Uncertainty in model estimates of global groundwater depth

Robert Reinecke, Sebastian Gnann, Lina Stein, Marc Bierkens, Inge de Graaf, Tom Gleeson, Gualbert Oude Essink, Edwin H Sutanudjaja, Claudia Ruz Vargas, Jarno Verkaik, & Thorsten Wagener

2024

Original citation:
Reinecke, R., Gnann, S., Stein, L., Bierkens, M., De Graaf, I., Gleeson, T., Essink, G. O., Sutanudjaja, E. H., Vargas, C. R., Verkaik, J., & Wagener, T. (2024). Uncertainty in model estimates of global groundwater depth. *Environmental Research Letters*, *19*(11), 114066. https://doi.org/10.1088/1748-9326/ad8587

IOP Publishing *Environ. Res. Lett.* **19** (2024) 114066 https://doi.org/10.1088/1748-9326/ad8587

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
15 February 2024

REVISED
31 August 2024

ACCEPTED FOR PUBLICATION
10 October 2024

PUBLISHED
17 October 2024

LETTER

Uncertainty in model estimates of global groundwater depth

Robert Reinecke[1,2,*], **Sebastian Gnann**[1,3], **Lina Stein**[1], **Marc Bierkens**[4,5], **Inge de Graaf**[6], **Tom Gleeson**[7], **Gualbert Oude Essink**[4,5], **Edwin H Sutanudjaja**[5], **Claudia Ruz Vargas**[8], **Jarno Verkaik**[4,5] **and Thorsten Wagener**[1]

[1] Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany
[2] Institute of Geography, Johannes Gutenberg-University Mainz, Mainz, Germany
[3] Chair of Hydrology, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany
[4] Unit Subsurface and Groundwater Systems, Deltares, Utrecht, The Netherlands
[5] Department of Physical Geography, Utrecht University, Utrecht, The Netherlands
[6] Earth Systems and Global Change, Wageningen University & Research, The Netherlands
[7] Department of Civil Engineering, University of Victoria, Victoria, Canada
[8] International Groundwater Resources Assessment Centre (IGRAC), Delft, The Netherlands
[*] Author to whom any correspondence should be addressed.

E-mail: reinecke@uni-mainz.de

Keywords: groundwater, global models, uncertainty

Supplementary material for this article is available online

Abstract

Knowing the depth at which groundwater can be found below the land surface is critical for understanding its potential accessibility by ecosystems and society. Uncertainty in global scale water table depth (WTD) limits our ability to assess groundwater's role in a water cycle altered by changing climate, land cover, and human water use. Global groundwater models offer a top–down pathway to gain this knowledge, but their uncertainty is currently poorly quantified. Here, we investigate four global groundwater models and reveal steady-state WTD disagreements of more than 100 m for one-third of the global land area. We find that model estimates of land areas with shallow groundwater at <10 m depth vary from 10% to 71% (mean of 23%). This uncertainty directly translates into subsequent assessments, as land areas with potential groundwater accessibility for forests, population, and areas equipped for irrigation, differ substantially depending on the chosen model. We explore reasons for these differences and find that contrary to observations, 3 out of 4 models show deeper water tables in humid than in arid climates and greatly overestimate how strongly topographic slope controls WTD. These results highlight substantial uncertainty associated with any global-scale groundwater analysis, which should be considered and ultimately reduced.

1. Introduction

Groundwater makes up 99% of all non-frozen 'fresh'water (deeper groundwater may be saline) on our planet (Gleeson *et al* 2016, United Nations 2022), sustaining ecosystems by providing water to vegetation (Koirala *et al* 2017, Roebroek *et al* 2020), rivers, lakes, and wetlands (Winter 1995, de Graaf *et al* 2019, Hare *et al* 2021, Jasechko *et al* 2021), and being a pivotal ecosystem by itself (Saccò *et al* 2023). Groundwater offers a relatively constant supply of freshwater to 43% of the world's irrigated agriculture (United Nations 2015, 2022) and safe drinking water to an estimated 3.7 billion people (United Nations 2015, 2022). While surface water supply is increasingly fragile due to climate change (IPCC 2022), groundwater is assumed to remain a more reliable source of freshwater (Taylor *et al* 2013). However, the recent IPCC 6th assessment report concluded that 'limitations in the spatio-temporal coverage of groundwater monitoring networks, abstraction data and numerical representations of groundwater recharge processes continue to constrain understanding of climate change impacts on groundwater' (IPCC 2022). This lack of knowledge has consequences in at least three critical aspects relevant to society: the accessibility of groundwater for ecosystems, drinking water supply, and agricultural use. Global-scale groundwater models have recently been developed to represent these key issues in hydrological assessments (Fan *et al* 2013, 2017, de Graaf *et al* 2015, 2019, Reinecke *et al* 2019b, Condon *et al* 2021), yet model uncertainties remain poorly explored (Gleeson *et al* 2021).

A key variable to understand groundwater in all three societal contexts is water table depth (WTD). Here, we define WTD as depth from the land surface to the top of the saturated zone (Baird and Low 2022). Recent efforts have shown the global complexity of WTD evolution with highly heterogeneous rising and falling patters of water levels (Jasechko *et al* 2024). Groundwater can quickly become inaccessible for ecosystems if WTD declines beyond the depth of vegetation roots (Fan *et al* 2017) or below the bottom of rivers and lakes (de Graaf *et al* 2019). Humans may build wells reaching down to hundreds of meters (Perrone and Jasechko 2019), yet below a certain depth, reduced permeability will decrease groundwater yield (Perrone and Jasechko 2019, Bierkens *et al* 2022), deeper wells are costlier to build and operate (Perrone and Jasechko 2019, Bierkens *et al* 2022), and water drawn from deeper aquifers may require desalinization (Kang and Jackson 2016, Kang *et al* 2019). We define groundwater as potentially accessible if the WTD is shallow enough to be used by ecosystems and/or humans. Notably, potentially accessible groundwater does not mean that (financial) resources are in place to access this water and to transport it to its destination, that the water is of adequate quality, or that the hydrogeological configuration allows for abstraction.

The availability of direct observations of WTD varies widely in quality and consistency (figure S1(a), Jasechko *et al* 2024). Models, on the other hand, can create consistency of information and, crucially, interpolate and extrapolate to regions where observations are unavailable (Gleeson *et al* 2021). Furthermore, they can be run with altered boundary conditions, which is key for understanding how Earth system dynamics may change with human interventions across scales (Wang-Erlandsson *et al* 2022). Through the connection of groundwater to streams, wetlands, and lakes, understanding WTD enables us to understand, for example, how streamflow will be affected by climate change and by groundwater abstractions (de Graaf *et al* 2019), and how it will affect sea level rise and water available for atmospheric circulation (Maxwell and Condon 2016). Global groundwater models are already utilized by different research communities; for example, their simulations are used in regional studies on groundwater accessibility by vegetation in the Amazon (Costa *et al* 2023), to quantify groundwater-vegetation interactions (Koirala *et al* 2017), or to identify ephemeral streams (Brinkerhoff *et al* 2024), to estimate regolith thickness (Pelletier *et al* 2016), to map groundwater dependent ecosystems (Link *et al* 2023), and their estimates are contained in datasets like HydroATLAS (Linke *et al* 2019).

However, the reliability of current models in support of water policy and scientific inquiry is highly debated within the community, and the need for model improvements has been discussed (Gleeson *et al* 2021). Central to this debate is the issue that these global models currently suffer from highly uncertain or missing global datasets for parameterization (e.g. for the representation of important vertical hydrogeological configurations) and missing or biased observations for evaluation. As a result of limited data for parameterization and of high computational demand they currently simulate at coarser spatial resolutions (here 5 arcmin to 1 km), and less vertical layers (two or continuous assumption of decreasing hydraulic conductivity) than local scale models. Despite groundwater's crucial role in the Earth system, we cannot yet provide a robust global picture of WTD. The ongoing discussions point to considerable uncertainties in global model estimates, which have not yet been quantified. Here, we take a close look at the current generation of models and quantify differences in steady-state WTD between models. While differences between models are to be expected due to different conceptual choices in model building and different input datasets, the reasons for these differences are not yet known. We thus take a first step in diagnosing the reasons for these differences, and demonstrate what they could entail for subsequent assessments. Simulating a steady-state first is standard practice for groundwater models (Hill and Tiedeman 2007), as this equilibrium state often initializes storage for a transient simulation. However, before moving to a time-varying analysis that contains additional uncertainties, e.g. due to poorly known pumping rates (Gleeson *et al* 2021, Bierkens *et al* 2022), we assessed here how far these models agree on their estimates of a natural steady-state. Steady state WTD estimates are also used in many existing studies and datasets (e.g. Pelletier *et al* 2016, Linke *et al* 2019, Link *et al* 2023), making it critical to better understand their uncertainties.

2. Methods

2.1. Models

Our analysis used the outputs of four published global groundwater models: Verkaik *et al* (2024), Fan *et al* (2017) (an update to Fan *et al* 2013), Reinecke *et al* (2019b), and de Graaf *et al* (2015) (table S1 summarizes the main model differences). In the following, they are indicated only by their first author names because some models do not have a specific name. The models exclude Greenland and Antarctica. All models represent a global steady-state WTD, which was calculated by removing the time component from the flow-equation and thus simulating an equilibrium water table, except for Fan, who reached this state by a transient daily simulation of the period 2004–2014. All simulations used here do not account for anthropogenic change, e.g. no pumping is implemented, even though the transient versions of the models may consider water abstractions (Reinecke, de Graaf, Verkaik). We chose also to focus on the steady-state here, because many assessments currently only use steady-state WTD (possibly also implicitly through datasets like HydroATLAS Linke *et al* 2019). In addition, groundwater moves slowly; thus, this starting point is important to understand model differences of transient simulations. Even though pumping might alter the system more rapidly, different steady-state WTD estimates also hint at conceptual differences between models that will persist in transient simulations.

Of the four models, two simulate groundwater on coarse spatial resolutions (5 arcmin Reinecke and de Graaf) and two on a finer spatial resolution of 30 arcsec (Fan, Verkaik). All models except Fan utilize an annual mean recharge of the global hydrological model they are coupled with (table S1). With respect to their parameterization, all models differ except for the Verkaik model, which utilized resampled inputs from the model of de Graaf and a subsequent model version of de Graaf *et al* (2017). Table S1 summarizes the central model properties and parameterizations.

For the calculation of the ensemble mean, model results were aggregated (resampling method = average) to a spatial resolution of 5 arcmin using GDAL. The uncertainty range was computed by calculating the difference between the maximum and minimum of the WTD for every grid cell of the ensemble. All assessments regarding relative area were calculated with cell areas based on a global equal area projection.

2.2. Separation into three categories

Whether models can simulate groundwater depths adequately has direct implications for their application and future model improvement. We thus created water table accessibility categories (potentially accessible to ecosystems (<10 m WTD), potentially accessible for irrigation or drinking water supply (<100 m), and costly to access or inaccessible (>100 m) based on global and large-scale datasets of rooting depth (Fan *et al* 2017), potential groundwater-stream connectivity (Jasechko *et al* 2021), and well depth (Jasechko *et al* 2021) (see also figure S2 and supplement for an extended discussion).

2.3. Uncertainty impact assessment

We evaluate model uncertainty as expressed via the three critical demands: ecosystem demand, drinking water supply and agricultural use. Here we analyze forests as a critical terrestrial ecosystem and carbon sink (Humphrey *et al* 2018), population as a proxy for where groundwater might be important to domestic use and industry (Ferguson *et al* 2018), and (current) irrigated area as a proxy for the potential use of groundwater for agriculture (Bierkens *et al* 2022).

To assess the impacts of model uncertainty on ecosystems, population, and agriculture, we used three different data sources. Global tree cover data (Hansen *et al* 2013) on 30 m resolution was aggregated to 5 arcmin. The data representing the % coverage was then converted to area using the land mask covered by the model ensemble. Population data for the year 2020 (constrained version; https://hub.worldpop.org) on a 100 m resolution was aggregated (resampling method = sum) to 5 arcmin and cut to the land mask covered by the model ensemble. This resulted in a slight decrease of the global population as coastal areas are not as well represented by the coarser global model mask. Global irrigated areas on 5 arcmin resolution (Siebert *et al* 2015) were used to calculate the areas equipped for irrigation. The three 5 arcmin data products were spatially joined using GDAL with the calculated uncertainty range of the ensemble.

2.4. Model evaluation

We can learn from comparing models with each other, with our expectations, and with available observations (Gleeson *et al* 2021). Large scale or global groundwater model evaluation is most commonly performed using point observations of WTD (often converted to hydraulic head) (Fan *et al* 2013, 2017, de Graaf *et al* 2015, 2017, Reinecke *et al* 2019b, 2020, de Graaf and Stahl 2022). This approach, however, provides little insight into the reasons for model disagreement, is limited to few (geographically biased) locations relative to the simulated domain, and suffers from commensurability issues Beven and Cloke (2012).

As an alternative, we can evaluate global-scale groundwater models by investigating so-called functional relationships between known drivers of groundwater flow and WTD (Gleeson *et al* 2021, Wagener *et al* 2022, Gnann *et al* 2023), including how well models reproduce such relationships in comparison to our current process understanding. For example, using the concept of water table ratio (Haitjema and Mitchell-Bruker 2005, Cuthbert *et al* 2019), we can conceptualize the natural water table as driven by four main factors: (i) climatic aridity, here defined as potential evapotranspiration divided by precipitation which serves as an indicator for the amount of groundwater recharge in figure 3, (ii) topography (approximated by topographic slope), (iii) subsurface permeability, and (iv) interactions with surface water bodies. We would, for example, expect deep water tables in dry, steep, highly permeable regions, far away from perennial streams.

Here we use WTD observations from Fan *et al* (2013), the aridity data are based on CHELSA data at 30 arcsec resolution (Karger *et al* 2017). Slope data are based on 250 m slope data from the Geomorpho90m dataset (Amatulli *et al* 2020) and elevation data (used in the supplement) are based on 250 m elevation data from Tomislav Hengl (2018); both are based on the MERIT DEM (Yamazaki *et al* 2017). All raster data (aridity, slope, WTD from all models) were resampled to 5 arcmin resolution using GDAL (resampling method = median) and aligned to exactly overlay. Resampling may influence driver-WTD relationships as it smooths out variability. Overall, however, the patterns are only slightly affected (figures S13 and 14).

Aridity was calculated by dividing potential evapotranspiration by precipitation (PET/P), both from CHELSA. Values below one indicate energy-limited environments, while values above one indicate water-limited environments. Using other data products, approaches and thresholds to calculate aridity will produce different aridity maps, yet this will not substantially change the fact that we, for example, expect deeper water tables in regions which tend to be water-limited.

3. Results and discussion

3.1. Global variation of potential groundwater accessibility for ecosystems and humans

Using our definitions of accessibility, we find that, averaged across all models, 23% (10%–71% uncertainty; minimum ensemble WTD—maximum ensemble WTD) of global groundwater is potentially accessible to ecosystems, 57% (28%–57%) is potentially accessible to humans, and 20% (2%–35%) is potentially costly or inaccessible. These numbers are calculated from the ensemble mean estimates of the four global groundwater models (figure 1(a); see also Methods and table S1). Shallow water tables accessible to ecosystems are located along coastlines and in regions with major aquifers, such as the Amazon Basin, the Central Valley aquifer, the Ganges-Brahmaputra Basin, and the Mississippi Alluvial Aquifer. Costly or inaccessible groundwater is mainly located in mountainous regions such as the Rocky Mountains, the Andes, and the Himalayas. The latter potentially shows the limits of models to simulate local groundwater systems in topographically diverse regions, even if groundwater provides a vital influx to streams in mountainous regions.

It is important to remember that the mean WTD shown here represents the long-term average (steady-state) of a natural world without human impacts (e.g. pumping). A consequence of choosing a steady-state here is a tendency to see shallower water tables because they do not include an anthropogenic fingerprint. For example, a shallow water table in the Central Valley aquifer, as shown in figure 1(a), is reasonable in a steady-state simulation if no groundwater abstraction is included.

3.2. Global uncertainty in steady-state WTD

While the ensemble mean WTD shows patterns that agree with our general conceptual understanding of groundwater processes, such as deeper water tables in drier or more mountainous regions (Cuthbert *et al* 2019), the inter-model differences are substantial. We show areas of considerable uncertainty as blank spaces to indicate where current models provide no robust estimates (Wagener *et al* 2021) in figure 1(b). For one-third of the global land area, the models show disagreements in WTD of more than 100 m. Green areas depict where models tend to agree in absolute terms, with differences of no more than 10 m amounting to only 12% of the global land area. Areas of high agreement include the Central Valley aquifer, the Mississippi Alluvial Aquifer, and the Ganges-Brahmaputra Basin.

Model differences broadly reflect topography and are exceptionally high in mountainous regions, such as the Rocky Mountains, the Andes, and the Himalayas. However, we also see significant differences in flatter regions if they are in dry climates, such as the Sahara, South Africa, and Australia. While models generally agree that water tables are deeper in these regions (>100 m), the models strongly disagree on how deep, often by several hundreds of meters. There is a strong positive correlation between the depth of the mean groundwater table and the degree of uncertainty (Spearman rank correlation $\rho_s = 0.96$, $p = 0.00$; figure S3).

However, while the models agree more in regions with lower topographic slopes and shallower water tables, the uncertainty in these regions might be more consequential. Relative uncertainty (i.e. uncertainty in relation to the mean WTD, see figure S4) is less correlated with topography and thus more strongly highlights flatter areas where models disagree (in relative terms), such as parts of the Amazon basin and the West-Siberian plain. In these flatter regions, a difference in WTD of 5 m can have an immense impact on the accessibility of water for roots (Fan *et al* 2017), capillary rise (Kroes *et al* 2018), and surface water connectivity (de Graaf *et al* 2019).

Looking closer into the model differences, figure 2 reveals that the model of de Graaf (figure 2(a)) and Verkaik (figure 2(b)) are generally deeper than the ensemble mean. However, de Graaf seems to be shallower in the Sahara desert. These two models mainly differ in their spatial resolution. While Reinecke (figure 2(c)) is conceptually similar to both models (number of layers and MODFLOW-like equations (Herrera *et al* 2023); see also table S1) it shows a shallower water table compared to the ensemble mean in mountains and deserts. Fan differs from

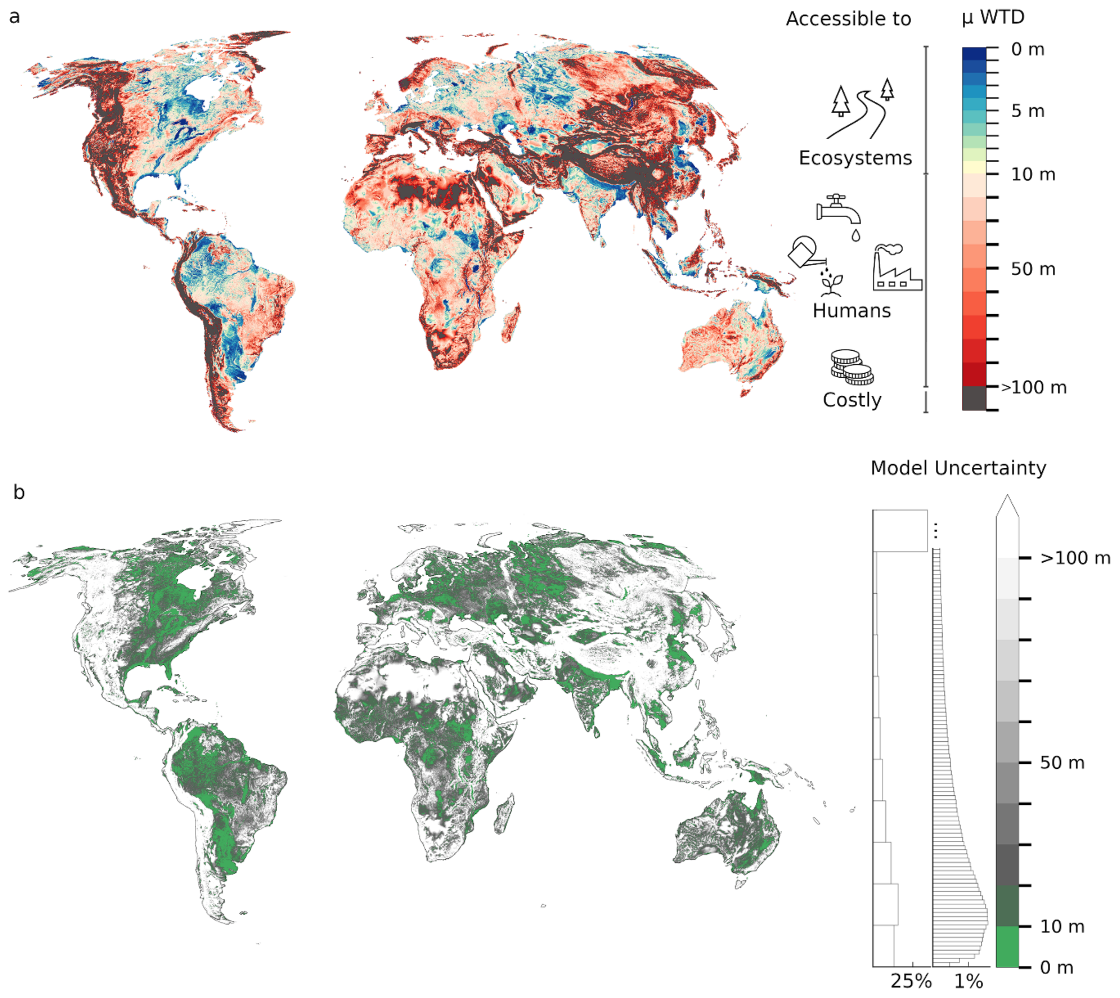

Figure 1. (a) Ensemble mean (μ) of steady-state water table depth (WTD) calculated using four steady-state global groundwater models with categorization into three categories of potential accessibility: by ecosystems, humans, and costly or inaccessible. (b) Uncertainty in WTD (highest minus lowest value per grid cell), also represented by two histograms (based on number of grid cells not area) with a bin size of 1 m and 10 m. Blank spaces in the map indicate areas with very uncertain model estimates (Wagener *et al* 2021).

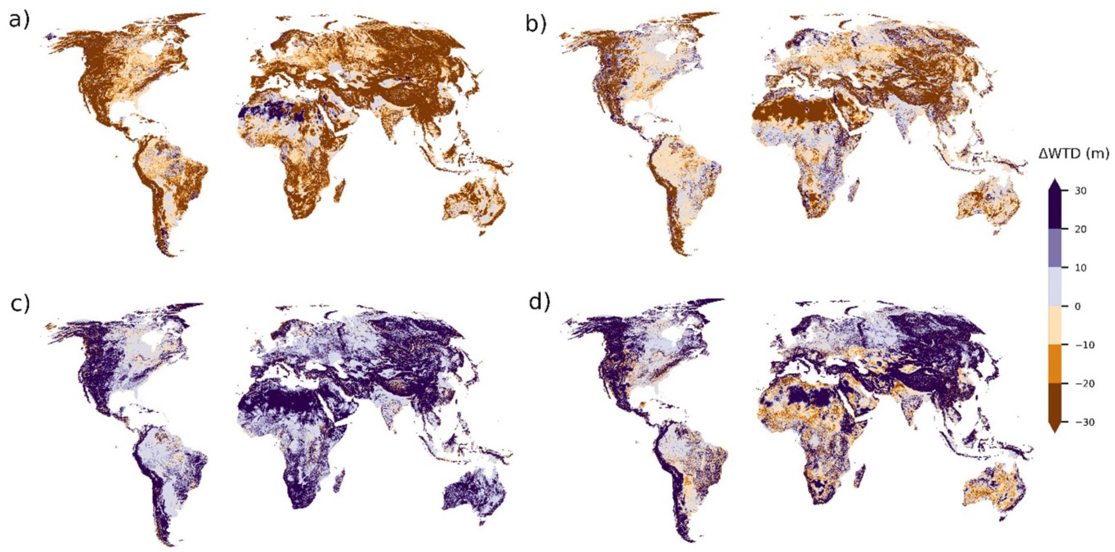

Figure 2. Differences in WTD (m) between each of the four models and the ensemble mean. (a) de Graaf, (b) Verkaik, (c) Reinecke, and (d) Fan.

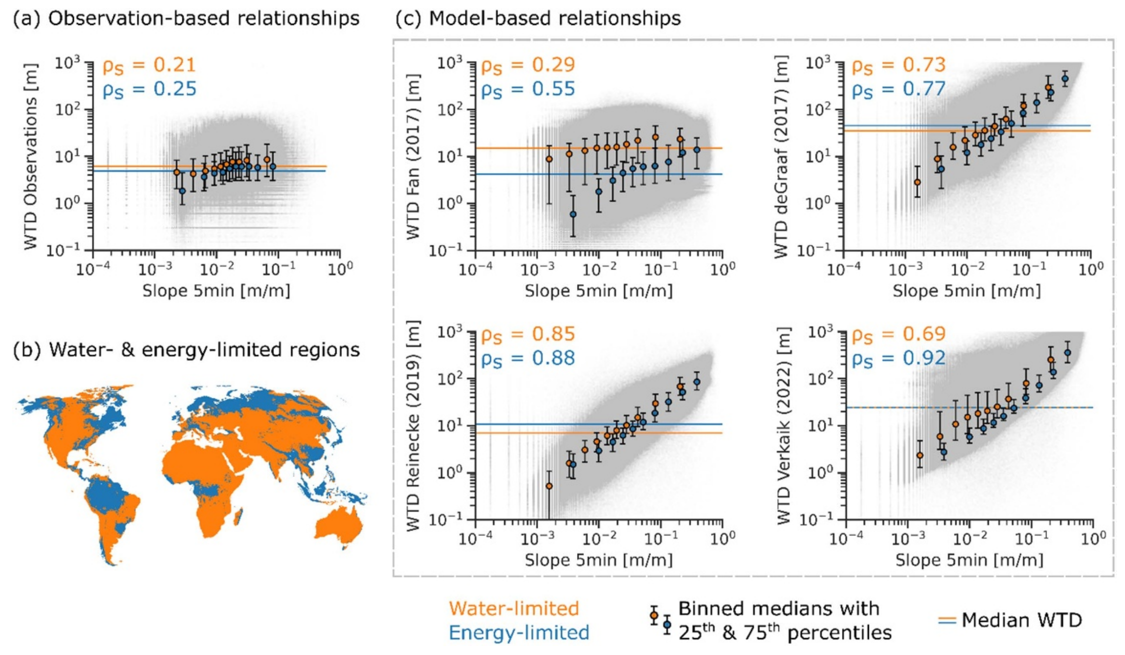

Figure 3. Relationship between topographic slope and (a) observed and (b) simulated WTD from four global models. (b) Location of water-limited (i.e. rather dry) and energy-limited (i.e. rather wet) regions. Spearman rank correlations shown in (a) and (c) are based on the point cloud, separated by aridity index. Bin averages are displayed as a visual aid and are separated based on the aridity index (orange and blue; see Methods for estimation and figure S9 for a global map). Topographic slope, aridity, and modeled WTD have been aggregated to a resolution of 5 arcmin. For the observations, the WTD values were compared with the 5 arcmin values of the grid cell in which the observations are located. Each bin contains 10% of the data (spread evenly across all slope values). The correlations are calculated using all data points and are therefore unaffected by the bins, which are primarily there for visualization. Observational data used is possibly highly affected by water abstractions or return flows. Model-based relationships are based on the entire model domain and not only on the model cells in which observations were available. The steady-state outputs of the models do not account for this anthropogenic impact.

all other models in showing deeper tables in some non-mountainous regions but overall shallower water tables than the ensemble mean. Differences of models in dry regions might be due to differences in groundwater recharge in these regions (West *et al* 2022).

3.3. Models disagree with observation-based relationships

The large disagreement in WTD estimates across current models suggests that there is potential for learning about the reasons for these differences. To achieve this, we explore driver-WTD relationships between models and between models and the largest available dataset (Fan *et al* 2013) (figure 3 and tables S3 and S4). With drivers of WTD we here mainly refer to topography and climate represented as slope and water/energy-limited regions.

In agreement with our conceptual understanding (Cuthbert *et al* 2019), observations show deeper water tables in water-limited regions than in energy-limited regions (6.1 m vs. 4.9 m, respectively), and deeper water tables for steeper slopes (Spearman rank correlations are $\rho_s = 0.21$ and 0.25, for water-limited and energy-limited regions, respectively). Deeper water tables in arid regions are estimated by Fan (15.0 m vs. 4.2 m), but not by Verkaik (24.4 m vs. 24.5 m), Reinecke (6.9 m vs. 10.7 m) and de Graaf (34.8 m vs. 45.2 m). The model of Fan shows medium correlations with slope (0.29 and 0.55), while the models of Reinecke (0.85 and 0.88), de Graaf (0.73 and 0.77), and Verkaik (0.69 and 0.92) show high correlations with slope, particularly in energy-limited regions. We find weak inverse relationships between permeability and WTD for all models (ρ_s ranges between −0.25 and −0.09 and is slightly higher for energy-limited regions; table S4), while observations show no clear relationship. Models also differ in how WTD correlates with distance to perennial streams, but there is no consistent pattern (ρ_s between −0.19 and 0.38 for water-limited regions, and between −0.04 and 0.16 for energy-limited regions; figure S8). In summary, we find topographic slope to be the dominant control in most models, while it is less pronounced in the observations.

3.4. Reasons for model differences

Multiple reasons contribute to the differences between the four models investigated here, including (i) uncertainties in groundwater recharge estimates, (ii) spatial resolution of the models, (iii) model choices regarding model parameterization, and (iv) conceptual choices in model implementation (e.g. subsurface layering and assigned permeabilities; see also table S1).

Groundwater recharge estimates are highly uncertain (Reinecke *et al* 2021, Berghuijs *et al* 2022, West *et al* 2022, Gnann *et al* 2023), and their evaluation is challenging due to sparse observations associated with significant biases and uncertainties (MacDonald *et al* 2021). Regardless of these problems, there are distinct differences between the models. For example, West *et al* (2022) shows distinctly different recharge patterns for PCR-GlobWB (Verkaik/ de Graaf) compared to recharge estimates from other models for very dry regions. However, this generally higher recharge does not translate into higher water tables as de Graaf and Verkaik is generally deeper than the ensemble (figure 2) and shallower in arid regions (figure 3). This contradicts the expectation that more recharge should lead to lower water tables (figure 3, table 1). It is likely that other parameters, such as surface water elevation, impact the model results more, which is consistent with findings in Reinecke *et al* (2019a).

The original spatial resolution of Reinecke and de Graaf is similar (5 arcmin), whereas Verkaik and Fan use a higher resolution (30 arcsec; see also table S1). Given that the Verkaik model is, in principle, a higher resolution version of the model by de Graaf, comparing these two models indicates the impact of resolution on WTD (see also (Krakauer *et al* 2014, Reinecke *et al* 2020). We find that aggregating to lower resolution has little effect on overall patterns of WTD (see S13 and S14), suggesting that model structure and forcing inputs might be more important than resolution (if no human impacts are considered).

Regarding other parametrization choices, Reinecke has a lower elevation of surface water bodies (at the 30th percentile of the higher resolution DEM) than de Graaf (mean of the DEM; table S1), which does not explain the generally deeper water table in the latter. Related to this, the static existence of surface water bodies in a steady-state simulation might lead to an overestimation of focused recharge through surface water bodies, as seasonal waterbodies usually exist continuously in the steady-state simulation (as there is no change in time). This issue is nonexistent in the Fan model as it does not specifically model surface water bodies but completely drains groundwater close to the surface (table S1). Other choices regarding the parameterization, such as the faster decrease of permeability with steeper slopes, seem to lead to shallower water tables in mountainous regions in Fan; however, similar approaches (but over fewer conceptual model layers; see table S1) are also implemented in the other models.

Lastly, one would expect that the representation of confined settings leads to a better representation which is implemented in de Graaf in some regions (or more specifically de Graaf *et al* 2017) and Verkaik, but not in Reinecke and Fan. A refinement of the vertical representation only exists in de Graaf and Verkaik who developed a calibrated global map of aquifer thickness for their models. Still, both models show a similarly strong relation to slope as Reinecke, which does not implement such an approach (table S1), thus it is unclear whether an improved representation can improve upon the WTD-Slope relationship.

Additional model comparison through structured intercomparison, such as harmonizing the recharge inputs across the models, is necessary to draw further conclusions. Furthermore, future development efforts should investigate the representation of complex hydrogeologies, such as karst systems, and the impacts of surface water body parameterization (regarding conductance and elevation).

3.5. Different global estimates of WTD may lead to different assessments

Even though uncertainties in areas with shallow water tables are necessarily small (figure 1), they are large enough to have major implications on the outcomes of global assessments of groundwater accessibility. Figure 4 translates the uncertainty in WTD into uncertainty of potential groundwater accessibility for forests, population, and areas equipped for irrigation (note that these classes are not mutually exclusive and are different from the defined potential accessibility classes). It shows that the uncertainty is high for all three classes (forest, population, irrigation) and all three categories of potential accessibility. We find that the global area covered by forest located in regions with ecosystem-accessible groundwater ($<$10 m) varies by +−12 million km^2 (global forest area is 30.5 million km^2) depending on what model estimate we use. How many people live in areas of potentially human-accessible groundwater ($>$10 m and $<$100 m) varies by +−0.9 billion, and the uncertainty of how much irrigated land is in areas of potential ecosystem accessible ($<$10 m) groundwater is +−66 billion ha (660 million km^2, global area is 305.7 million ha). We do not suggest that forests, people, or agriculture necessarily depend on groundwater in these areas, but it highlights the potential lack of robustness of any subsequent application of these simulations without considering these uncertainties.

This large uncertainty directly affects our ability to provide critical global assessments and support decision-making. For example, assessing the likelihood of ecosystems losing connection to groundwater is pivotal for carbon policy and ecosystem protection (Huggins *et al* 2023, Saccò *et al* 2023). Mapping these potentially fragile ecosystems would indicate where ecosystem protection policy would provide the biggest impact (Huggins *et al* 2023). Without improving our ability to support such decisions, we are ultimately jeopardizing our ability to achieve multiple SDGs such as climate action (SDG 13) and terrestrial ecosystems (SDG 15).

The uncertainty shown here affects scientific insights we can gain with these models and it contextualizes existing studies. It is the first time multiple global groundwater models are analyzed and compared directly, and thus provides an important

Table 1. A list of hypotheses for model differences, evidence for or against these hypotheses, and possible reasons.

Hypothesis	Conclusions	Evidence	Possible reasons
More groundwater **recharge leads to shallower water tables.**	Not supported by analysis	The mean recharge of PCR-GlobWB (used in de Graaf) is higher than in WaterGAP (used in Reinecke), but water tables are generally deeper in de Graaf than in Reinecke.	Other factors, such as the parameterization of the aquifer geometry, hydraulic conductivity, and surface water interaction, have a stronger influence on the model results (supported partially by Reinecke *et al* (2019a)). Reinecke and deGraaf use an average recharge from different periods (100 year mean vs. 50 year mean).
Higher spatial **resolution leads to a better representation of WTD-slope relationships** because it leads to a more realistic representation of topography.	Not supported by analysis	The global relationships between WTD and slope in the de Graaf and Verkaik models are the same.	Model structure, forcings, and numerics (Herrera *et al* 2023) may have a larger influence (e.g. groundwater recharge of PCR-GlobWB is still in a coarse scale resolution).
Fixed steady-state heads of surface water bodies lead to shallower water tables in arid regions.	Partially supported	Reinecke and de Graaf show shallower water tables in arid regions. Both simulate water bodies with fixed heads in the steady-state. Verkaik also, but does not show this effect (possibly this is related to the model resolution as well as Verkaik simulates on 1 km). Fan does not simulate any surface water bodies.	When the groundwater table drops below the bottom of a surface water body, MODFLOW-based models simulate water leakage from the surface water body to the groundwater. In a steady-state this flux is constant as the head of the water body does not change, likely leading to higher simulated groundwater levels.
A **decrease** in **permeability with slopes** steepness **leads to shallower water tables**	Supported	Compared to all other models, Fan includes this assumption in the model and shows shallower water tables in mountainous regions compared to other models.	A decrease in permeability leads to a lower hydraulic conductivity and, thus, a decrease in lateral flow even with high gradients. At coarse resolutions (5′), gradients in steep regions can be very large and likely induce unrealistic lateral fluxes.
Confining layers are important to accurately represent WTD, even at large spatial scales.	Inconclusive	Fan does not account for confining layers compared to de Graaf but still shows a closer match with observations.	The observations are likely biased towards shallower water tables. Available global datasets do not indicate whether an aquifer is confined or unconfined. It is unclear if a better representation would directly affect the WTD relationships discussed here.
A refined vertical representation of the hydrogeology improves the simulation of **WTD.**	Inconclusive	De Graaf implements a ‘calibrated’ aquifer thickness. However, the relationships investigated here are similar to Reinecke, which does not implement a varying aquifer thickness.	Currently no other global dataset besides the two-layer GLHYMPS 2.0 datasets (Huscroft *et al* 2018) is available to scrutinize the parameterization further. It is further unclear if a better representation would directly affect the WTD relationships discussed here.

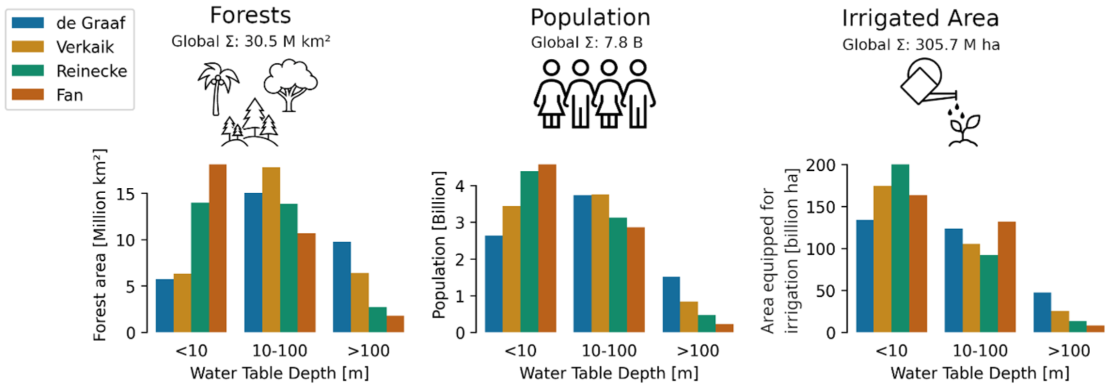

Figure 4. Uncertainties associated with model choice affecting forest area, population, and area equipped for irrigation (AEI) with respect to uncertain regions of three depths of groundwater. Each plot quantifies the uncertainty of how much forest, population, or irrigation is potentially located in areas of ecosystem (<10 m), human (10–100 m), or costly (>100 m) accessible groundwater for each model. The global total of each category is given at the top of each category.

foundation for subsequent research. Previous studies focused on individual global groundwater models and applied sensitivity experiments to demonstrate how their conclusions are affected by key uncertainties (de Graaf *et al* 2019, Reinecke *et al* 2019a). However, a single model experiment likely did not cover the full uncertainty range. Our results suggest that future analyzes would benefit from quantifying and attributing uncertainty more widely using multiple models. All in all, this study also adds to existing evidence that assessing model differences is much easier than attributing these differences to their causes, given the multitude of modeling choices and assumptions that make models dissimilar (e.g. Ceola *et al* 2015).

4. Conclusion

Groundwater is a pivotal source of freshwater for terrestrial and aquatic ecosystems. It functions as an ecosystem, provides drinking water to humans, and remains a reliable source for irrigation during dry periods. To assess how global change (e.g. climate change, land use change, water abstractions) affects the water cycle and potentially freshwater sustainability, we require a global perspective on groundwater. Our analysis shows, however, that current global groundwater models are still in their infancy with significant differences in their simulations of global distributions of depth to groundwater. Here, we focused on steady-state simulations frequently used in existing studies and datasets. Future studies should consider using the ensemble provided here to reflect this uncertainty, bearing in mind that this dataset ignores the considerable impacts of groundwater pumping.

Our results provide first insights into model differences and their potential reasons. They invite a more in-depth investigation to understand and explain inter-model and model-observation differences in the future. Such a comparison would greatly benefit from a structured Model-Intercomparison Project (MIP) specifically focused on groundwater, comparable to the Inter-Sectoral Impact MIP (Frieler *et al* 2017), to provide a consistent framework for model simulations (e.g. standardized forcing data, output resolution, and variable names) (Condon *et al* 2021, Gleeson *et al* 2021). While steady-state disagreements already suggest pathways for improvement, future work needs to investigate transient differences which are central in understanding how abstractions and climate variability affect the model results.

With improved global models (including better representation of human impacts), we would be able to more robustly assess impacts of climate change and human activities on global groundwater resources, filling a current gap in the IPCC reports (IPCC 2022). Information on where groundwater is accessible, abstracted, and potentially remains accessible for future irrigation will enable international organizations like the FAO and the World Bank to guide programs on irrigation infrastructure and crop adaptation. To reach these goals, we need to acknowledge that current global-scale groundwater models must be improved, evaluated more thoroughly, and the scientific community needs to work jointly to compile existing local knowledge into global knowledge shared across communities.

Data availability statement

The ensemble mean on 5 arcmin resolution including the uncertainty bounds can be accessed here: https://doi.org/10.5281/zenodo.7538161 CCby4.0.

The source code of the modeling framework of Reinecke *et al* (2019b) can be accessed at:

http://globalgroundwatermodel.org

CHELSA data are available from

https://envicloud.wsl.ch/#/?prefix=chelsa%2Fchelsa_V2%2FGLOBAL%2F

Geomorpho90m data are available from https://doi.pangaea.de/10.1594/PANGAEA.899135

Global DEM derivatives at 250 m based on the MERIT DEM are available from https://doi.org/10.5281/zenodo.1447209

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5281/zenodo.7538161.

Acknowledgment

RR, SJG, LS, and TW acknowledge support from the Alexander von Humboldt Foundation in the framework of the Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research (BMBF). IdG acknowledges funding from the European Research Council (ERC) under the grant agreement GROW-101041110. We thank Ying Fan Reinfelder for her helpful remarks on an early draft.

Author contributions

RR led the analyses and writing of the manuscript. TW and RR conceived the idea, SG and LS supported the analysis and development of the manuscript. SG led the analysis presented in figure 3. All authors reviewed the manuscript and provided suggestions on text and figures.

ORCID iDs

Robert Reinecke https://orcid.org/0000-0001-5699-8584
Sebastian Gnann https://orcid.org/0000-0002-9797-5204
Lina Stein https://orcid.org/0000-0002-9539-9549
Marc Bierkens https://orcid.org/0000-0002-7411-6562
Inge de Graaf https://orcid.org/0000-0001-7748-868X
Gualbert Oude Essink https://orcid.org/0000-0003-0931-6944
Claudia Ruz Vargas https://orcid.org/0000-0002-4698-3685
Jarno Verkaik https://orcid.org/0000-0001-7420-8304
Thorsten Wagener https://orcid.org/0000-0003-3881-5849

References

Amatulli G, McInerney D, Sethi T, Strobl P and Domisch S 2020 Geomorpho90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers *Sci. Data* **7** 1–18

Baird A J and Low R G 2022 The water table: its conceptual basis, its measurement and its usefulness as a hydrological variable *Hydrol. Process.* **36** e14622

Berghuijs W R, Luijendijk E, Moeck C, van der Velde Y and Allen S T 2022 Global recharge data set indicates strengthened groundwater connection to surface fluxes *Geophys. Res. Lett.* **49** e2022GL099010

Beven K J and Cloke H L 2012 Comment on "Hyperresolution global land surface modeling: meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F Wood *et al* *Water Resour. Res.* **48** W01801

Bierkens M, Graaf I E D, Lips S, Perrone D, Reinhard A J, Jasechko S, van der Himst T and van Beek R 2022 Global economic limits of groundwater when used as a last resort for irrigation *Preprint* (https://doi.org/10.21203/rs.3.rs-1874539/v1) (Accessed 29 July 2022)

Brinkerhoff C B, Gleason C J, Kotchen M J, Kysar D A and Raymond P A 2024 Ephemeral stream water contributions to United States drainage networks *Science* **384** 1476–82

Ceola S *et al* 2015 Virtual laboratories: new opportunities for collaborative water science *Hydrol. Earth Syst. Sci.* **19** 2101–17

Condon L E *et al* 2021 Global groundwater modeling and monitoring: opportunities and challenges *Water Resour. Res.* **57** e2020WR029500

Costa F R C, Schietti J, Stark S C and Smith M N 2023 The other side of tropical forest drought: do shallow water table regions of Amazonia act as large-scale hydrological refugia from drought? *New Phytol.* **237** 714–33

Cuthbert M O, Gleeson T, Moosdorf N, Befus K M, Schneider A, Hartmann J and Lehner B 2019 Global patterns and dynamics of climate-groundwater interactions *Nat. Clim. Change* **9** 137–41

de Graaf I E M, Gleeson T, van Rens Beek L P H, Sutanudjaja E H and Bierkens M F P 2019 Environmental flow limits to global groundwater pumping *Nature* **574** 90–94

de Graaf I E M and Stahl K 2022 A model comparison assessing the importance of lateral groundwater flows at the global scale *Environ. Res. Lett.* **17** 44020

de Graaf I E M, Sutanudjaja E H, van Beek L P H and Bierkens M F P 2015 A high-resolution global-scale groundwater model *Hydrol. Earth Syst. Sci.* **19** 823–37

de Graaf I, van Beek R L, Gleeson T, Moosdorf N, Schmitz O, Sutanudjaja E H and Bierkens M F 2017 A global-scale two-layer transient groundwater model: development and application to groundwater depletion *Adv. Water Resour.* **102** 53–67

Fan Y, Li H and Miguez-Macho G 2013 Global patterns of groundwater table depth *Science* **339** 940–3

Fan Y, Miguez-Macho G, Jobbágy E G, Jackson R B and Otero-Casal C 2017 Hydrologic regulation of plant rooting depth *Proc. Natl Acad. Sci.* **114** 10572–7

Ferguson G, McIntosh J C, Perrone D and Jasechko S 2018 Competition for shrinking window of low salinity groundwater *Environ. Res. Lett.* **13** 114013

Frieler K *et al* 2017 Assessing the impacts of 1.5 °C global warming - simulation protocol of the inter-sectoral impact model intercomparison project (ISIMIP2b) *Geosci. Model Dev.* **10** 4321–45

Gleeson T *et al* 2021 GMD perspective: the quest to improve the evaluation of groundwater representation in continental- to global-scale models *Geosci. Model Dev.* **14** 7545–71

Gleeson T, Befus K M, Jasechko S, Luijendijk E and Cardenas M B 2016 The global volume and distribution of modern groundwater *Nat. Geosci.* **9** 161–7

Gnann S *et al* 2023 Functional relationships reveal differences in the water cycle representation of global water models *Nat. Water* **1** 1079–90

Haitjema H M and Mitchell-Bruker S 2005 Are water tables a subdued replica of the topography? *Groundwater* **43** 781–6

Hansen M C *et al* 2013 High-resolution global maps of 21st-century forest cover change *Science* **342** 850–3

Hare D K, Helton A M, Johnson Z C, Lane J W and Briggs M A 2021 Continental-scale analysis of shallow and deep groundwater contributions to streams *Nat. Commun.* **12** 1450

Hengl T 2018 Global DEM derivatives at 250 m, 1 km and 2 km based on the MERIT DEM *Zenodo*

Herrera P A, Langevin C and Hammond G 2023 Estimation of the water table position in unconfined aquifers with MODFLOW 6 *Groundwater* **61** 648–62

Hill M C and Tiedeman C R 2007 *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty* 1st edn (Wiley)

Huggins X, Gleeson T, Serrano D, Zipper S, Jehn F, Rohde M M, Abell R, Vigerstol K and Hartmann A 2023 Overlooked risks and opportunities in groundwatersheds of the world's protected areas *Nat. Sustain.* **6** 855–64

Humphrey V, Zscheischler J, Ciais P, Gudmundsson L, Sitch S and Seneviratne S I 2018 Sensitivity of atmospheric CO_2 growth rate to observed changes in terrestrial water storage *Nature* **560** 628–31

Huscroft J, Gleeson T, Hartmann J and Börker J 2018 Compiling and mapping global permeability of the unconsolidated and consolidated earth: GLobal HYdrogeology MaPS 2.0 (GLHYMPS 2.0) *Geophys. Res. Lett.* **45** 1897–904

IPCC 2022 *Climate Change 2022: Impacts, Adaptation and Vulnerability: Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press)

Jasechko S, Seybold H, Perrone D, Fan Y and Kirchner J W 2021 Widespread potential loss of streamflow into underlying aquifers across the USA *Nature* **591** 391–5

Jasechko S, Seybold H, Perrone D, Fan Y, Shamsudduha M, Taylor R G, Fallatah O and Kirchner J W 2024 Rapid groundwater decline and some cases of recovery in aquifers globally *Nature* **625** 715–21

Kang M, Ayars J E and Jackson R B 2019 Deep groundwater quality in the southwestern United States *Environ. Res. Lett.* **14** 34004

Kang M and Jackson R B 2016 Salinity of deep groundwater in California: water quantity, quality, and protection *Proc. Natl Acad. Sci. USA* **113** 7768–73

Karger D N, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza R W, Zimmermann N E, Linder H P and Kessler M 2017 Climatologies at high resolution for the earth's land surface areas *Sci. Data* **4** 1–20

Koirala S *et al* 2017 Global distribution of groundwater-vegetation spatial covariation *Geophys. Res. Lett.* **44** 4134–42

Krakauer N Y, Li H and Fan Y 2014 Groundwater flow across spatial scales: importance for climate modeling *Environ. Res. Lett.* **9** 34003

Kroes J, Supit I, van Dam J, van Walsum P and Mulder M 2018 Impact of capillary rise and recirculation on simulated crop yields *Hydrol. Earth Syst. Sci.* **22** 2937–52

Link A, El-Hokayem L, Usman M, Conrad C, Reinecke R, Berger M, Wada Y, Coroama V and Finkbeiner M 2023 Groundwater-dependent ecosystems at risk—global hotspot analysis and implications *Environ. Res. Lett.* **18** 94026

Linke S *et al* 2019 Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution *Sci. Data* **6** 283

MacDonald A M *et al* 2021 Mapping groundwater recharge in Africa from ground observations and implications for water security *Environ. Res. Lett.* **16** 34012

Maxwell R M and Condon L E 2016 Connections between groundwater flow and transpiration partitioning *Science* **353** 377–80

Pelletier J D, Broxton P D, Hazenberg P, Zeng X, Troch P A, Niu G-Y, Williams Z, Brunke M A and Gochis D 2016 A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling *J. Adv. Model. Earth Syst.* **8** 41–65

Perrone D and Jasechko S 2019 Deeper well drilling an unsustainable stopgap to groundwater depletion *Nat. Sustain.* **2** 773–82

Reinecke R *et al* 2021 Uncertainty of simulated groundwater recharge at different global warming levels: a global-scale multi-model ensemble study *Hydrol. Earth Syst. Sci.* **25** 787–810

Reinecke R, Foglia L, Mehl S, Herman J D, Wachholz A, Trautmann T and Döll P 2019a Spatially distributed sensitivity of simulated global groundwater heads and flows to hydraulic conductivity, groundwater recharge, and surface water body parameterization *Hydrol. Earth Syst. Sci.* **23** 4561–82

Reinecke R, Foglia L, Mehl S, Trautmann T, Cáceres D and Döll P 2019b Challenges in developing a global gradient-based groundwater model (G^3M v1.0) for the integration into a global hydrological model *Geosci. Model Dev.* **12** 2401–18

Reinecke R, Wachholz A, Mehl S, Foglia L, Niemann C and Döll P 2020 Importance of spatial resolution in global groundwater modeling *Groundwater* **58** 363–76

Roebroek C T J, Melsen L A, Hoek van Dijke A J, Fan Y and Teuling A J 2020 Global distribution of hydrologic controls on forest growth *Hydrol. Earth Syst. Sci.* **24** 4625–39

Saccò M *et al* 2023 Groundwater is a hidden global keystone ecosystem *Glob. Change Biol.* **30** e17066

Siebert S, Kummu M, Porkka M, Döll P, Ramankutty N and Scanlon B R 2015 A global data set of the extent of irrigated land from 1900 to 2005 *Hydrol. Earth Syst. Sci.* **19** 1521–45

Taylor R G *et al* 2013 Ground water and climate change *Nat. Clim. Change* **3** 322–9

United Nations 2015 *Water for a sustainable world (The United nations world water development report)* vol 2015 (UNESCO)

United Nations 2022 *The United Nations World Water Development Report 2022: groundwater: making the invisible visible* (UNESCO)

Verkaik J, Sutanudjaja E H, Oude Essink G H P, Lin H X and Bierkens M F P 2024 GLOBGM v1.0: a parallel implementation of a 30 arcsec PCR-GLOBWB-MODFLOW global-scale groundwater model *Geosci. Model Dev.* **17** 275–300

Wagener T, Gleeson T, Coxon G, Hartmann A, Howden N, Pianosi F, Rahman M, Rosolem R, Stein L and Woods R 2021 On doing hydrology with dragons: realizing the value of perceptual models and knowledge accumulation *WIREs Water* **8** e1550

Wagener T, Reinecke R and Pianosi F 2022 On the evaluation of climate change impact models *WIREs Clim. Change* **13** e772

Wang-Erlandsson L *et al* 2022 A planetary boundary for green water *Nat. Rev. Earth Environ.* **3** 380–92

West C, Reinecke R, Rosolem R, MacDonald A M, Cuthbert M O and Wagener T 2022 Ground truthing global-scale model estimates of groundwater recharge across Africa *Sci. Total Environ.* **858** 159765

Winter T C 1995 Recent advances in understanding the interaction of groundwater and surface water *Rev. Geophys.* **33** 985–94

Yamazaki D, Ikeshima D, Tawatari R, Yamaguchi T, O'Loughlin F, Neal J C, Sampson C C, Kanae S and Bates P D 2017 A high-accuracy map of global terrain elevations *Geophys. Res. Lett.* **44** 5844–53