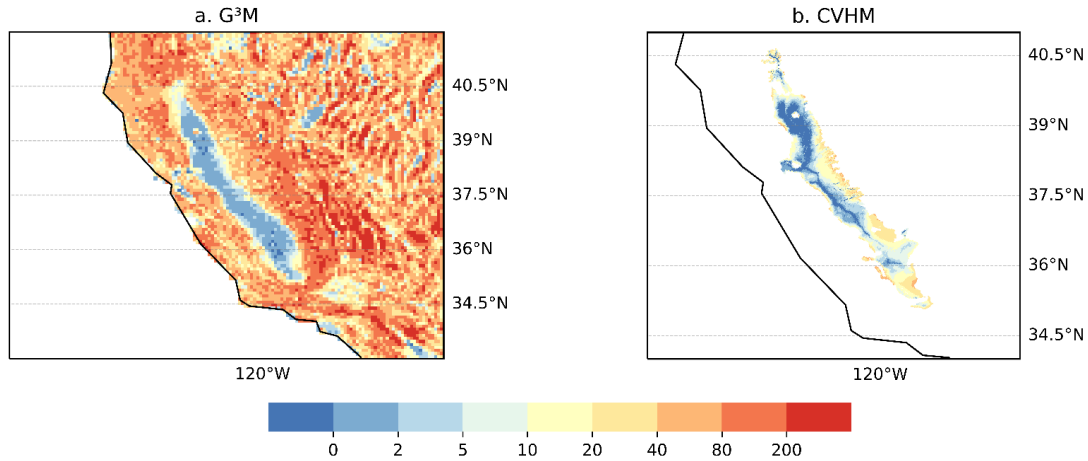



Figure 10 Plots of depth to GW [**m**] as calculated by G³M for the Central Valley and the Great Basin (a), and as used by CVHM as the natural state and starting condition (Faunt, 2009) (b).

4 Discussion

5   The main aim of global gradient-based groundwater flow modelling with G³M is to better simulate water exchange between SW and GW, for example for an improved estimation of GW resources in dry regions of the globe that are augmented by focused recharge from SW bodies. A major challenge for simulating GW-SW interactions (but also capillary rise) at the global scale is the large size of grid cells that is required due to computational constraints. Within the 5' grid cells, land surface elevation at the scale of 30" very often varies by more than 20 m, and often by 200 m and more (Fig. S1), while the vertical

10   position of the cell and the hydraulic head are approximated in the model by just one value. The question is whether head-dependent flows between grid cells, between GW and SW and from GW to soil (capillary rise) can be simulated successfully at the global scale, i.e. whether an improved quantification of these flows as compared to the simple linear reservoir model currently used in most GHMs can be achieved by this approach. This question cannot be answered yet as we have not yet achieved a dynamic coupling of G³M with a global hydrological model but one may speculate that some innovative approach

15   to take into account the elevation variations within the grid cells may be needed.

The presented development of the uncoupled steady-state global GW flow model enabled us to better understand how the spatial hydraulic head pattern relates to the fundamental drivers topography, climate and geology (Fan et al., 2007) and how the interaction to SW bodies govern the global head distribution. Simulated depth to groundwater is particularly affected by the assumed hydraulic head in SW bodies, the major GW drainage component in the model. As rivers represent a natural

20   occurring drainage at the lowest point in a given topography, one would assume that the minimum elevation 30" land surface elevation per 5' grid cell is a reasonable choice. Experiments have shown that this will induce a head distribution well below the average 5' elevation that is much below observations of Fan et al. (2013). We also tested setting h_{swb} to the average elevation of all "blue" cells (with a depth to GW of less than 0.25 m) of the steady-state 30" water table results of Fan et al. (2013) that indicate the locations were GW discharges to the surface. This leads to an overall underestimation of the observed

25   hydraulic heads (Fig. S9). Furthermore, it leads to an increase in losing SW bodies (comp. Fig. S11 with Fig. 4). However, it is difficult to judge whether this improves the simulation. More stretches of the Nile and its adjacent wetlands and also of the Niger wetlands and rivers in Northeastern Brazil are losing in case of lower h_{swb}, which appears to reasonable. Additionally, choosing the average as SW elevation provides on the on hand a better fit to observations (Fig. S9) but leads to a wold wide flooding with largely overestimated heads (Fig. S10) and a much longer convergence time due to an increased oscillation

30   between gaining and losing conditions.

The problem is very likely one of scale. This is supported by the fact that both high-resolution models, even the simple one of Fan et al. (2013) fit better to observations than the low-resolution model G³M (Fig. 9). In case of high resolution (e.g. 30"), there are a number of grid cells at an elevation above the average 5' land surface elevation, leading to higher hydraulic heads in parts of the 5' area that drain towards the SW body in a lower 30" grid cell. In case of the low spatial resolution of 5'

5   in which h_{swb} is set to the elevation of the fine-resolution drainage cell, the 5' hydraulic head is rather close to this (low) elevation (Fig. S12), resulting in an underestimation of hydraulic head and thus an overestimation of depth to GW. While it is plausible and necessary to assume that there is SW-GW interaction within each of in the approximately 80 km², this is not the case for the two orders of magnitude smaller 30" grid cells. Thus, with the high resolution, heads are not strongly controlled everywhere by the head in SW bodies. Selecting the 30th percentile of the 30" land surface elevation as h_{swb} was found, by

10   trial-and-error, to lead to a hydraulic head distribution that fits reasonably well to observed head. It avoids that the simulated GW table drops to low while avoiding the excessive flooding that occurs if h_{swb} is set to the average of 30" land surface elevations, i.e. the 5' land surface elevation (Fig. S9).

The constraint that the selected h_{swb} value puts on simulated hydraulic heads is also linked to the conductance of the SW bodies. A higher conductance will lead to aquifer heads closer to h_{swb}. If the hydraulic head drops below the bottom level

15   of the SW body, the hydraulic gradient is assumed to become 1 and the SW body recharges the GW with a rate of K_{aq} per unit SW body area. In case of a K_{aq} value of 10^{-5} m s⁻¹, the SW body would lose approximately 1 m of water each day. It is to be investigated how the sensitivity to choice of SW body elevation and conductance leads to a solution that fits observations best. A lower conductance may lead to a higher groundwater table as SW bodies don't drain as much water; on the other hand, they seem to provide an important recharge mechanism in the steady-state model for some regions preventing an even higher depth

20   to GW. The simple conductance approach applied in G³M could possibly be approved by the approach proposed by Morel-Seytoux et al. (2017).

De Graaf et al. (2015) set their SW head (h_{swb}) to the land surface elevation of the 6' grid cells minus river depth at bankfull conditions plus water depth at average river discharge. Together with the missing interaction between lakes and wetlands and a different approach to river conductance, this might be a reason for the additional drainage above the floodplain

25   that was necessary to avoid excessive flooding. On the other hand, this adaption allows the drainage of water even if the hydraulic head is below the SW elevation that might have led to the global underestimation of hydraulic heads. Thus, the difference in model heads seems to be closely related to the sensitivity of SW body elevation.

As described above, G³M differs from regional groundwater models due to grid cell size in that it is more conceptual and cannot capture actual variability of topography, aquifer depth (Richey et al., 2015) and (vertical) heterogeneity of

30   subsurface properties. Because of the relatively thin model layers and the assumed confined conditions, groundwater flows may occur even if the cell itself should be dry as the hydraulic head is below the model domain. Prohibiting any lateral flow from these cells would yield large mountainous areas as non-active cells, not only defeating the purpose of the model but also inducing large numerical problems. The lack of information about the three-dimensional distribution of hydraulic conductivity is expected to negatively impact the quality of simulated GW flow. For example, the lateral conductivity and connectivity of

35   groundwater along thousands of kms from e.g. the Rocky Mountains in the Central USA to the coast as well as the vertical connectivity is likely to be overestimated by G³M, as vertical faults and interspersed aquitards are not represented; this leads to an underestimation of hydraulic head in those mountainous areas.

5 Conclusions

We have presented the concept and first results of a new global gradient-based GW flow model that is to be coupled to the

40   GHM WaterGAP. The uncoupled steady-state model has provided important insights into challenges of global GW flow modelling mainly related to the necessarily large grid cells size (5' by 5') as well as first global maps of SW-GW interactions.

Simulated heads were found to be strongly impacted by assumption regarding the interaction with SW bodies, in particular the selected elevation of the SW table and the prescribed conductance. We have demonstrated that simulated G³M hydraulic heads fit better to observed heads than the heads of the comparable steady-state GW model of de Graaf et al. (2015), without requiring additional drainage that would prevent a full coupling to a GHM.

5   The presented results are the first step towards a fully coupled model in which SW heads are computed as a function of surface water hydrology and GW abstractions can be taken into account. Especially the interaction with SW bodies that can run dry will make the model behavior more realistic. The fully coupled model will simulate transient behaviour reflecting climate variability and change. Simulated hydraulic head dynamics will be compared to observed head time series as well as to the output of large-scale regional models, while total water storage variations will be compared to GRACE satellite data.

10   However, it will be challenging to judge the quality of simulated GW-SW interactions due to a scarcity of observations.

6 Code and data availability

The model-framework code is published in Reinecke (2018) and available at globalgroundwatermodel.org with a description on how to compile and run a basic GW model. The code is available under the GNU General Public License 3. The input data and simulations presented here may be obtained by contacting the lead author via email.

15   **Acknowledgments**

Part of this study was funded by a Friedrich-Ebert foundation Ph.D. fellowship. We are very grateful to Ying Fan and Gonzalo Miguez-Macho for fruitful discussions and data provisioning.

References

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., and Siebert, S.: Development and testing of the WaterGAP 2 global
20   model of water use and availability, Hydrological Sciences Journal, 48, 317–337, doi:10.1623/hysj.48.3.317.45290, 2003.

Allen, P. M., Arnold, J. C., and Byars, B. W.: Downstream channel geometry for use in planning-level models, J Am Water Resources Assoc, 30, 663–671, doi:10.1111/j.1752-1688.1994.tb03321.x, 1994.

Belcher, W. R. and Sweetkind, D. S.: Death Valley regional groundwater flow system, Nevada and California-Hydrogeologic framework and transient groundwater flow model, U.S. Geological Survey Professional Paper, 398, 2010.

25   Costa, A. C., Foerster, S., de Araújo, J. C., and Bronstert, A.: Analysis of channel transmission losses in a dryland river reach in north-eastern Brazil using streamflow series, groundwater level series and multi-temporal satellite data, Hydrol. Process., 27, 1046–1060, doi:10.1002/hyp.9243, 2013.

de Graaf, I. E.M. de, Sutanudjaja, E. H., van Beek, L. P. H., and Bierkens, M. F. P.: A high-resolution global-scale groundwater model, Hydrol. Earth Syst. Sci., 19, 823–837, doi:10.5194/hess-19-823-2015, 2015.

30   de Graaf, I. E.M. de, van Beek, R. L.P.H., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E. H., and Bierkens, M. F.P.: A global-scale two-layer transient groundwater model: Development and application to groundwater depletion, Advances in Water Resources, 102, 53–67, doi:10.1016/j.advwatres.2017.01.011, 2017.

Dogrul, E., Brush, C., and Kadir, T.: Groundwater Modeling in Support of Water Resources Management and Planning under Complex Climate, Regulatory, and Economic Stresses, Water, 8, 592, doi:10.3390/w8120592, 2016.

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

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

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

Döll, P., Müller Schmied, H., Schuh, C., Portmann, F. T., and Eicker, A.: 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, 2014.

Eisner, S.: Comprehensive evaluation of the WaterGAP3 model across climatic, physiographic, and anthropogenic gradients, University of Kassel, Kassel, Germany, 2016.

Elsawwaf, M., Feyen, J., Batelaan, O., and Bakr, M.: Groundwater-surface water interaction in Lake Nasser, Southern Egypt, Hydrol. Process., 28, 414–430, doi:10.1002/hyp.9563, 2014.

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

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

FAO: Irrigation potential in Africa: http://www.fao.org/docrep/w4347e/w4347e0i.htm, 1997, last access: 1 May 2018.

Faunt, C. C. (Ed.): Groundwater Availability of the Central Valley Aquifer, California, U.S. Geological Survey Professional Paper, 1776, 225 pp., 2009.

Faunt, C. C., Sneed, M., Traum, J., and Brandt, J. T.: Water availability and land subsidence in the Central Valley, California, USA, Hydrogeol J, 24, 675–684, doi:10.1007/s10040-015-1339-x, 2016.

Gleeson, T., Befus, K. M., Jasechko, S., Luijendijk, E., and Cardenas, M. B.: The global volume and distribution of modern groundwater, Nature Geosci, 9, 161–167, doi:10.1038/ngeo2590, 2016.

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

Gleeson, T., Wada, Y., Bierkens, M. F. P., and van Beek, L. P. H.: Water balance of global aquifers revealed by groundwater footprint, Nature, 488, 197–200, doi:10.1038/nature11295, 2012.

Gossel, W., Ebraheem, A. M., and Wycisk, P.: A very large scale GIS-based groundwater flow model for the Nubian sandstone aquifer in Eastern Sahara (Egypt, northern Sudan and eastern Libya), Hydrogeology Journal, 12, 698–713, doi:10.1007/s10040-004-0379-4, 2004.

Harbaugh, A. W.: MODFLOW-2005, the US Geological Survey modular ground-water model: the ground-water flow process, US Department of the Interior, US Geological Survey Reston, 2005.

Hartmann, J. and Moosdorf, N.: The new global lithological map database GLiM: A representation of rock properties at the Earth surface, Geochemistry, Geophysics, Geosystems, 13, doi:10.1029/2012GC004370, 2012.

Konikow, L. F.: Contribution of global groundwater depletion since 1900 to sea-level rise, Geophys. Res. Lett, 38, doi:10.1029/2011GL048604, 2011.

Lange, S.: EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI), 2016.

Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, Journal of Hydrology, 296, 1–22, doi:10.1016/j.jhydrol.2004.03.028, 2004.

Maxwell, R. M., Condon, L. E., and Kollet, S. J.: 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, 2015.

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

Morel-Seytoux, H. J., Miller, C. D., Miracapillo, C., and Mehl, S.: River Seepage Conductance in Large-Scale Regional Studies, Ground Water, 55, 399–407, doi:10.1111/gwat.12491, 2017.

Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F. T., Flörke, M., and Döll, P.: 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, 2014.

Naff, Richard L., and Edward R. Banta: The US Geological Survey modular ground-water model-PCGN: a preconditioned conjugate gradient solver with improved nonlinear control, Geological Survey (US), 1331, 2008.

Reinecke, R.: G³M-f a global gradient-based groundwater modelling framwork, JOSS, 3, 548, doi:10.21105/joss.00548, 2018.

Richey, A. S., Thomas, B. F., Lo, M.-H., Famiglietti, J. S., Swenson, S., and Rodell, M.: Uncertainty in global groundwater storage estimates in a Total Groundwater Stress framework, Water Resour. Res, 51, 5198–5216, doi:10.1002/2015WR017351, 2015.

Saad, Y.: ILUT: A dual threshold incomplete LU factorization, Numer. Linear Algebra Appl., 1, 387–402, doi:10.1002/nla.1680010405, 1994.

Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., and McMahon, P. B.: Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley, PNAS, 109, 9320–9325, doi:10.1073/pnas.1200311109, 2012.

Stonestrom, D. A., Constantz, J., Ferre, T. P.A., and Leake, S. A.: Ground-water recharge in the arid and semiarid southwestern United States, U.S. Geological Survey Professional Paper 1703, 414 p, 2007.

Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F. P.: Large-scale groundwater modeling using global datasets: a test case for the Rhine-Meuse basin, Hydrol. Earth Syst. Sci., 15, 2913–2935, doi:10.5194/hess-15-2913-2011, 2011.

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., Gurdak, J. J., Allen, D. M., Shamsudduha, M., Hiscock, K., Yeh, P. J.-F., Holman, I., and Treidel, H.: Ground water and climate change, Nature Climate change, 3, 322–329, doi:10.1038/nclimate1744, 2012.

UNESCO-IHP, IGRAC, WWAP: GEF-TWAP Methodology Transboundary Aquifers.: http://isarm.org/sites/default/files/resources/files/TWAP%20Methodology%20Groundwater%20Component%20%28Revised%20Aug%202012%29.pdf, 2012, last access: 9 October 2017.

van Beek, L. P. H., Wada, Y., and Bierkens, M. F. P.: Global monthly water stress: 1. Water balance and water availability, Water Resour. Res., 47, doi:10.1029/2010WR009791, 2011.

Vergnes, J.-P., Decharme, B., Alkama, R., Martin, E., Habets, F., and Douville, H.: A Simple Groundwater Scheme for Hydrological and Climate Applications: Description and Offline Evaluation over France, J. Hydrometeor, 13, 1149–1171, doi:10.1175/JHM-D-11-0149.1, 2012.

Vergnes, J.-P., Decharme, B., and Habets, F.: Introduction of groundwater capillary rises using subgrid spatial variability of topography into the ISBA land surface model, J. Geophys. Res. Atmos., 119, 11, doi:10.1002/2014JD021573, 2014.

Verzano, K., Bärlund, I., Flörke, M., Lehner, B., Kynast, E., Voß, F., and Alcamo, J.: Modeling variable river flow velocity on continental scale: Current situation and climate change impacts in Europe, Journal of Hydrology, 424-425, 238–251, doi:10.1016/j.jhydrol.2012.01.005, 2012.

Wada, Y.: Modeling Groundwater Depletion at Regional and Global Scales: Present State and Future Prospects, Surv Geophys, 37, 419–451, doi:10.1007/s10712-015-9347-x, 2016.

Wada, Y., van Beek, L. P. H., and Bierkens, M. F. P.: Nonsustainable groundwater sustaining irrigation: A global assessment, Water Resour. Res., 48, doi:10.1029/2011WR010562, 2012.

Table 1 Model parameter values, input data sources and other information about the steady-state simulation.

Parameter	Symbol	Units	Description	Eq. No.
Landmask	-	-	Location and area of 2161074 cells at 5' resolution based on WaterGAP (Eisner, 2016))	-
GW recharge	R_g	$L^3 T^{-1}$	Mean annual diffuse GW recharge 1901–2013 of WaterGAP 2.2c (Müller Schmied et al., 2014) forced with EWEMBI (Lange, 2016), spatial resolution 0.5° (Fig. S4)	1,3,7
Hydraulic conductivity	K_{aq}	LT^{-1}	Derived from Gleeson et al., 2014 (Fig. S3)	2,5
Hydraulic head	$h_{(aq)}$	L	Head of the aquifer in a computational cell, initial estimate based on 5' average of 30" head of Fan et al. (2013)	2,4,7,8
Ocean boundary conductivity	c_{ocean}	$L^2 T^{-1}$	$10\ m^2\ day^{-1} = 0.1\ m\ day^{-1}\ 10\ km\ 10\ km^{-1}\ 100\ m$, with K of $10^{-6}\ m\ s^{-1}$ and a distance of 10 km from the cell center to the boundary with a cell thickness of 100 m	3,4
Ocean boundary head	h_{ocean}	L	Global mean sea-level of 0 m	4
SW head	h_{swb}	L	30 % quantile (P$_{30}$) of 30" land surface elevation of Fan et al. (2013) per 5' grid cell	5
SW bottom elevation	B_{swb}	L	2 m (wetlands), 10 m (local lakes), 100 m (global lakes) below P$_{30}$	5
Area of global and lcal lakes and global and local wetlands	WL	L^2	Per 5' grid cell, based on WaterGAP 3 (Eisner, 2016),	6
Length of the river	L	L	Per 5' grid cell, based on WaterGAP 3 (Eisner, 2016)	6
Width of the river	W	L	Per 5' grid cell, based d on WaterGAP 3 (Eisner, 2016)	6
River head	h_{riv}	L	h_{swb}	6,7,8
River bottom elevation	B_{riv}	L	$h_{riv} - 0.349 \times Q_{bankfull}^{0.341}$ (Allen et al., 1994)	6
Equilibrium hydraulic head	h_{eq}	L	Steady-state hydraulic head of Fan et al. (2013) (averaged to 5' from original spatial resolution of 30")	7
Layers	-	-	2 confined, 100 m thick each	-
Land surface elevation	-	L	5' average of 30" digital elevation map of Fan et al. (2013) (Fig. S2)	-
E-folding factor	-	-	Applied only to lower layer for 150 m depth, based on area-weighted average of Fan et al. (2013)	-
Timestep	t	T	Daily timestep	-
Convergence criterion	-	L	$\|$hydraulic head residuals $\|_{inf} < 10^{-100}$ and max head change < 10 m	-
Inner iterations	-	-	50 inner iterations between Picard iterations (Naff, Richard L., and Edward R. Banta, 2008)	-