Geoscientific
Model Development

Challenges in developing a global gradient-based groundwater model (G^{3}M v1.0) for the integration into a global hydrological model

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

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

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

Abstract. In global hydrological models, groundwater (GW) is typically represented by a bucket-like linear groundwater reservoir. Reservoir models, however, (1) can only simulate GW discharge to surface water (SW) bodies but not recharge from SW to GW, (2) provide no information on the location of the GW table, and (3) assume that there is no GW flow among grid cells. This may lead, for example, to an underestimation of groundwater resources in semiarid areas where GW is often replenished by SW or to an underestimation of evapotranspiration where the GW table is close to the land surface. To overcome these limitations, it is necessary to replace the reservoir model in global hydrological models with a hydraulic head gradient-based GW flow model.

We present G^{3}M, a new global gradient-based GW model with a spatial resolution of 5′ (arcminutes), which is to be integrated into 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, which allows sensitivity analyses, calibration, and data assimilation. This paper presents the G^{3}M concept and model design decisions that are specific to the large grid size required for a global-scale model. Model results under steady-state naturalized conditions, i.e., neglecting GW abstractions, are shown. Simulated hydraulic heads show better agreement to observations around the world compared to the model output of de Graaf et al. (2015). Locations of simulated SW recharge to GW are found, as is expected, in dry and mountainous regions but areal extent of SW recharge may be underestimated. Globally, GW discharge to rivers is by far the dominant flow component such that lateral GW flows only become a large fraction of total diffuse and focused recharge in the case of losing rivers, some mountainous areas, and some areas with very low GW recharge. A strong sensitivity of simulated hydraulic heads to the spatial resolution of the model and the related choice of the water table elevation of surface water bodies was found. We suggest to investigate how global-scale groundwater modeling at 5′ spatial resolution can benefit from more highly resolved land surface elevation data.

1 Introduction

Groundwater (GW) is the source of about 40 % of all human water abstractions (Döll et al., 2014) and is also an essential source of water for freshwater biota in rivers, lakes, and wetlands. GW strongly affects river flow regimes and supplies the majority of river water during ecologically and economically critical periods with little precipitation. 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 (Döll et al., 2014; Scanlon et al., 2012; Wada et al., 2012; Konikow, 2011). 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 (Wada, 2016; Döll et al., 2014; Gleeson et al., 2012, 2016). The strategic importance of GW for global water and food security will probably intensify under climate change as more frequent and intense climate extremes increase variability in surface water (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 (IS-ARM) (http://isarm.org, last access: 1 June 2019) and the ongoing GW component of the Transboundary Waters Assessment Program (TWAP) (http://www.geftwap.org, last access: 1 June 2019). To support prioritization for investment among transboundary aquifers, as 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 interactions with SW and soil is needed not only at the local and regional scale 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; Wada, 2016; Döll et al., 2012, 2014). In particular, they serve to quantify GW recharge (Döll and Fiedler, 2008). Like typical 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. Losing streams typically occur in semiarid and arid regions but also seasonally 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 a 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 and soil, it is therefore necessary to replace the linear GW reservoir model in GHMs by a hydraulic gradient-based GW flow model.

Large-scale gradient-based GW flow models are still rare and mainly available for data-rich regions, e.g., Death Valley (Belcher and Sweetkind, 2010) and the Central Valley (Belcher and Sweetkind, 2010; 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 northern Africa; Gossel et al., 2004). However, they are in most cases not integrated within hydrological models that quan-

tify GW recharge based on climate data and provide information on the condition of SW (e.g., streamflow and storage). 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, 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 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″ (arcseconds) 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 by poor 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). Recently, 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, 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, to achieve plausible discharge performance, they found it necessary to increase drainage from GW to rivers beyond the drainage driven by the hydraulic head difference between GW and river. This additional drainage, which accounts for about 50 % of global GW flow into SW, is simulated as a function of GW storage above the floodplain.

In this study, we present the Global Gradient-based Groundwater Model (G³M), which is to be integrated into the GHM WaterGAP 2 to improve estimation of flows between SW and GW (affecting both streamflow and groundwater recharge and thus water availability for humans and ecosystems) and implement capillary rise (affecting evapotranspiration). Table 1 provides a comparative summary of G³M as well as the global groundwater models by Fan et al. (2013), de Graaf et al. (2015, 2017), and the continental-scale model ParFlow (Maxwell et al., 2015).

The objective of this paper is to learn from a steady-state model, a well-established first step in groundwater model development, to (1) understand the basic model behavior by limiting model complexity and degrees of freedom and thus

Table 1. Comparison of global- and continental-scale groundwater models. DEM is digital elevation model. n/a means not applicable.

Aspect	G^3M	de Graaf et al. (2015, 2017)	Fan et al. (2013)	ParFlow
Extent	global	global	global	continental USA
Resolution	$5'$	$6'$	$30''$	1 km
Software	G^3M-f	MODFLOW	unnamed	ParFlow
Computational expense	medium	medium	high	very high
Flow representation	3-D saturated	3-D saturated	2-D saturated	3-D saturated or unsaturated
Timescale	steady-state or (transient)	steady-state or transient	steady-state	steady-state or transient
Vertical layers	2	2	1	5
Full coupling possible	yes	yes	yes	yes (already coupled)
In-memory coupling	yes	no	n/a	yes
Constant saturated thickness	yes	yes	no	no
Impermeable bottom	no	no	no	yes
Surface water body location	in every cell	in almost every cell	no surface water	created during simulation
Surface water body elevation	P_{30} of $30''$ DEM	avg. of $30''$ DEM	n/a (outflow if WTD < 0.25 m)	n/a
Deviation from observations	large	very large	medium	medium

(2) providing insights into dominant processes and uncovering potential model-inherent characteristics difficult to observe in a fully coupled transient model. A transient model might obfuscate model-inherent trends due to the slow-changing nature of groundwater processes e.g., trends towards large overestimation or underestimation due to wrong parameterization. A fully coupled model furthermore adds complexity and uncertainty to the model outcome. The presented steady-state model is furthermore used to (3) investigate parameter sensitivity and sensitivity to spatial resolution. In addition, the steady-state solution can be used as (4) an initial condition for future fully coupled transient runs.

The model concept and equations as well as applied data and parameter values are presented in Sect. 2. In Sect. 3, we show steady-state results of G^3M driven by WGHM data. Simulated hydraulic heads are compared to observations world-wide and to the output of existing large-scale GW models (Table 1). Furthermore, sensitivity to parameters and grid size is shown for the example of New Zealand. Finally, the implications of modeling decisions and grid size are discussed (Sect. 4) and conclusions are drawn (Sect. 5).

2 Model description

2.1 G^3M model concept

Although G^3M is based on principles of the well-known GW flow modeling software MODFLOW (Harbaugh, 2005), G^3M differs in its parameterization 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 G^3M, 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 size of $5'$ by $5'$ (approx. 9 km by 9 km at the Equator) as this size fits well to WaterGAP and is smaller than the suggested $6'$ of Krakauer et al. (2014). WaterGAP 3 (Eisner, 2016) has the same cell size, and 36 such cells fit into one $0.5°$ Water-GAP 2 cell. Global climate data are only available for $0.5°$ grid cells. The land mask of G^3M, i.e., location and size of $5'$ grid cells, is that of WaterGAP 3 and encompasses 2.2 million $5'$ grid cells on each layer.

Due to the lack of the spatial distribution of hydrogeological properties, we chose to use, in the current version of G^3M, two GW layers with a vertical size of 100 m each (Fig. 1). We performed a sensitivity analysis that confirmed the findings of others (de Graaf et al., 2015) that the aquifer thickness has a relatively small impact on the model results. Therefore, selecting a uniform thickness of 100 m (motivated by the assumed depth of validity of the lithology data) (Fig. 1) worldwide for the first layer and also for the second layer is expected to lead to less uncertainties as compared to hydraulic conductivities and the surface water table elevation.

G^3M focuses on a plausible simulation of water flows between GW and SW, and we deemed it suitable to have an upper GW layer that interacts with SW and soil (the soil layer
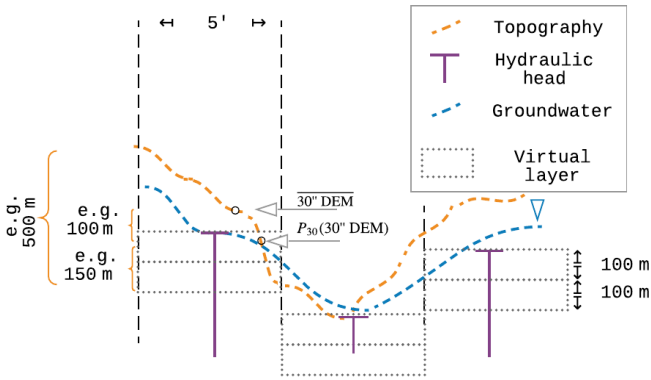
Figure 1. Schematic of G^{3}M's spatial structure with $5'$ grid cells, hydraulic head per cell, and the conceptual virtual layers (virtual because at this stage only confined conditions are computed). The underlying variability in the topography changes the perception of simulated depth to groundwater depending on what metrics are used to represent it on a coarser resolution. Layers in G^{3}M are of a conceptual nature and describe the saturated flow between locations of head laterally and vertically. The P_{30} is used in the presented steady-state model as SW elevation instead of an average or minimum per grid cell.

of WaterGAP is described in detail in Sect. S1 in the Supplement) and a lower one in which GW may flow laterally without such interactions. As land surface elevation within each $5'$ grid cell, with an area of approximately $80\,\mathrm{km}^2$, may vary by more than $200\,\mathrm{m}$ (Fig. S4.1), neighboring cells in G^{3}M may not be adjacent anymore (Fig. 1), in contrast to (regional) GW models with smaller grid cells. This makes G^{3}M a rather conceptual model in which water exchange between groundwater cells is driven by hydraulic head gradients but flow can no longer be conceptualized as occurring through continuous pore space. In addition, due to the coarse spatial scale and the possible large variations in land surface elevations within each grid cell, the upper model layers should not be considered to be aligned with an average land surface elevation. The model layers can be rather thought to be vertically aligned with the elevation of the surface water body table, as this prescribed elevation is, together with the sea level, the only elevation included in the groundwater flow equation (Eq. 1).

The simulation of aquifers that contain dry cells and/or cells that oscillate between wet and dry states poses great challenges to solving Eq. (1) (Niswonger et al., 2011). G^{3}M-f (the framework code used to implement G^{3}M) implements the traditional wetting approach from Harbaugh (2005) as well as the approach proposed by Niswonger et al. (2011) along with the proposed damping scheme. Both approaches have proven to be insufficient to simulate head-based transmissivities (unconfined conditions) on the global scale. Large mountainous areas would be excluded if unconfined conditions are assumed from the beginning of the solution step, as the head is often far below the deepest model layer, resulting

in a no-flow condition and imposing convergence issues to the matrix solver. We choose to simulate both layers with a specific saturated thickness even though the upper layer can be expected to decrease in water level and thus in transmissivity (hydraulic conductivity times saturated depth). The large uncertainties regarding hydraulic conductivities (possibly an order of magnitude), further justifies using the computationally more efficient assumption of specified saturated thickness. This approach is consistent with findings that this is accurate for large, complex groundwater models (Sheets et al., 2015). Furthermore, it is consistent with recent presented large-scale studies, e.g., for the Rhine–Meuse basin of Sutanudjaja et al. (2011) (using one confined layer), the Death Valley Regional Flow Model (Belcher, 2004; Faunt et al., 2011), and the global groundwater model of de Graaf et al. (2017) (two layers and partially unconfined conditions are simulated by parametrization of the model input and not by a head-dependant transmissivity).

Three-dimensional groundwater flow is described by a partial differential equation (approximated in the model implementation by using the finite differences method, Sect. 2.4)

$$\frac{\mathrm{dGWS}}{\mathrm{d}t} = \left(\frac{\partial}{\partial x}\left(K_x \frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_z \frac{\partial h}{\partial z}\right) + W\right)\Delta x \Delta y \Delta z = S_s \frac{\partial h}{\partial t}\Delta x \Delta y \Delta z, \tag{1}$$

where $K_{x,y,z}$ is the hydraulic conductivity $[LT^{-1}]$ along the x, y, and z axes between the cells (harmonic mean of grid cell 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 $[L]$. In- and out-flows in the groundwater are accounted for as

$$W\Delta x \Delta y \Delta z = R_g + Q_{swb} - NA_g - Q_{cr} + Q_{ocean}, \tag{2}$$

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

Q_{swb} in Eq. (3) replaces k_g GWS and R_{g_swb} in the linear storage equation of WaterGAP (Eq. S1), such that losing 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

$$Q_{swb} = \begin{cases} c_{swb}\left(h_{swb} - h_{aq}\right) & h_{aq} > B_{swb}, \\ c_{swb}\left(h_{swb} - B_{swb}\right) & h_{aq} \leq B_{swb}, \end{cases} \tag{3}$$

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

$$c_{\text{swb}} = \frac{K\,L\,W}{h_{\text{swb}} - B_{\text{swb}}}, \tag{4}$$

where K is hydraulic conductivity, L is length and W is width of the SW body per grid cell. For lakes (including reservoirs) and wetlands, c_{swb} is estimated based on hydraulic conductivity of the aquifer K_{aq} and SW body area (Table 2). 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 values so that, for steady-state conditions, the river is the sink for all the inflow to the grid cell (GW recharge and inflow from neighboring cells) that is not transported laterally to neighboring cells such that

$$c_{\text{riv}} = \frac{R_{\text{g}} + Q_{\text{eq}_{\text{lateral}}}}{h_{\text{eq}} - h_{\text{riv}}} \quad h_{\text{aq}} > h_{\text{riv}}. \tag{5}$$

For G^{3}M, 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, $Q_{\text{eq}_{\text{lateral}}}$ $[L^3T^{-1}]$, is the net lateral inflow into the cell computed based on the h_{eq} distribution as well as G^{3}M K_{aq} and cell thickness (Table 2). Elevation of the river water table, h_{riv} $[L]$, is provided by WGHM. Using a fully dynamic approach, i.e., utilizing the hydraulic head and lateral flows from the current iteration to recalculate c_{riv} in each iteration towards the steady-state solution, has proven to be too unstable due to its nonlinearity affecting convergence. We limit c_{riv} to a maximum of 10^7 m^2 d^{-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^{-1}.

If the river recharges the GW (losing river), Eq. (5) cannot be used as the Fan et al. (2013) high-resolution equilibrium model only models groundwater outflows but not inflows from SW bodies. If h_{aq} drops below h_{riv}, Eq. (4) is used to compute c_{riv}, with K equal to K_{aq}.

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

$$Q_{\text{ocean}} = c_{\text{ocean}} \left(h_{\text{ocean}} - h_{\text{aq}} \right), \tag{6}$$

where 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 2). We assume that the density difference to sea water is negligible at this scale. Q_{cr} is not yet implemented in G^{3}M.

2.2 The steady-state uncoupled model version

In the first implementation stage, G^{3}M was developed as a steady-state (right-hand side of Eq. 1 is zero) stand-alone 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 2 and described below.

Gleeson et al. (2014) provided a global subsurface permeability dataset from which K_{aq} was computed. The dataset 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, 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 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. S4.3). The original data were 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 by multiplying the upper layer value by $\exp(-50m\,f^{-1})^{-1}$ (Fan et al., 2007).

Mean annual GW recharge computed by WaterGAP 2.2c for the period 1901–2013 is used as input (Fig. S4.4), 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. The 0.5° data of WaterGAP were equally distributed to the pertaining cells. Regarding the ocean boundary condition, h_{ocean} is set to 0 m and c_{ocean} to 10 m^2 d^{-1} (Table 2), reflecting a global average conductance based on hydraulic conductivity and lateral surface area.

It is assumed that there is exchange of water between GW and one river stretch in each $5'$ grid cell, and additionally 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, in reality temporal variable, area of lakes and wetlands that

Table 2. Model parameter values, input data sources, and other information about the steady-state simulation.

Parameter	Symbol	Units	Description	Eq. no.
Land mask	–	–	location and area of 2 161 074 cells at $5'$ resolution based on WaterGAP 3 (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. S4.4)	2, 5, S1
Hydraulic conductivity	K_{aq}	LT^{-1}	derived from Gleeson et al. (2014) (Fig. S4.3)	1, 3
Hydraulic head	$h_{(aq)}$	L	head of the aquifer in a computational cell, initial estimate based on $5'$ average of $30''$ head by Fan et al. (2013)	1, 6, 5
Ocean boundary conductivity	c_{ocean}	$L^2 T^{-1}$	$10\,\mathrm{m}^2\,\mathrm{d}^{-1}$	2, 6
Ocean boundary head	h_{ocean}	L	global mean sea level of 0 m	6
SW head	h_{swb}	L	30th quantile (P_{30}) of $30''$ land surface elevation by Fan et al. (2013) per $5'$ grid cell	3
SW bottom elevation	B_{swb}	L	2 m (wetlands), 10 m (local lakes), 100 m (global lakes) below P_{30}	4
Area of global and local lakes and global and local wetlands	WL	L^2	per $5'$ grid cell, based on WaterGAP 3 (Eisner, 2016),	4
Length of the river	L	L	per $5'$ grid cell, based on WaterGAP 3 (Eisner, 2016)	4
Width of the river	W	L	per $5'$ grid cell, based on WaterGAP 3 (Eisner, 2016)	4
River head	h_{riv}	L	h_{swb}	4, 5
River bottom elevation	B_{riv}	L	$h_{riv} - 0.349 \times Q_{bankfull}^{0.341}$ (Allen et al., 1994)	5
Equilibrium hydraulic head	h_{eq}	L	steady-state hydraulic head by Fan et al. (2013) (averaged to $5'$ from original spatial resolution of $30''$)	5
Layers	–	–	two confined, 100 m thick each	–
Land surface elevation	–	L	$5'$ average of $30''$ digital elevation map by Fan et al. (2013) (Fig. S4.2)	–
E-folding factor	–	–	applied only to lower layer for 150 m depth, based on area-weighted average by Fan et al. (2013)	–
Time step	t	T	daily time step	–
Head convergence criterion (outer loop)	–	L	max head change globally < 10 m in three consecutive iterations	–
Residual convergence criterion (inner loop)	–	–	$\|$ conjugate gradient residuals $\|_{inf} < 10^{-100}$ in $\mathrm{m}^3\,\mathrm{d}^{-1}$	–
Maximum number of inner iterations	–	–	maximum 50 inner iterations between outer Picard iterations (Naff and Banta, 2008)	–

affect the flow exchange between SW body and GW. Maps generally show the maximum spatial extent of SW bodies. This maximum extent is seldom reached, in particular in the 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 80 % of their maximum extent is reached. In the transient model SW body areas change over time. 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 losing river may in reality dry out and therefore cease to lose any more water. For rivers, B_{swb} is set to $h_{riv} - 0.349 \times Q_{bankfull}^{0.341}$ (Allen et al., 1994), where $Q_{bankfull}$ is the bank-full river discharge in the $5'$ grid cell

(Verzano et al., 2012). Globally constant values are used for B_{swb} for wetlands, local lakes, and global lakes (Table 2).

For the steady-state model, all SW bodies in a grid cell are assumed to have the same head, i.e., $h_{riv} = h_{swb}$. We found that for both gaining and losing conditions, Q_{swb} and thus computed hydraulic heads are highly sensitive to h_{swb}. The overall best agreement with the hydraulic head observations by Fan et al. (2013) was achieved if h_{swb} (Eqs. 3, 4, and 5) was set to the 30th percentile (P_{30}) of the $30''$ land surface elevation values by Fan et al. (2013) per $5'$ cell, e.g., the $30''$ elevation that is exceeded by 70 % of the one-hundred $30''$ elevation values within one $5'$ cell. To decrease convergence time, we used h_{eq} derived from the high-resolution steady-state hydraulic head distribution by Fan et al. (2013) as initial guess of h_{aq}. In each outer iteration (Sect. 2.4) gaining and losing conditions may change depending on the current head solution.

2.3 Integration into WGHM

We intend to integrate G^3M into WaterGAP 2, i.e., the $0.5°$ version of WaterGAP (for details see Sect. S1), to keep computation time low enough for performing sensitivity analyses and ensemble-based data assimilation and calibration, instead of integrating it into WaterGAP 3 (Eisner, 2016), which has the same spatial resolution as G^3M. However, data from WaterGAP 3 were used to set up G^3M. Location and area of the $5'$ grid cells of G^3M are the same as in the land mask 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 (Table 2).

2.4 Model implementation

G^3M is implemented using a newly developed open-source model framework G^3M-f (Reinecke, 2018a). The main motivation to develop a new model framework is the efficient in-memory coupling to the GHM and flexible adaptation to the specific requirements of global-scale modeling. Written in C++14, the 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 an extension to WaterGAP but allows for an in-memory coupling to any GHM or can be used as a stand-alone groundwater model. It provides a unit-tested (Dustin, 2006) environment offering different modules that can couple in-memory results to a different model or write out data flows to different file formats. G^3M-f has the following advantages over using an established GW modeling software such as MODFLOW. G^3M-f enables an improved coupling capability. Unlike MODFLOW, it provides a clear development interface to the programmer coupling a model to G^3M-f. It can be easily compiled as a library and provides

a clearly separated logic between computation and data read-in or write-out). It is written in the same language as the target GHM enabling a straight-forward in-memory access to arrays without the need to write data to disk, required when coupling with MODFLOW (a very computationally expensive operation even if that disk is a RAM-disk). Even though it is possible to call FORTRAN functions from C++, it is very complicated to pass file pointers properly, as the I/O implementation of both languages differs substantially and it is widely considered bad practice to handle I/O in two different languages at once. As MODFLOW was never designed to be coupled or integrated to or into other models, it is not possible to separate the I/O logic fully from the computational logic without substantial code changes that are hard to test. To this end, G^3M-f provides a highly modularized framework that is written with extensibility as design goal while implementing all required groundwater mechanisms.

Equation (1) is reformulated as finite-difference equation and solved using a conjugate gradient approach and an incomplete LUT (incomplete lower–upper factorization with dual-threshold strategy) preconditioner (Saad, 1994). In order to keep the memory footprint low, the conjugate gradient method makes use of the sparse matrix. Furthermore, it solves the equations in parallel (preconditioner currently nonparallel). As internal numerical library, G^3M uses Eigen3 (http://eigen.tuxfamily.org, last access: 1 June 2019). G^3M can compute the presented steady-state solution (with the right-hand side of Eq. 1 being zero and the heads by Fan et al., 2013, as initial guess, Tables 1, 2) on a commodity computer with four computational cores and a standard solid-state drive in about 30 min while occupying 6 GB of RAM.

Similar to MODFLOW, G^3M-f solves Eq. (1) in two nested loops using a Picard iteration (Mehl, 2006): (1) the outer iteration checks the head and residual convergence criterion (if the maximum head change between iterations is below a given value in three consecutive iterations and/or the norm of the residual vector of the conjugate gradient (Harbaugh, 2005; Niswonger et al., 2011) is below a given value). It adjusts head-dependant values, for example, from gaining to losing conditions and updates the system of linear equations if flows are no longer head dependent. (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 2), solving the current system of linear equations.

Because switching between Eqs. (4) and (5), which occurs if, for example, 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} between Eqs. (4) and (5) 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 or vice versa.

Different from Vergnes et al. (2014), G^{3}M's computations are not based on spherical coordinates directly but on an irregular grid of quadratic cells of different size depending on the latitude. Cell sizes are provided by WaterGAP3 and are derived from their spherical coordinates maintaining their correct area and center location. The model code will be adapted in the future to account for the different lengths in x and y directions per cell correctly.

3 Results

3.1 Global hydraulic head and water table depth distribution under natural steady-state conditions

As expected, the computed global distribution of steady-state hydraulic head (in the upper model layer) under natural conditions (Fig. 2a) largely follows the land surface elevation (Fig. S4.2), albeit with a lower range and locally different ratios between the hydraulic head and land surface gradients (Fig. S4.6). Water table depth (WTD), i.e., the distance between the groundwater table and the land surface, can be computed by subtracting the hydraulic head computed by G^{3}M for the upper layer of each 5' grid cell from the arithmetic mean of the land surface elevations of the one-hundred 30″ grid cells within each 5' cells (Fig. S4.2). The global map of steady-state WTD (Fig. 2b) clearly resembles the map of differences between surface elevation and P_{30}, the assumed water level of SW bodies h_{swb}, shown in Fig. S4.1, which indicates that simulated WTD is strongly governed by the assumed water level in SW bodies.

Deep GW, i.e., a large WTD, occurs mainly in mountainous regions (Fig. 2b). These high values of WTD are mainly a reflection of the steep relief in these areas as quantified either by the differences in mean land surface elevations between neighboring grid cells (Fig. S4.7) or the difference between mean land surface elevation and P_{30}, the 30th percentile of the 30″ land surface elevations (Fig. S4.1). When computed hydraulic head is subtracted not from average land surface elevation but from P_{30} (the assumed water table elevation of SW bodies), the resulting map shows that the groundwater table is mostly above P_{30}, in both flat and steep terrain (Fig. 2c). Thus, high WTD values at the 5' resolution do not indicate deep unsaturated zones and losing rivers but just high land surface elevation variations within a grid cell. In steep terrain, 5' water tables are higher above water level in the surface water bodies than in flat terrain (Fig. 2c). Deep GW tables that are not only far below the mean land surface elevation but also below the water table of surface water bodies are simulated to occur in some (steep or flat) desert areas with very low GW recharge. Negative WTD only occurs in places were the P_{30} is above the mean surface elevation, e.g., parts of the Netherlands (Fig. 2b). Fewer than 10 cells experience WTD less than -10 m, which is very likely due to a not fully converged head solution.
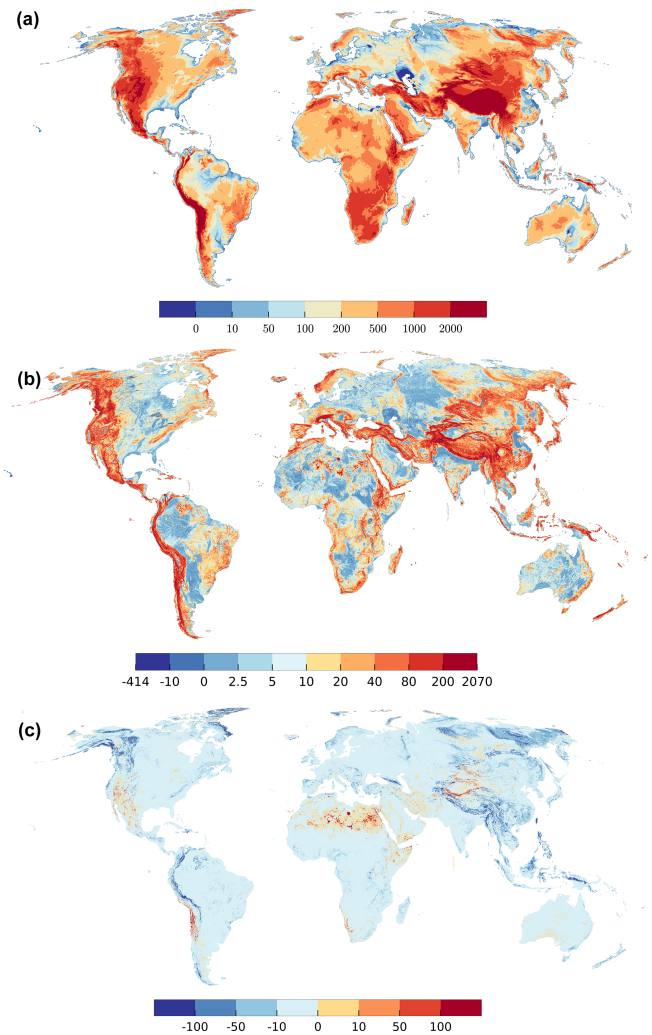
Figure 2. (a) Simulated steady-state hydraulic head of groundwater above sea level (meters). Maximum value 6375 m, minimum -414 m (extremes included in dark blue and dark red). **(b)** Water table depth (meters). **(c)** Difference between 30th percentile of the 30″ land surface elevation per 5' grid cell (chosen elevation for surface water bodies h_{swb}) and simulated groundwater head (meters). Maximum value 1723 m, minimum value -1340 m (extremes included in dark blue and dark red).

In 2.1 % of all cells, GW head is simulated to be above the average land surface elevation by more than 1 m in 0.3 % and by more than 100 m in 0.004 % of the cells. 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 (Sect. 2.2). Please note that the computed steady-state WTD certainly underestimates the steady WTD in GW depletion areas such as the High Plains Aquifer and the Central Valley in the USA (Sect. S2), northwestern India, the North China Plain, and parts of Saudi Arabia and Iran (Döll et al., 2014) as ground-
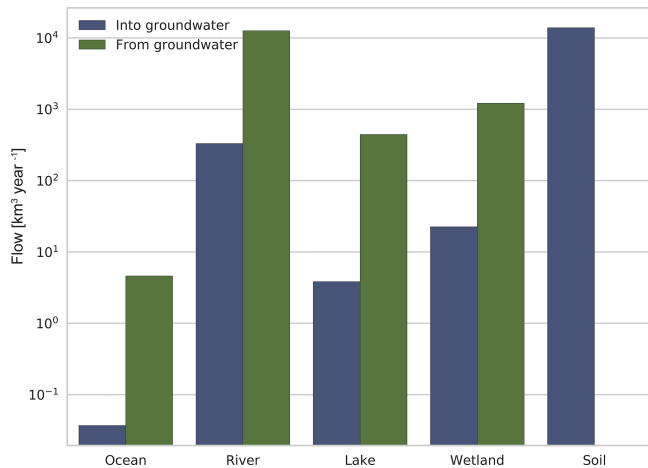
Figure 3. Global sums of flows from different compartments into or from GW at steady state. Flows into the GW are denoted by the color blue, flows out of the GW into the different compartments by green. The compartment soil is the diffuse GW recharge from soil calculated by WaterGAP.

water withdrawals are not taken into account in the presented steady-state simulation of G^3M.

3.2 Global water budget

Inflows to and outflows from GW of all G^3M grid cells were aggregated according to the compartments ocean, river, lake, wetland, and diffuse GW recharge from soil (Fig. 3). The difference between the global sum of inflows and outflows is less than 10^{-6} %. This small volume balance error indicates the correctness of the numerical solution.

Total diffuse GW recharge, model input from WaterGAP, from soil is 10^4 km^3 yr^{-1} and approximately equal to the simulated flow of GW to rivers (Fig. 3). Rivers are the ubiquitous drainage component of the model, followed by wetlands, lakes, and the ocean boundary. According to G^3M, the amount of river water that recharges GW is more than 1 order of magnitude smaller than GW flow to rivers (Fig. 3). Possibly, flow from SW bodies to GW is overestimated, as outflow from SW is not limited by water availability in the SW, and depending on the hydraulic conductivity, Eqs. (4) and (5) can lead to rather large flows. Inflow from the ocean, which is more than 2 orders of magnitude smaller than outflow to ocean, occurs in regions where $h_{\text{swb}} = P_{30}$ is below h_{ocean}, e.g., the Netherlands. Globally, lakes and wetlands are computed to receive up to 10^3 km^3 yr^{-1} of water from GW, and lose 1–2 orders of magnitude less.

3.3 GW–SW interactions

Figure 4 plots the spatial distribution of simulated flows from and to lakes and wetlands (Fig. 4a) as well as from and to rivers (Fig. 4b). Parallel to the overall budget (Fig. 3), the map reveals the globally large, but locally strongly varying,
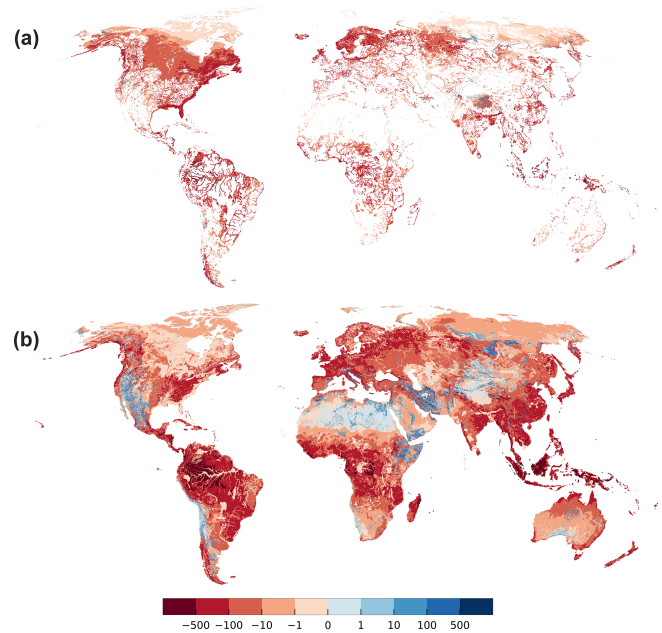
Figure 4. Flow Q_{swb} (mm yr^{-1}) from or to wetlands, lakes **(a)**, and losing or gaining streams **(b)** with respect to the $5'$ grid cell area. Gaining surface water bodies are shown in red, surface water bodies recharging the aquifer in blue. Focused aquifer recharge occurs in arid regions, e.g., alongside the river Nile, and in mountainous regions where the average water table is well below the land surface elevation.

influence of lakes and wetlands (Fig. 4a). 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 Fig. 4a and b). Similarly, areas dominated by wetlands and lakes (e.g., parts of Canada and Scandinavia) show less inflow for rivers (Fig. 4b). In G^3M, all SW bodies (rivers, lakes, and wetlands) in a grid cell either lose or gain water. Consistent with negative or positive differences between h_{swb} and h_{aq} (Fig. 2c), 93 % of all grid cells contain gaining rivers and only 7 % losing 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.

Gaining rivers, lakes, and wetlands with very high absolute Q_{swb} values over 500 mm yr^{-1} (averaged over the grid cell area of approximately 80 km^2) can be found in the Amazon and Congo basins as well in Bangladesh and Indonesia, where GW recharge in very high (Fig. S4.4). Values below 1 mm yr^{-1} occur in dry and permafrost areas where groundwater recharge is low.

Losing SW bodies are caused by a combination of low GW recharge from soil (Fig. S4.4) and steep mountainous terrain (Fig. S4.7). While the steep Himalayas receive enough GW recharge to have gaining SW bodies, this is not the case for the much drier mountain ranges around the Taklamakan desert in central Asia or mountainous Iran where SW bod-

ies are losing. In the Sahara, GW recharge is so low that SW bodies are losing even in relatively high terrain.

Rivers lose more than $100\,\mathrm{mm\,yr^{-1}}$ 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, and lower values in the Sahara and Kazakhstan. The river Nile in northern Republic of Sudan and Egypt is correctly simulated to be a losing river (Fig. 4b), being an allogenic river that is mainly sourced from the upstream humid areas, including the artificial Lake Nasser (Elsawwaf et al., 2014) (Fig. 4a). Furthermore, 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 simulated along the Niger River and its wetlands and almost none in the northeastern Brazil even though losing conditions are known to occur there (Costa et al., 2013; FAO, 1997). This is also true when the minimum elevation for SW bodies is assumed (compare Fig. S4.10) leading to the conclusion that the misrepresentation might be linked to an inadequate representation of the local geology.

Simulated flows between GW and SW depend on assumed conductances for both rivers and lakes or wetlands (Eqs. 3, 4, 5) shown in Fig. 5. Q_{swb} (Fig. 4) correlates positively with conductance. Conductance for gaining rivers correlates positively with GW recharge (Eq. 5 and Fig. S4.4). High river conductance values are reached in the tropical zone due to a high GW recharge but are capped at a plausible maximum value of $10^7\,\mathrm{m^2\,d^{-1}}$ in the case of a river (Sect. 2.1) (Fig. 5b). Lakes and wetlands can have larger values of conductance due to their large areas, e.g., in Canada or Florida.

3.4   Lateral flows

Figure 6 shows lateral GW flow (between grid cells, summing up over all 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 neighboring cells depends on five main factors: (1) hydraulic conductivity (Fig. S4.3), (2) diffuse GW recharge (Fig. S4.4), (3) losing or gaining SW bodies (Fig. 4), (4) their conductance (Fig. 5), and (5) the head gradients (Fig. 2a).

On large areas of the globe, where GW discharges to SW bodies, the lateral flow percentage is less than 0.5 % of the total GW recharge to the grid cell, as most of the GW recharge in a grid cell is simulated to leave the grid cell by discharge to SW bodies. For example, in the permafrost regions, the assumed very low hydraulic conductivity limits the outflow to neighboring 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 con-
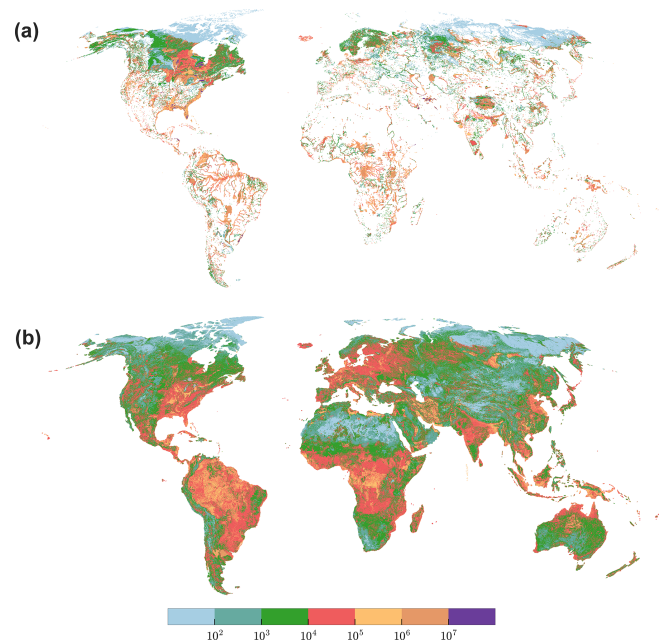
Figure 5. Conductance $(\mathrm{m^2\,d^{-1}})$ of lakes and wetlands **(a)** and rivers **(b)**. In regions close to the pole, conductance is in general lower due to the influence of the low aquifer conductivity (losing conditions), and relatively small GW recharge due to permafrost conditions (only applies for gaining conditions). Max conductance of wetlands is $10^8\,\mathrm{m^2\,d^{-1}}$.
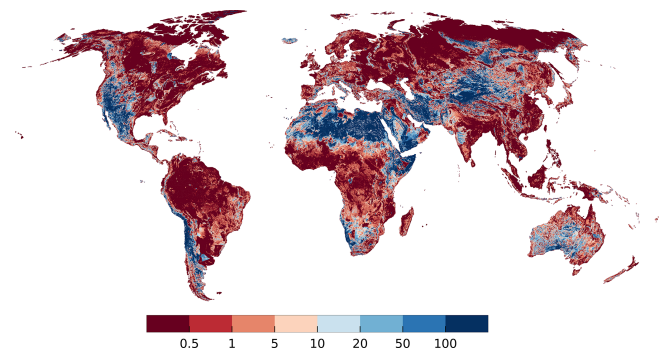
Figure 6. Percentage of GW recharge from soil and surface water inflow that is transferred to neighboring cells through lateral out flow (sum of both layers). Grid cells with zero total GW recharge are shown in white (a few cells in the Sahara and the Andes).

ductivities, 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 in the case of gaining SW bodies in dry areas like Australia or in the Taklamakan desert. They can also be observed 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 the lateral directions, and values often exceed 100 % (Fig. 6).

3.5 Comparison to groundwater well observations and the output of two higher-resolution models

Global observations of WTD 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 a total of 1 070 402 WTD observations. An "observed head" per 5′ model cell was then calculated by first computing the hydraulic head of each observation by subtracting WTD from the 5′ average of the 30″ land surface elevation and then calculating the arithmetic mean of all observations within the 5′ model cell. This resulted in 78 664 grid cells with observations out of a total of 2.2 million G^3M top-layer grid cells. Multiple obstacles limit the comparability of 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. 7. Shallow GW 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 G^3M due to lacking information. Because the steady-state model cannot take into account the impact of GW abstraction, the computed WTD 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 the High Plains Aquifer in the American Midwest of the USA. Still, the elevation of the GW table in the nondepleted Rhine valley in Germany is overestimated, too. Overestimates in the Netherlands may partially be due to artificial draining. Figure 8a shows the hydraulic head comparison as a 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) (their Fig. 6).

To compare the performance of G^3M to the steady-state results of two high-resolution models by Fan et al. (2013) and ParFlow (Maxwell et al., 2015) (Table 1), heads in 30″ (Fan et al., 2013) and 1 km (ParFlow) grid cells were averaged to the G^3M 5′ grid cells. The comparison of 5′ observations to the 5′ average of ParFlow seems to be consistent with the 1 km model comparison in Maxwell et al. (2015) (their Fig. 5), even though overestimates or underestimates

in the original resolution seem to be smoothed out by averaging to 5′ (not shown). The heads by Fan et al. (2013) fit better to observations than G^3M heads, with less underestimation (Fig. 8b) and a RMSE (root mean square error) of 26.0 m compared to the 32.4 m RMSE of G^3M. The comparison of G^3M heads to values by Fan et al. (2013) for all 5′ grid cells, which are also the initial heads of G^3M and the basis to compute river conductances, shows that heads computed with the G^3M are mostly much lower except in regions with a shallow GW (Fig. 8c); RMSE is 46.7 m. 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 G^3M heads than the initial values (Fan et al., 2013), are mainly occurring in steep mountainous areas like the Himalayas where the 5′ model is not representing smaller valleys with a lower head. For the continental US, the computationally expensive 1 km integrated hydrological model ParFlow (Maxwell et al., 2015) fits much better to observations than G^3M (Fig. 8d, e), with a RMSE of 14.3 m (ParFlow) compared to 34.2 m (G^3M). G^3M produces a generally lower water table (Fig. 8f), a main reason being that ParFlow assumes an impermeable bedrock at a depth of 100 m below the land surface elevation.

The global map of head comparison (Fig. 7) suggests that G^3M performs reasonably well in flat areas compared to mountainous regions. This is corroborated by Fig. 9, which shows the difference between observed and simulated hydraulic heads for five land surface elevation categories. It is evident that model performance deteriorates with increasing land surface elevation and positively correlates with variations in land surface elevation within each grid cell (Fig. S4.7).

Plotting hydraulic head instead of WTD 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. Well observations provide WTD and only sometimes contain complementary data specifying the elevation at which the measurements were taken. Even though hydraulic heads are a direct result of the model and are forcing lateral GW flows, WTD is more relevant for processes like capillary rise. For G^3M, there is almost no correlation between WTD observations and simulated values. To our knowledge, no publication on large-scale GW modeling has presented correlations of simulated with observed WTD.

## 3.6 Testing sensitivity of computed steady-state hydraulic heads to parameter values and spatial resolution

To limit the computational effort for assessing model sensitivity to both parameters and grid size, we selected New Zealand as a representative "small world" that includes a complex topography and the ocean as a clear boundary con-
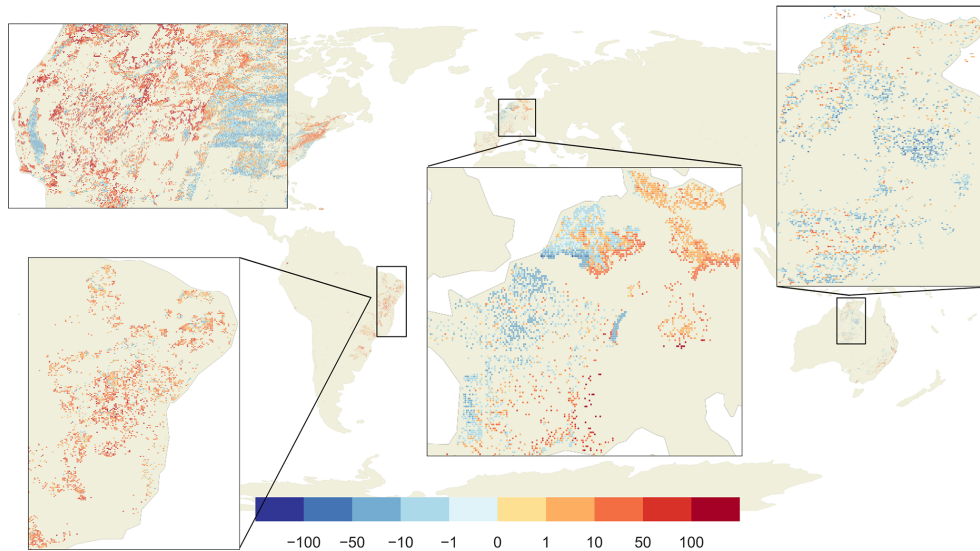
Figure 7. Differences between observed and simulated hydraulic head (metres). Red dots show areas where the model simulated deeper GW as observed, blue shallower GW. In the gray areas, no observations are available.
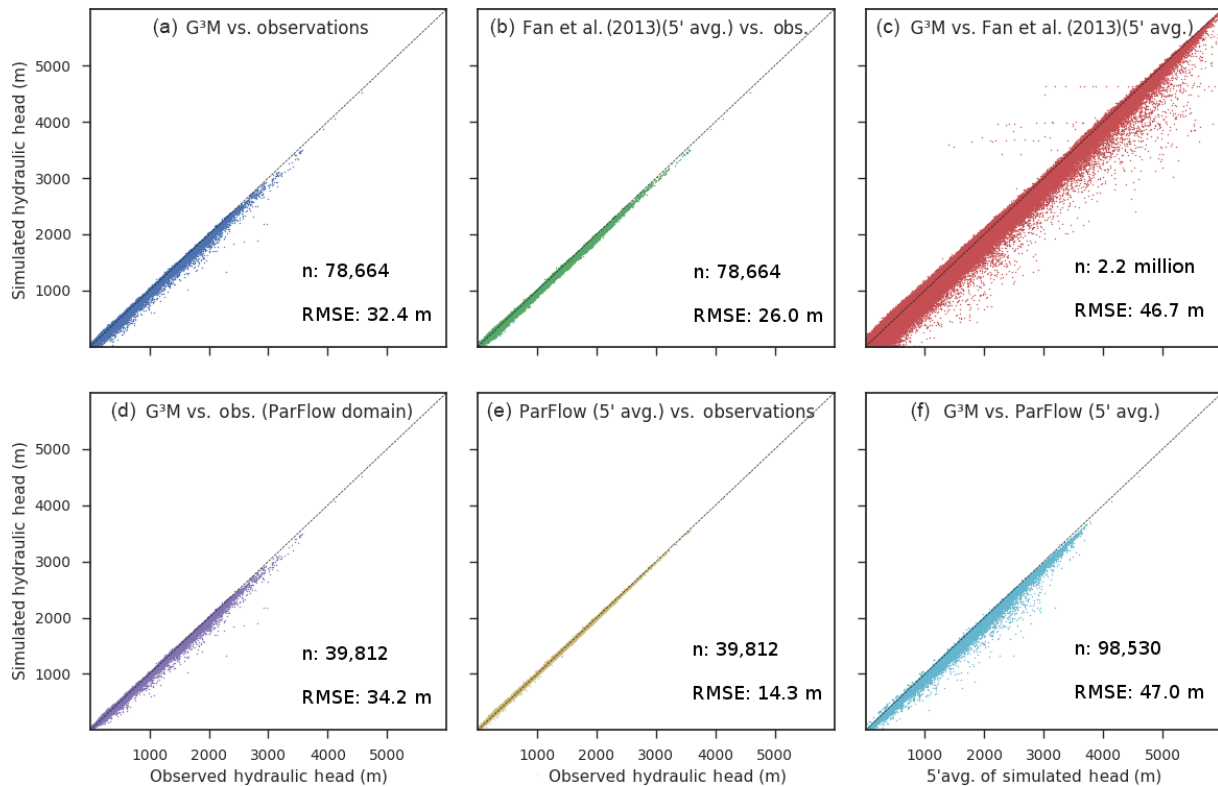
Figure 8. Scatter plots of simulated vs. observed hydraulic head and inter-model comparison of heads. The steady-state run of G^3M vs. observations (**a**), the $5'$ average of the equilibrium head by Fan et al. (2013) vs. observations (**b**), and the avg. equilibrium vs. G^3M (**c**). The steady-state run of G^3M 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 G^3M vs. $5'$ average of the ParFlow average annual GW table (**f**).
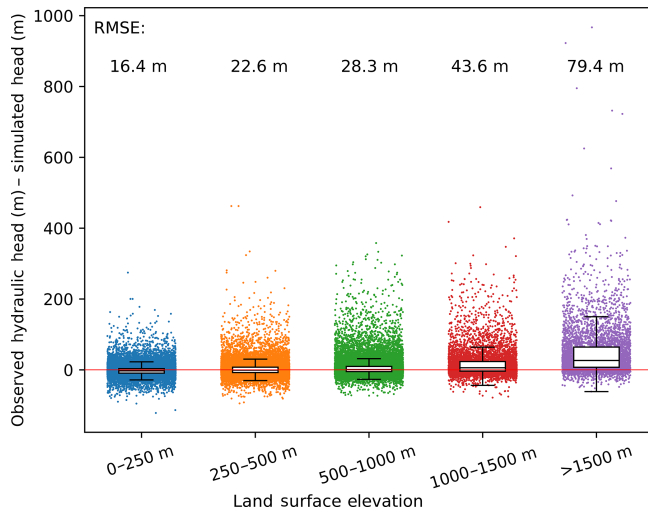
Figure 9. Observed minus simulated hydraulic head for different land surface elevation categories. The whiskers of the boxplots show the interquartile range.

Figure 10. Low (5′) vs. high (30″) spatial resolution for the Canterbury region in New Zealand: comparison of observed vs. simulated hydraulic head for both resolutions. The observed head is the geometric mean per 5′ and 30″, respectively.

dition. All inputs and parameters are the same as in the global 5′ model.

3.6.1 Parameter sensitivity

To determine which parameters simulated hydraulic heads are most sensitive to, we used the established sensitivity tool UCODE 2005 (Hill and Tiedeman, 2007) to compute composite scaled sensitivity (CSS) values for seven model parameters (Sect. S3). CSS of h_{swb} is orders of magnitude larger than the CSS of the other parameters. This confirms our observations during model development when an appropriate value had to be found (Sect. 2.2). The second-most important parameter is K_{aq} and the third-most important R_g. CSS of the conductance of lakes is 1 order of magnitude less than CSS of R_g but as only few cells contain lakes, the CSS value that averages over all grid cells indicates a large sensitivity to c_{Lakes} for grid cells with lakes. Simulated hydraulic heads were found to be rather insensitive to changes in the conductance of rivers, wetlands, or ocean boundary.

3.6.2 Sensitivity to spatial resolution

The extremely high sensitivity of simulated hydraulic heads to the choice of h_{swb} (Sect. 3.6.1) and the better agreement of the continental models with a higher spatial resolution of approx. 30″ (the Fan et al., 2013, model and ParFlow, Sect. 3.5) motivated us to run G^{3}M for New Zealand with a spatial resolution of 30″ to understand the impact of spatial resolution on simulated hydraulic heads. The 30″ G^{3}M model uses the same input as the 5′ model except for the land surface elevation, h_{swb}, and the location of rivers. While the total lengths and widths of the rivers are equal in both models – a river is assumed to exist in all 5′ grid cells – the river is concen-

trated – in the 30″ model – to a few 30″ grid cells within each 5′ grid cell. The river cell locations at 30″ are determined based on 30″ HydroSHEDS (https://hydrosheds.org/, last access: 1 June 2019) information on flow accumulation. Starting with the 30″ cell with the highest number of upstream cells per 5′ cell, a river is added to this 30″ cell using the length and information of HydroSHEDS until the size of the river of the 5′ model is reached for all 30″ cells within a 5′ cell. The areal fraction of all other SW bodies from 5′ grid data is used for all 30″ grid cells within the 5′ grid cell. h_{swb} is set to the land surface elevation.

Figure 10 compares the performance of the two model versions. The comparison of simulated hydraulic head to observations for the Canterbury region (Westerhoff et al., 2018) shows that the overall performance of the 30″ model is better, with a smaller RMSE of 26.7 m as compared to a RMSE of 53.8 m in the case of the original spatial resolution of 5′. The 30″ model results in generally lower simulated hydraulic heads leading to a closer fit to the observed values. This is likely caused by the improved estimation of SW body elevation, which generally leads to lower estimates of h_{swb}. On the other hand, overestimates of observed hydraulic heads prevails in the 30″ model, even though h_{swb} was set to the land surface elevation, indicating that further investigation is necessary. The underestimates are likely due to large GW abstractions for irrigation in the particular region (Westerhoff et al., 2018).

4 Discussion

The objective of global gradient-based groundwater flow modeling with G^{3}M is to better simulate water exchange

between SW and GW in the GHM WaterGAP, e.g., for an improved estimation of GW resources in dry regions of the globe that are augmented by focused recharge from SW bodies. We assume that the fully coupled model will lead to an improved WaterGAP performance during droughts with an increased drop in streamflow due to the now possible switch from gaining to losing conditions. It is presumable that a calibration of c_{swb} and c_{riv} is necessary to achieve a good discharge performance. The presented steady-state model is the first step in this direction. It helped to understand basic model behavior, e.g., the sensitivity to SW body elevation, and the necessary improvement of its parameterization, before moving to the more complex integrated transient model. The reduced runtime of the steady-state model in comparison to a fully integrated transient run supported the investigation of parameter sensitivity and sensitivity to spatial resolution. Additionally, the presented steady-state model can be used in future fully integrated transient runs as initial condition.

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. S4.1), while the vertical 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 before a dynamic coupling of G³M with a global hydrological model has been achieved, but one may speculate that some innovative approach to take into account the elevation variations within the grid cells is needed.

It is difficult to the assess quality of the presented steady-state G³M results. Model performance assessment is hindered by poor data availability and the coarse model resolution. (1) To our knowledge the data collection of depth to groundwater by Fan et al. (2013) is unique. However, they do not represent steady-state values. Apart from depth to groundwater observations, hardly any relevant data are available at the global scale. Especially the exchange between surface water and groundwater is difficult to measure even at the local scale. Therefore, we compared G³M results with the results from other large-scale models. Comparison to the results of catchment-scale groundwater flow models is planned for transient runs that will be possible after integration into WaterGAP. (2) Scale differences make the comparison to point observations of depth to groundwater difficult. Often, observations are biased towards alluvial aquifers in valleys. The calculated hydraulic head of the grid cell may represent the average groundwater level per grid cell correctly but can be still far off the local observations of depth to groundwater.

As the current model only represents an uncalibrated natural steady state, a comparison to observations only provides the first indicator where the model and the performance measurements need to be improved as we move to a fully transient model.

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 of topography, climate, and geology (Fan et al., 2007) and how the interaction to SW bodies governs 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 naturally 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 by Fan et al. (2013). We also tested setting h_{swb} to the average elevation of all "blue" cells (with a WTD of less than 0.25 m) of the steady-state 30″ water table results by Fan et al. (2013) that indicate the locations where GW discharges to the surface or SW bodies. This leads to an overall underestimation of the observed hydraulic heads (Fig. S4.9) as the assumed SW elevation is too low. Furthermore, it leads to an increase in losing SW bodies (compare Fig. S4.10 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 the case of lower h_{swb}, which appears to be reasonable. Additionally, choosing the average as SW elevation provides, on the on hand, a better fit to observations (Fig. S4.9 right) but leads to a worldwide flooding (Fig. S4.9) and a much longer convergence time due to an increased oscillation between gaining and losing conditions.

The problem is very likely one of scale. All three models (Fan et al., 2013; ParFlow, and G³M 30″) (Table 1), even the simple one by Fan et al. (2013), fit better to observations than the 5′ model G³M (Figs. 8, 10). In the case of high resolution, 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 the case of the low spatial resolution of 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. S4.8 center), resulting in an underestimation of hydraulic head and thus an overestimation of WTD. While it is plausible and necessary to assume that there is SW–GW interaction within each of the approximately 80 km² cells, this is not the case for the 2 orders of magnitude smaller 30″ grid cells. Thus, with 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 trial-and-error – to

lead to a hydraulic head distribution that fits reasonably well to observed head values. This avoids the situation when the simulated GW table drops too low while also 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. S4.9).

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 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 the case of a K_{aq} value of $10^{-5}\,\mathrm{m\,s^{-1}}$, the SW body would lose approximately 1 m of water each day. Further investigations are needed regarding the appropriate choice of SW body elevation and conductance. The simple conductance approach applied in G^{3}M could possibly be improved by the approach by Morel-Seytoux et al. (2017), who proposes an analytical, and physically based, estimate of the leakance coefficient for coarse-scale models based on river and aquifer properties.

De Graaf et al. (2015) set their SW head (h_{swb}) to the mean land surface elevation (Table 1) of the $6'$ grid cells minus river depth at bank-full conditions plus water depth at average river discharge as compared to P_{30} in the $5'$ G^{3}M. 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 that was necessary to improve the discharge to rivers (Sutanudjaja et al., 2014). On the other hand, the additional drainage leads to drainage of water even if the hydraulic head is below the SW elevation, which might have led to the global underestimation of hydraulic heads. Thus, the difference in model heads seems to be closely related to the sensitivity to SW body elevation.

Due to the course spatial resolution and lack of data, G^{3}M does not capture the actual variability in topography, aquifer depth (Richey et al., 2015) or (vertical) heterogeneity of subsurface properties. 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 groundwater along thousands of kilometers from the Rocky Mountains in the central USA to the coast as well as the vertical connectivity are likely to be overestimated by G^{3}M, as vertical faults and interspersed aquitards are not represented; this is expected to lead 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 $5'$ GW flow model G^{3}M that is to be integrated into the 0.5° GHM WaterGAP. The uncoupled steady-state model has provided important insights into challenges of global GW flow modeling mainly related to the necessarily large grid cell sizes ($5'$ by $5'$). In addition, first global maps of SW–GW interactions were generated. Simulated heads were found to be strongly impacted by assumptions regarding the interaction with SW bodies, in particular the selected elevation of the SW table. We have demonstrated that simulated G^{3}M hydraulic heads fit better to observed heads when compared to the heads of the comparable steady-state GW model by de Graaf et al. (2015), without requiring additional drainage. Furthermore, we provided insights into how the choice of surface water body elevation h_{swb} affects model outcome. In a next step, approaches for utilizing high-resolution topographic data to improve the selection of h_{swb} will be investigated.

The presented results are the first step towards a fully coupled model in which SW heads are jointly computed, also taking into account the impact of SW and GW abstraction. Especially the interaction with SW bodies that can run dry will make the G^{3}M behavior more realistic. The fully coupled model will simulate transient behavior 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. However, it will be challenging to judge the dynamics of GW and the quality of simulated GW–SW interactions due to a scarcity of observations.

References

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

Belcher, W. R. (Ed.): Death Valley Regional Ground-water Flow System, Nevada and California-Hydrogeologic Framework and Transient Ground-water Flow Model, U.S. Geological Survey Professional Paper, 408 pp., available at: https://pubs.usgs.gov/sir/2004/5205/ (last access: 1 June 2019), 2004.

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 1711, 398 pp., available at: https://pubs.usgs.gov/pp/1711/ (last access: 1 June 2019), 2010.

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, https://doi.org/10.1002/hyp.9243, 2013.

de Graaf, I. E. M., 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, https://doi.org/10.5194/hess-19-823-2015, 2015.

de Graaf, I. E. M., 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, Adv. Water Resour., 102, 53–67, https://doi.org/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, https://doi.org/10.3390/w8120592, 2016.

Döll, P. and Fiedler, K.: Global-scale modeling of groundwater recharge, Hydrol. Earth Syst. Sci., 12, 863–885, https://doi.org/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, J. Geodyn., 59–60, 143–156, https://doi.org/10.1016/j.jog.2011.05.001, 2012.

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, https://doi.org/10.1002/2014WR015595, 2014.

Dustin, E.: Effective software testing: 50 specific ways to improve your testing, 5th printing, Addison-Wesley, Boston, Mass., USA, XV, 271 s., 2006.

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, https://doi.org/10.1002/hyp.9563, 2014.

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, D10125, https://doi.org/10.1029/2006JD008111, 2007.

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

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

Faunt, C. C. (Ed.): Groundwater Availability of the Central Valley Aquifer, California, U.S. Geological Survey Professional Paper, 1766, 225 pp., available at: https://pubs.usgs.gov/pp/1766/ (last access: 1 June 2019), 2009.

Faunt, C. C., Provost, A. M., Hill, M. C., and Belcher, W. R.: Comment on "An unconfined groundwater model of the Death Valley Regional Flow System and a comparison to its confined predecessor" by R.W.H. Carroll, G.M. Pohll and R.L. Hershey [J. Hydrol. 373/3–4, pp. 316–328], J. Hydrol., 397, 306–309, https://doi.org/10.1016/j.jhydrol.2010.11.038, 2011.

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, https://doi.org/10.1038/nature11295, 2012.

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, https://doi.org/10.1002/2014GL059856, 2014.

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

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), Hydrogeol. J., 12, 698–713, https://doi.org/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, USA, 2005.

Hartmann, J. and Moosdorf, N.: The new global lithological map database GLiM: A representation of rock properties at

the Earth surface, Geochem. Geophy. Geosy., 13, Q12004, https://doi.org/10.1029/2012GC004370, 2012.

Hill, M. C. and Tiedeman, C. R.: Effective groundwater model calibration: With analysis of data, sensitivities, predictions, and uncertainty, Wiley-Interscience, Hoboken, NJ, USA, 2007.

Konikow, L. F.: Contribution of global groundwater depletion since 1900 to sea-level rise, Geophys. Res. Lett, 38, L17401, https://doi.org/10.1029/2011GL048604, 2011.

Krakauer, N. Y., Li, H., and Fan, Y.: Groundwater flow across spatial scales: importance for climate modeling, Environ. Res. Lett., 9, 34003, https://doi.org/10.1088/1748-9326/9/3/034003, 2014.

Lange, S.: EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI), GFZ Data Services, https://doi.org/10.5880/pik.2016.004, 2016.

Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, J. Hydrol., 296, 1–22, https://doi.org/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, https://doi.org/10.5194/gmd-8-923-2015, 2015.

Mehl, S.: Use of Picard and Newton iteration for solving nonlinear ground water flow equations, Ground Water, 44, 583–594, https://doi.org/10.1111/j.1745-6584.2006.00207.x, 2006.

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, https://doi.org/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, https://doi.org/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, https://doi.org/10.5194/hess-18-3511-2014, 2014.

Naff, R. L. and Banta, E. R.: The US Geological Survey modular ground-water model-PCGN: a preconditioned conjugate gradient solver with improved nonlinear control, Geological Survey (US), 2008-1331, available at: https://pubs.er.usgs.gov/publication/ofr20081331 (last access: 1 June 2019), 2008.

Niswonger, R. G., Panday, S., and Ibaraki, M.: MODFLOW-NWT, a Newton formulation for MODFLOW-2005, US Geological Survey Techniques and Methods, 6-A37, https://doi.org/10.3133/tm6A37, 2011.

Reinecke, R.: G^{3}M-f a global gradient-based groundwater modelling framwork, JOSS, 3, 548, https://doi.org/10.21105/joss.00548, 2018a.

Reinecke, R.: G^{3}M framework, Zenodo, https://doi.org/10.5281/zenodo.1175540, 2018b.

Reinecke, R.: Model Output, Zenodo, https://doi.org/10.5281/zenodo.1315471, 2018c.

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, https://doi.org/10.1002/2015WR017351, 2015.

Saad, Y.: ILUT: A dual threshold incomplete LU factorization, Numer. Linear Algebr., 1, 387–402, https://doi.org/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, P. Natl. Acad. Sci. USA, 109, 9320–9325, https://doi.org/10.1073/pnas.1200311109, 2012.

Sheets, R. A., Hill, M. C., Haitjema, H. M., Provost, A. M., and Masterson, J. P.: Simulation of water-table aquifers using specified saturated thickness, Ground Water, 53, 151–157, https://doi.org/10.1111/gwat.12164, 2015.

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 pp., available at: https://pubs.usgs.gov/pp/pp1703/ (last access: 1 June 2019), 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, https://doi.org/10.5194/hess-15-2913-2011, 2011.

Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F. P.: Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data, Water Resour. Res., 50, 687–705, https://doi.org/10.1002/2013WR013807, 2014.

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, Nat. Clim. Change, 3, 322–329, https://doi.org/10.1038/nclimate1744, 2012.

UNESCO-IHP, IGRAC, WWAP: GEF-TWAP Methodology Transboundary Aquifers, available at: http://isarm.org/sites/default/files/resources/files/TWAP Methodology Groundwater Component (Revised Aug 2012).pdf (last access: 9 October 2017), 2012.

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, W07517, https://doi.org/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. Hydrometeorol., 13, 1149–1171, https://doi.org/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, 11065–11086, https://doi.org/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, J. Hydrol., 424–425, 238–251, https://doi.org/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, https://doi.org/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, W00L06, https://doi.org/10.1029/2011WR010562, 2012.

Westerhoff, R., White, P., and Miguez-Macho, G.: Application of an improved global-scale groundwater model for water table estimation across New Zealand, Hydrol. Earth Syst. Sci., 22, 6449–6472, https://doi.org/10.5194/hess-22-6449-2018, 2018.