



Global hyper-resolution modeling of historical and future groundwater dynamics

Barry van Jaarsveld¹, Niko Wanders¹, Nicole Gyakowah Otoo¹, Edwin H. Sutanudjaja¹, Jarno Verkaik², Daniel Zamrsky^{1,3}, and Marc F.P. Bierkens¹

¹Utrecht University, Department of Physical Geography, Princeton laan 8a, Utrecht, The Netherlands

²Deltares, Groundwater and Water Security, Utrecht, The Netherlands

³IHE Delft Institute for Water Education, Department of Water Resources and Ecosystems, Delft, The Netherlands

Correspondence: Barry van Jaarsveld (a.s.vanjaarsveld@uu.nl)

Abstract. The sustainable management of global groundwater resources is a key societal challenge and is central to the Sustainable Development Goals. The localized dynamics of groundwater abstraction, topography, and surface-water interactions, as well as the sensitivity of groundwater-dependent ecosystems, call for high-resolution information to support effective groundwater management. At the same time, groundwater observations are very limited and concentrated in a few regions, rendering large parts of groundwater resources ungauged. To address limited observations and coarse global models, we applied the global groundwater model GLOBGM (v1.1) to simulate past and future groundwater heads and water table depth at 30 arc-seconds (~1 km) on a monthly time step. Model calibration improved mean bias in water table depth predictions from -4.8 m to 3.6 m compared to GLOBGM v1.0, with depth-weighted bias reduced from 34.2 m to 32.5 m across 34 800 observation wells. Groundwater dynamics are simulated for a historical reference period (1960–2019) to support model evaluation and attribution of observed impacts to climate variability and change. Baselines (1960–2014) and three combined socioeconomic-climate scenarios (2015–2100; SSP1-RCP2.6, SSP3-RCP7.0, SSP5-RCP8.5) are simulated with five global climate models, supporting detection and impact assessment of future change. Validation against monthly observations yielded skillful predictions (KGE-NP > -0.41) in approximately 75 % of deep wells (>60 m) and 90 % of shallow to intermediate wells (0–20 m). When validated against annual observations, approximately 80 % of sites showed skillful predictions regardless of depth. Historical trend analysis (1960–2019) accurately reproduced known groundwater depletion regions such as the U.S. High Plains, Arabian Peninsula, and Indo-Gangetic Plain, while also identifying rising water table depths in northern latitudes and Arctic regions potentially linked to climate-driven recharge changes. Future scenario-based simulations suggest rising water table depths for most continents in the next century, with Europe being a notable exception. However, known regions of groundwater depletion are expected to persist. Regions of reduced reliability are mapped, and quality assurance flags are provided to guide the appropriate use and interpretation of the results. The resulting data set offers high-resolution information to assess groundwater dynamics for the past and future, supporting improved global water resource management and climate impact assessments.



1 Introduction

Groundwater is the most abundant source of fresh liquid water; globally, 98 % of accessible fresh liquid water is in the form of groundwater (Döll et al., 2012; Gleeson et al., 2020). People increasingly rely on these resources to supplement surface water abstractions to meet total water demands (Custodio, 2010). Groundwater is vulnerable to overexploitation when abstraction rates persistently exceed natural groundwater recharge — the process by which surface water, largely from precipitation, infiltrates the soil to replenish aquifers. Driven by population growth and associated increases in water demand for agriculture, domestic, and industrial purposes, groundwater abstraction has increased significantly in recent decades, ultimately causing a decline in groundwater storage in many regions (Bierkens and Wada, 2019; Niazi et al., 2024; Wada and Bierkens, 2014; Kuang et al., 2024; Rodell et al., 2018; Döll et al., 2012; Leijnse et al., 2024). Beyond anthropogenic pressures, groundwater is sensitive to decadal and annual climate variability and its influence on groundwater recharge (Taylor et al., 2013; Moeck et al., 2020b). These effects are complex, vary geographically, and are mediated by interactions between topography, vegetation, underlying geological properties, and the interactions between surface water and groundwater (Moeck et al., 2020b; Maples et al., 2019). For example, groundwater in deep confined aquifers with low permeability is relatively insensitive to changes in short-term recharge dynamics, whereas groundwater in highly permeable shallow aquifers with vegetation characteristics that facilitate exchanges across the land-atmosphere continuum is more tightly coupled to recharge variability (Ilstedt et al., 2016; West et al., 2022; Maples et al., 2019). Consequently, the future of groundwater storage depends on complex interactions among human water demands, climate, and ecosystems (Vörösmarty et al., 2000; Wada and Bierkens, 2014).

Overexploitation of groundwater causes water table depths to decline, a consequence of groundwater abstraction evident in many regions, especially in arid and dryland areas (Jasechko and Perrone, 2021; Jasechko et al., 2024). This reduces groundwater's reliability as a water source, given the physical and economic constraints associated with abstracting deeper reserves (Perrone and Jasechko, 2019). Since half of global river flows are groundwater-sustained, deepening water tables threaten downstream surface water supplies (De Graaf et al., 2024; Xie et al., 2024; Bierkens et al., 2021) and may increase vulnerability to hydrological drought (Wada et al., 2013; Vega-Briones et al., 2024). Beyond these supply constraints, reduced reserves aggravate land subsidence, accelerate sea level rise, and degrade groundwater quality (Kuang et al., 2024; Bierkens and Wada, 2019). Deeper water table depths also influence the composition and functioning of groundwater-dependent ecosystems and the services they provide (Saccò et al., 2024; Fan et al., 2017; Otoo et al., 2025). Ultimately, groundwater forms an important component of global socio-ecological systems and the Earth system as a whole (Huggins et al., 2023).

Sustainable management of the world's remaining groundwater reserves is a key societal challenge and is cited in several Sustainable Development Goals (Garriek et al., 2017; UN Water, 2022). Policymakers must implement adaptation and mitigation strategies to secure water availability by maintaining groundwater levels within safe operating limits. However, this requires reliable estimates of current and future groundwater volumes. Relying solely on observed groundwater data is insufficient given the severe limitations in global spatial and temporal coverage (Condon et al., 2021; UN Water, 2022). For example, the current state-of-the-art dataset contains 180 000 groundwater time series from 41 countries, with over 90 % of these time series occurring in North America, Australia, or Europe (Bäthge et al., 2025). To overcome this challenge and provide



global-scale information, water managers rely on physically based numerical global groundwater models (Condon et al., 2021). However, until now, most global groundwater flow models have had limited spatial extent, allowing regional-scale analyzes ($\leq 1\,000\,000\text{ km}^2$) at most (Zamrsky et al., 2025).

Groundwater models of global extent were first formulated in the past decade (Fan et al., 2013), with more recent efforts enabling transient simulations (Reinecke et al., 2019; Costantini et al., 2023; de Graaf et al., 2019). However, these models provide information at spatial resolutions that are often too coarse to meet the demands of policymakers and stakeholders, who require finer-scaled information (Condon et al., 2021; Bierkens et al., 2015). To overcome this limitation, spatial resolution must be enhanced to make global models more locally relevant and actionable (Bierkens et al., 2015; Wood et al., 2011). To this end, hyper-resolution groundwater models have been highlighted for their potential to include climate, vegetation, and anthropogenic influences on groundwater at appropriate spatial resolutions for past and future (Bierkens et al., 2015). This is especially relevant, as hyper-resolution models will not only enable understanding of current water scarcity hotspots but also project future hotspots and enable planning of effective adaptation strategies (Leijnse et al., 2024; Bierkens, 2015). Recent work demonstrates that global groundwater models at $\sim 1\text{ km}$ spatial resolution are possible with today's high performance computers (Verkaik et al., 2024; Kollet et al., 2026). This technical development offers the opportunity to generate global hyper-resolution estimates of past and future groundwater reserves at scales relevant to policymakers.

This study aims to produce the first hyper-resolution monthly estimates of global groundwater storage, incorporating the impacts of climate change and human activities across past and future periods. To this end, using scenarios and global climate models (GCMs) from the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP), monthly groundwater heads and water table depths are simulated at 30 arc-seconds ($\sim 1\text{ km}$ at the equator). We use the ISIMIP3a climate forcing to simulate groundwater reserves for the historical reference period (1960–2019), which can be used for detection and attribution of past climate and socioeconomic change impacts on groundwater reserves. To project future trends, we employ ISIMIP3b inputs from a suite of five GCMs, covering both the baseline (1960–2014) and future periods (2015–2100) under combined socio-economic and climate change scenarios (SSP1-RCP2.6, SSP3-RCP7.0, and SSP5-RCP8.5). These simulations are analyzed by calculating long-term trends, providing a global initial assessment of historical shifts in the global groundwater system and how it may change in the future.

2 Methods

2.1 The GLOBGM v1.1 model

Monthly estimates of groundwater heads and water table depths were simulated using GLOBGM v1.1, a global 30 arc-seconds, two-layer transient, MODFLOW-based groundwater flow model. GLOBGM is the successor to the 5 arc-minute version originally developed by de Graaf et al. (2017). For brevity, we refer to Verkaik et al. (2024) for a detailed description of model inputs, parametrization, and runtime configurations. Briefly, GLOBGM is based on a prototype version of MODFLOW6, which is parallelised using a message-passing interface and makes use of unstructured grids. Independent unstructured grids are distributed over three continental-scale groundwater models — Afro-Eurasia, the Americas, Australia — and one remain-



ing model for the islands (Fig. 1). Each of the four groundwater models is partitioned into non-overlapping sub-models that
 90 are coupled within the MODFLOW linear solver. The two-layer configuration represents unconfined, confining, or confined
 aquifer systems (Verkaik et al., 2024). For cells where a confining layer is present, the top model layer represents the confining
 layer, and the bottom layer is the confined aquifer; in regions where a confining layer is absent, the bottom model layer repre-
 sents the unconfined aquifer. GLOBGM employs an offline coupling approach whereby 5 arc-minutes resolution local runoff,
 groundwater abstraction, and groundwater recharge from PCR-GLOBWB2 (Sutanudjaja et al., 2018) are resampled to match
 95 GLOBGM's 30 arc-seconds grid.

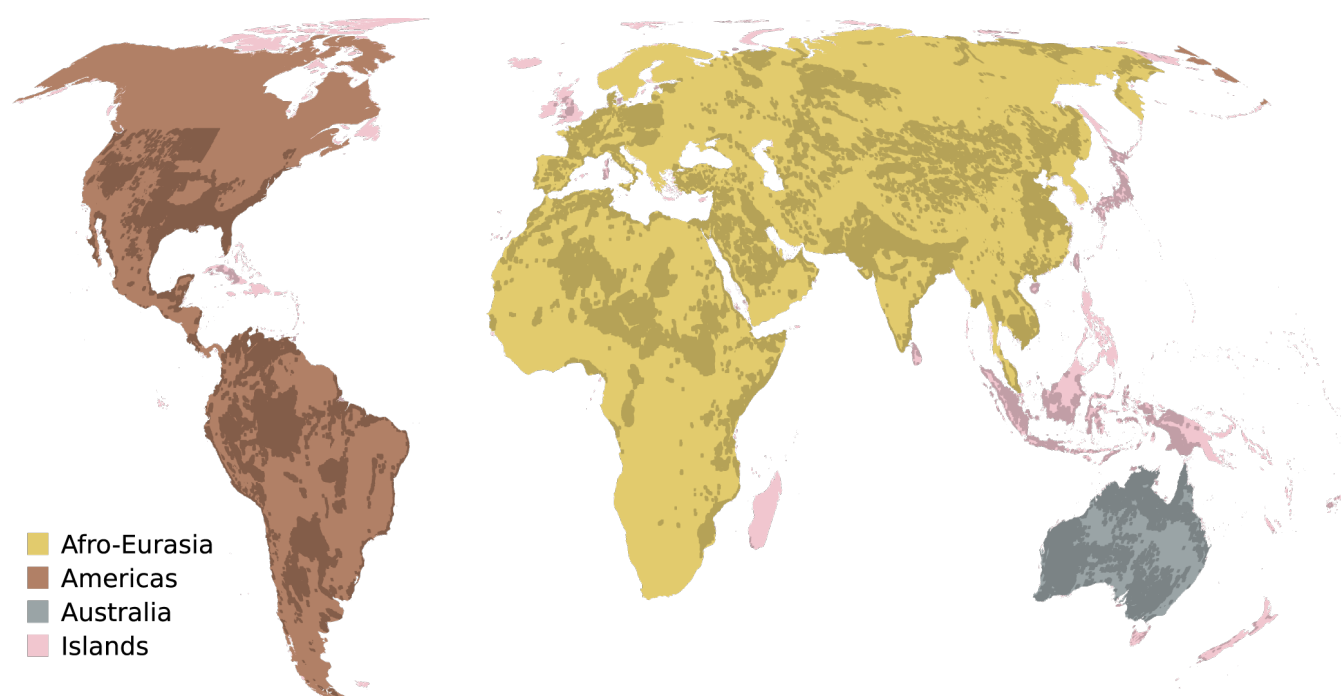


Figure 1. The four groundwater models that constitute GLOBGM, a global groundwater model. Dark overlays indicate the presence of a confining layer.

GLOBGM v1.1 builds upon GLOBGM v1.0 (Verkaik et al., 2024), and in contrast to GLOBGM v1.0, a dynamic drainage
 elevation routine was implemented to better represent permanent wetland areas. To address a major source of uncertainty in
 global modelling, the updated version also incorporates a new downscaling method for groundwater recharge using observed
 data (Reinecke et al., 2024; Gnann et al., 2023; Bierkens and Wada, 2019). The treatment of discharge and abstraction is
 100 consistent, whereby local runoff from PCR-GLOBWB2 was converted to surface water discharge by accumulating it through
 the 30 arc-seconds HydroSHEDS river drainage network (Lehner et al., 2008), and groundwater abstraction was resampled to
 30 arc-seconds using a nearest neighbor algorithm.



2.1.1 Dynamic Drainage Elevation

In the dynamic drainage elevation routine, each cell's drain elevation adjusts based on soil saturation. As soils become wetter, the drainage elevation rises closer to the surface; as they dry, it lowers toward the base elevation. Discharge occurs only when the hydraulic head exceeds this dynamic drain level (Eq. 1). By coupling drain elevation to saturated area fractions, this approach maintains groundwater levels closer to the surface in wetland areas, thereby reducing unrealistic head fluctuations and highly intermittent discharge events. To simulate variations in the saturated area fraction, the improved ARNO scheme (Todini, 1996; Hagemann and Gates, 2003), an integral part of PCR-GLOBWB2, identifies areas subject to surface runoff.

$$f_{sat} = 1 - \left(\frac{W_{max} - W}{W_{max} - W_{min}} \right)^{\frac{\beta}{\beta+1}} \quad (1)$$

where W_{max} is the maximum water storage, W_{min} is the minimum water storage, W is the actual water storage, β is a dimensionless shape parameter in PCR-GLOBWB2.

2.1.2 Groundwater Recharge Downscaling

Accurately predicting groundwater recharge remains a challenge for current global hydrological models (Reinecke et al., 2024; Gnann et al., 2023; Bierkens and Wada, 2019). Previous studies have shown that PCR-GLOBWB2, to which GLOBGM is coupled, tends to underestimate recharge rates compared to observations (Berghuijs et al., 2024). The bias is most pronounced in arid or semi-arid regions where recharge rates are low, intermittent, and unpredictable (Berghuijs et al., 2024). Furthermore, GLOBGM v1.0 predicts groundwater heads that are generally lower than *in situ* observations and other large-scale groundwater models (Reinecke et al., 2024; Verkaik et al., 2024). To address these challenges, recharge supplied to GLOBGM v1.1 was downscaled and corrected using observed groundwater recharge data.

A Generalized Additive regression approach using pyGam (Servén et al., 2018) was employed to produce a long-term average 30 arc-seconds recharge field ($GWR_{observed}$) from known drivers and covariates (Table 1). This approach has previously been applied in groundwater modeling contexts (Mohan et al., 2018; Zamrsky et al., 2025). In contrast to these studies, land cover was not explicitly included as a predictor here; instead, the 5 arc-minute PCR-GLOBWB2 recharge was included to capture land cover-associated residuals by proxy. To prevent unrealistic recharge rates, $GWR_{observed}$ was further processed to ensure it did not exceed local precipitation rates (Brun et al., 2022b, a). Finally, to calculate a correction factor (GWR_{cf}), $GWR_{observed}$ was compared with PCR-GLOBWB2 5 arc-minute recharge (bilinearly interpolated to 30 arc-seconds) from a GSWP3-W5E5 forced naturalized run (Eq. 2). For the steady-state simulation, the 5 arc-minute PCR-GLOBWB2 recharge was interpolated to the 30 arc-second GLOBGM v1.1 grid via bicubic interpolation and corrected (Eq. 3). For the transient simulations, for each time step, PCR-GLOBWB2 was interpolated and corrected (Eq. 3) and, to avoid unrealistic estimates, recharge values were constrained by local precipitation rates (Eq. 4).

$$GWR_{cf} = \frac{GWR_{observed}}{GWR_{5arc'}} \quad (2)$$



Table 1. Response and predictor variables used for groundwater recharge downscaling.

Type	Variable	Units	Resolution	Reference
Response	Observed Recharge	m year ⁻¹	Point Data	Moeck et al. (2020a)
	Aridity Index	—	30 arc-seconds	Zamrsky et al. (2024)
	Clay, Sand, & Silt Fraction	%	30 arc-seconds	Poggio et al. (2021)
Predictor	Potential Ref. Evaporation	m year ⁻¹	30 arc-seconds	Brun et al. (2022b)
	Precipitation	m year ⁻¹		
	Recession Coefficient	—	30 arc-seconds	Verkaik et al. (2024)
	Slope	—		
	Elevation	m		
	Hydraulic Conductivity	m day ⁻¹		
	PCR-GLOBWB2 Recharge	m year ⁻¹	5 arc-minutes	Sutanudjaja (2024)

where GWR_{observed} is the observed recharge field, $GWR_{5\text{arc}'}$ is the 5 arc-minutes recharge from PCR-GLOBWB2 interpolated to 30 arc-seconds.

$$135 \quad GWR_{\text{corrected}} = GWR_{5\text{arc}'} \times GWR_{\text{cf}} \quad (3)$$

where GWR_{cf} is the correction field and $GWR_{5\text{arc}'}$ is the 5 arc-minutes recharge from PCR-GLOBWB2 interpolated to 30 arc-seconds.

$$GWR_{\text{corrected}} = \begin{cases} P_{\text{month}} & \text{if } GWR_{\text{corrected}} > P_{\text{month}} \\ GWR_{\text{corrected}} & \text{if } GWR_{\text{corrected}} \leq P_{\text{month}} \end{cases} \quad (4)$$

where $GWR_{\text{corrected}}$ is the corrected recharge field and P_{month} is the local precipitation for the month.

140 2.2 Calibration, spin-up and initial conditions

GLOBGM v1.0 was uncalibrated, and initial conditions were calculated using a method inherited from its predecessor, PCR-GLOBWB2 (Verkaik et al., 2024; de Graaf et al., 2019). This process involved simulating steady-state groundwater heads under naturalized conditions, followed by a transient spin-up period in which the first year of forcing was repeated back-to-back. However, uncalibrated models prove less accurate than calibrated counterparts, such as the CONUS groundwater and global-scale inverse models (Verkaik et al., 2024; Zell and Sanford, 2020; Fan et al., 2013). Furthermore, spurious trends were evident in GLOBGM v1.0 simulations (Verkaik et al., 2024). To address these limitations, GLOBGM v1.1 presented here was calibrated, and an alternative spin-up procedure was implemented.



2.2.1 Calibration

To improve the accuracy of groundwater heads and water table depths simulated by GLOBGM v1.1, hydraulic conductivity, anisotropy, and entrance resistance parameters were calibrated using the steady-state solution. For anisotropy and entrance resistance, a range of global prefactor adjustments was assessed (Table 2). While a broader range of prefactors was possible, adjustments were limited to within two orders of magnitude. This constraint prevented significant departures from the original model parameterization, which was necessary given the spatially concentrated distribution of observations. Previous calibration attempts on the 5 arc-minute predecessor reported limited sensitivity to variations in hydraulic conductivity when applied as a global prefactor to the entire domain (Otoo et al., 2025). Therefore, to incorporate regional sensitivities to changes in hydraulic conductivity, different combinations of prefactors were assessed (Table 2), which varied by lithological class (Hartmann and Moosdorf, 2012). In total, 162 permutations were assessed.

Table 2. Parameters and prefactors used during calibration. Lithologies are from the GLiM dataset (Hartmann and Moosdorf, 2012).

Parameter	Lithology Class	Candidate values
Ks	Fine & coarse unconsolidated	0.1–1.0–10.0
	Fine & coarse	0.1–1.0–10.0
	Carbonate	0.1–1.0–10.0
Anisotropy	Consistent across all lithological classes	0.1–1.0
Entrance resistance	Consistent across all lithological classes	0.1–1.0–10.0

Candidate calibration values were assessed by running the steady-state solution for the historical reference simulation (Section 2.3) under pristine natural conditions (i.e., unaltered by anthropogenic influence on surface water and groundwater dynamics). Note that updates regarding groundwater recharge downscaling and saturated area fraction adjustments were included prior to calibration (Section 2.1.1 & 2.1.2). To select optimal calibration settings, long-term averaged groundwater table depths from observations that overlapped with the forcing period (i.e., 1960–2019) were used, obtained from IGRAC’s Global Groundwater Monitoring Network (IGRAC, 2024). Given that the steady-state solution of GLOBGM v1.1 was forced with data representative of the hydrological cycle in a pristine state, observations were filtered to exclude regions with substantial groundwater abstraction. Specifically, observations were filtered to exclude known hotspots for water scarcity (Lange and Büchner, 2021), as these regions likely experience extensive groundwater abstraction. Using the filtered dataset of 34 800 observation wells, calibration permutations were evaluated by the mean bias over all locations (Bias_{WTD} ; Eq. 5).

$$\text{Bias}_{\text{WTD}} = \text{WTD}_{\text{obs}} - \text{WTD}_{\text{sim}} \quad (5)$$

where WTD_{sim} simulated water table depths and WTD_{obs} the observed water table depths.

To enable a direct comparison between model simulations and observations, we used the water table depth outputs from GLOBGM v1.1, derived from heads and surface elevation. Note that due to this conversion and Eq. 5, a positive bias indicates that simulated water table depths are shallower than observed, whereas a negative bias indicates that they are deeper. For



unconfined aquifers, where no confining layer is present, observations were assigned to the bottom model layer. Where a confining layer was present in the model, observations were assigned to the correct model layer based on the reported filter. If filter depth was not reported, model layers were assigned using an approach that evaluates the correlation between remotely sensed soil moisture data and simulated groundwater heads, as done by Verkaik et al. (2024). More specifically, if the observed groundwater timeseries, when expressed as groundwater heads, correlated well (> 0.5) with soil moisture (ESA CCI SM; Dorigo et al. (2017)), it was assigned to the upper model layer (i.e., confining aquifer); if no correlation was found, it was assigned to the bottom model layer (i.e., confined aquifer).

2.2.2 Spin-up and initial conditions

We aimed to derive a set of initial groundwater head states for the simulation start year (i.e., 1960), such that initial conditions resemble the observed data available for 1960. First, the steady-state estimates of groundwater heads under pristine natural conditions were obtained. These steady-state results served as the starting point for a transient simulation for the year 1960, without abstraction, for 75 iterations back-to-back. Subsequently, a second transient simulation for the year 1960, with abstraction activated, was conducted for 50 iterations, allowing simulated groundwater levels to approach observed levels. GLOBGM v1.1 was forced with data relevant to the historical reference simulation described in section 2.3 for this spin-up protocol. Similar to the calibration methodology, observed water table depth data were obtained from IGRAC's Global Groundwater Monitoring Network. Observations were filtered for the year 1960 (375 observation wells). At the end of each iteration, Bias_{WTD} (Eq. 5) and PBIAS (Eq. 6) were calculated. Simulated and observed water table depths were assigned to the correct model layer as done in section 2.2.1. Our goal was to minimise the difference between modelled and observed groundwater depths of the starting year (i.e., 1960), using the convergence of Bias_{WTD} and PBIAS as indicators of stability. To assess the evolution of head estimates over the model domain, the average groundwater head for each sub-model was calculated ($\bar{h}$; Eq. 7).

$$\text{PBIAS} = \frac{\sum (\text{WTD}_{\text{sim}} - \text{WTD}_{\text{obs}})}{\sum \text{WTD}_{\text{obs}}} \times 100 \quad (6)$$

where WTD_{sim} and WTD_{obs} denote the simulated and observed water table depths, respectively.

$$\bar{h} = \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m h_{i,j} \quad (7)$$

where $h_{i,j}$ groundwater head at grid cell in row i and column j , n and m are the numbers of rows and columns, respectively.

Achieving absolute dynamic equilibrium in a global groundwater model at the hyper-resolution is computationally prohibitive and would require thousands of iterations (Verkaik et al., 2024). Furthermore, given the significant structural uncertainties in global groundwater models, further refinement of the initialization is expected to provide negligible benefit (Reinecke et al., 2024). Therefore, groundwater heads estimated at the final iteration were used as initial conditions for all subsequent simulations presented herein. To account for potential spin-up issues in the final dataset, a procedure was developed to identify locations and time periods where the model is likely not in dynamic equilibrium. Focusing on the historical reference simulation, the valid month for each grid cell was identified, defined as the month after which the model is considered to be in a



spun-up state. Non-equilibrium states were identified using a temporal correlation analysis. This approach relies on the fact that
 205 in areas affected by spin-up artifacts, hydraulic heads undergo a systematic adjustment that overpowers natural groundwater
 dynamics, and the magnitude of this adjustment decreases over time until the spin-up effect dissipates. To quantify this, the
 absolute difference in hydraulic heads between consecutive months was calculated (Eq. 8).

$$\Delta h_t = |h_t - h_{t-1}| \quad (8)$$

where h_t is the hydraulic head (in meters) at time step t , and Δh_t represents the absolute month-to-month change in hydraulic
 210 head.

As a first pass, to filter out locations where groundwater heads are static, if the mean absolute difference of the entire
 simulation was less than 0.001 m, the cell was considered static and marked the entire simulation period as valid. For the
 remaining cells, a 12-month rolling window was applied to compute the mean difference and Pearson correlation coefficient
 between absolute differences and time (Eq. 9). A strong negative correlation indicates that the differences are decreasing,
 215 indicating incomplete spin-up. In contrast, a weak correlation suggests the model responds to dynamic forcing rather than
 initialization artifacts. The start valid date was determined by comparing the correlation coefficient, r_t , against a threshold of
 -0.2. If the correlation in the first window was greater than -0.2, the cell was marked as valid for the entire simulation. However,
 if the initial correlation indicated a strong negative trend, subsequent years were scanned to identify the first instance where r_t
 rose above -0.2, designating the valid simulation time step. To prevent false negatives caused by numerical noise, if the mean
 220 change ($\mu_h = \frac{\mu \Delta h_t}{\tau}$) in head within the window was less than 0.001 m (i.e., the convergence criteria), the window was also
 marked as valid regardless of the correlation value.

$$r_t = \frac{\text{cov}(\Delta h, T)}{\sigma_{\Delta h} \cdot \sigma_T} \quad (9)$$

where r_t is the Pearson correlation coefficient at time step t , T represents the vector of time indices within the rolling window,
 $\text{cov}(\Delta h, T)$ is the covariance between head differences and time, and $\sigma_{\Delta h}$ and σ_T denote the standard deviations of head
 225 differences and time indices, respectively.

2.3 Simulations

Groundwater abstraction, recharge, and discharge from the HYPFLOWSCI6 (HYdrological Projection of Future gLObal Water
 States with CMIP6) dataset was used as forcing for the simulations presented here (Sutanudjaja, 2024; Kuzma et al., 2023). The
 HYPFLOWSCI6 dataset represents the integration of PCR-GLOBWB 2 with CMIP6 climate forcings to project future water
 230 stress under three combined socioeconomic-climate scenarios: SSP1-RCP2.6, SSP3-RCP7.0, and SSP5-RCP8.5 (Sutanudjaja,
 2024; Kuzma et al., 2023). Following the HYPFLOWSCI6 simulation framework, simulations were divided into two parts:
 (1) historical reference simulation and (2) scenario-based simulations. For the historical reference simulation, GSWP3-W5E5
 forcing was used for the period 1960–2019 (Lange et al., 2021). For the scenario-based simulations the baselines (1960–2014)
 and future periods (2015–2100) for three SSP-RCP scenarios and five bias corrected GCM's were used (Table 3; Lange and
 235 Büchner (2021)).



Table 3. Climate forcing datasets and future scenarios for groundwater simulations with GLOBGM v1.1.

Simulation	Forcing	Period	Description
Historical Reference	GSWP3-W5E5 (Lange et al., 2021)	1960–2019	Reanalysis-based reconstruction which allows validation and impact attribution against observed climate records.
	GFDL-ESM4 (Dunne et al., 2020)	1960–2014	Baseline simulation.
Future Scenarios	IPSL-CM6A-LR (Boucher et al., 2020)	2015–2100	SSP1-RCP2.6: Global warming of 2° C, radiative forcing of 2.6 W m ⁻² in 2100, optimistic sustainable development with low barriers to mitigation and adaptation.
	MPI-ESM1-2-HR (Mauritsen et al., 2019)		SSP3-RCP7.0: 3.6° C global warming, 7.0 W m ⁻² radiative forcing in 2100, business as usual world with regional rivalry and medium-high emissions pathway.
	MRI-ESM2-0 (Yukimoto et al., 2019)		
	UKESM1-0-LL (Sellar et al., 2020)		SSP5-RCP8.5: 4.4° C global warming, 8.5 W m ⁻² radiative forcing in 2100, pessimistic scenario with high economic growth and high fossil fuel dependence.

The simulations were initiated from the best-fit set of initial conditions as described in section 2.2.2. These simulations include an additional spin-up year to avoid sudden jumps between initial conditions and the simulation start year. All simulations were conducted on Snellius, the Dutch national supercomputer. The MODFLOW6 prototype used in GLOBGM requires substantial preprocessing of the PCR-GLOBWB2 inputs into exactly one binary file per unstructured grid. In addition, post-
240 processing is required to convert the binary outputs created by MODFLOW6 into data formats more commonly used by the community. To streamline the multi-model simulations, the pre-processing, MODFLOW6 simulation, and post processing steps were collected into a single snakemake workflow constructed to fit the HPC infrastructure on Snellius(Mölder et al., 2021). Satisfactory simulation times were achieved when running GLOBGM v1.0 on 12 compute nodes (Verkaik et al., 2024); each simulation presented here was assigned 12 compute nodes. For the historical reference simulation (1960–2019), the total
245 run time was approximately 21 hours. Whereas for a single GCM simulation, with its baseline (1960–2014) and three future runs (2015–2100), the total simulation time was approximately 106 hours.

2.4 Quality Assurance, Validation and Bias Correction

2.4.1 Quality Assurance

The hydrogeological schematization in GLOBGM targets large sedimentary aquifer systems. In regions dominated by karst,
250 permafrost, or mountainous terrain, the simulations have been shown to be relatively less accurate (Verkaik et al., 2024; de



Graaf et al., 2017; Gädeke et al., 2020). Therefore, a map of regions where predictions are likely less reliable or fall outside the range of applicability of GLOBGM v1.1 is provided. Regions dominated by karst aquifer systems were obtained from the World Karst Aquifer Map (WHYMAP WOKAM; Chen et al. 2017) and rasterized to match the grid resolution used in GLOBGM. The Global Permafrost Zonation Index Map was used to obtain regions dominated by permafrost; a threshold value of 0.9; corresponds to areas with permanent and continuous permafrost (Gruber, 2012; Ford and Frauenfeld, 2016). To identify mountainous regions, the standard deviation of slope within each GLOBGM grid cell was utilized. Grid cells within the top 1 % of topographical heterogeneity, characterized by slope standard deviation values exceeding 77, were delineated as mountainous terrain.

2.4.2 Validation

The simulations from the historical reference simulation (Section 2.3) were validated against observations obtained from IGRAC's Global Groundwater Monitoring Network (IGRAC, 2024). Observations were filtered to include only those wells with a minimum of 8 years of data for the period 1965–2019; observations prior to 1965 were excluded from validation to avoid the effects of spin-up artifacts (see 2.2.2). To assess the accuracy of the simulations, the non-parametric Kling-Gupta Efficiency (KGE-NP; Eq. 10; Pool et al. (2018)) was calculated between observed and simulated groundwater heads. Validation focused on groundwater heads, not water table depth, to avoid numerical instability in the KGE-NP bias term caused by small denominators. KGE-NP values range from $-\infty$ to 1.0, with values greater than -0.41 implying that the model is a better predictor than the mean of the data (Knoben et al., 2019). To match the observed data to the correct model layer, the same procedure relying on correlations between soil moisture and simulated water table depths was used as in the calibration (Section 2.2.1) and initial conditions (Section 2.2.2).

$$\text{KGE-NP} = 1 - \sqrt{(r_s - 1)^2 + (\alpha_{\text{NP}} - 1)^2 + (\beta - 1)^2} \quad (10)$$

where r_s is the Spearman rank correlation between the simulated and observed time series; α_{NP} is the variability ratio, calculated from the relative flow duration curve as $1 - \text{MAE}(\text{FDC}_{\text{sim}}, \text{FDC}_{\text{obs}})$, where FDC represents the normalized flow duration curves; and β is the bias ratio, defined as the ratio of the mean of simulations to the mean of observations.

2.4.3 Machine Learning Bias Correction

To correct for systematic bias in simulated water table depths, a machine learning-based bias correction approach was implemented. An extreme gradient boosting model (XGBoost) was trained to predict mean annual bias (Eq. 11) between observed groundwater heads from historical reference simulations (see section 2.3). As features, a select number of GLOBGM input parameters (steady-state water table depth, drain conductance, elevation) and Google's AlphaEarth Foundations geospatial embeddings (Brown et al., 2025) were used. Observations were obtained from IGRAC's Global Groundwater Monitoring Network (IGRAC, 2024) and included all available annual observation wells (126 179) that overlapped with the simulation period.



$$\overline{\text{Bias}_{\text{HDS}}} = \frac{1}{n} \sum_{i=1}^n (\text{HDS}_{\text{sim},i} - \text{HDS}_{\text{obs},i}) \quad (11)$$

where HDS_{sim} is the simulated groundwater head and HDS_{obs} is the observed groundwater head.

The FLAML framework (Wang et al., 2019) was used to automate hyper-parameter optimization and model training. The model was trained using 80 % of observations (100 934) through five fold cross-validation and the mean absolute bias as the objective function. The trained model was evaluated on the remaining 20 % (25 234). To assess the effectiveness of the bias correction, KGE-NP values were calculated before and after applying the bias correction to the simulated groundwater heads and compared. The performance of the model was assessed using coefficient of determination, MAE, and RMSE metrics on the holdout set. We evaluated how the bias correction improved simulations for the training and holdout sets, as well as for monthly observations, using the non-parametric Kling–Gupta Efficiency (KGE-NP; Eq. 10). For this, stations were filtered to include only those with a minimum of 8 years of data and that did not contain artesian groundwater conditions, resulting in a total of 7150 monthly and 25 234 annual observation wells.

The trained model was then used to predict the bias (ML_{bias}) for all grid cells in the model domain, which was then applied to bias-correct GLOBGM v1.1 groundwater heads and water table depths described in section 2.3. To prevent overcorrections and preserve the original model physics, an additive bias correction was applied and constrained based on the mean simulated water table depth (Eq. 12). For the historical reference simulation, the mean over the entire simulation (1965–2019) was used and for the scenario-based simulations the mean for the historical baseline period (1965–2014) was used. To prevent numerical instabilities and over-correction in shallow zones, we introduced a scaling factor. The scaling factor modulates the intensity of the bias correction based on the mean simulated water table depth and ensures a spatially smooth correction field (Eq. 13 & 14). Finally the correction was applied to the simulated heads and water table depths (Eq. 15 & 16).

$$\overline{\text{WTD}}_{\text{sim}} = \frac{1}{T} \sum_{t=1}^T \text{WTD}_{\text{sim}}(t) \quad (12)$$

$$\alpha = \begin{cases} 0 & \text{if } \bar{W}_{\text{sim}} \leq d_{\text{protect}} \\ \frac{\bar{W}_{\text{sim}} - d_{\text{protect}}}{w_{\text{ramp}}} & \text{if } d_{\text{protect}} < \bar{W}_{\text{sim}} < d_{\text{protect}} + w_{\text{ramp}} \\ 1 & \text{if } \bar{W}_{\text{sim}} \geq d_{\text{protect}} + w_{\text{ramp}} \end{cases} \quad (13)$$

where d_{protect} is a protection depth below which no correction is applied (set to 0.1 m), and w_{ramp} is the ramp width over which the correction scales from 0 to 1 (set to 5 m).

$$\Delta_{\text{corr}} = \begin{cases} -0.5|\bar{W}_{\text{sim}}| & \text{if } \text{ML}_{\text{bias}} \cdot \alpha < -0.5|\bar{W}_{\text{sim}}| \\ \text{ML}_{\text{bias}} \cdot \alpha & \text{if } -0.5|\bar{W}_{\text{sim}}| \leq \text{ML}_{\text{bias}} \cdot \alpha \leq 0.5|\bar{W}_{\text{sim}}| \\ 0.5|\bar{W}_{\text{sim}}| & \text{if } \text{ML}_{\text{bias}} \cdot \alpha > 0.5|\bar{W}_{\text{sim}}| \end{cases} \quad (14)$$



$$\text{HDS}_{\text{corrected}}(t) = \text{HDS}_{\text{sim}}(t) - \Delta d_{\text{correction}} \quad (15)$$

$$\text{WTD}_{\text{corrected}}(t) = \text{WTD}_{\text{sim}}(t) + \Delta d_{\text{correction}} \quad (16)$$

2.5 Analysis

To provide insights into the spatial distribution of 30 arc-seconds groundwater reserves simulated by GLOBGM v1.1 the long-term average annual groundwater table depth was calculated at the global-scale (Eq. 17) to display spatial variation and highlight the benefit of increased resolution. For visualization purposes, a ‘from ground level downwards’ view of groundwater table depths is provided; for locations where a confining layer exists, only the top model layer is visualized.

$$\overline{\text{WTD}}_{i,j} = \frac{1}{N} \sum_{k=1}^N \text{WTD}_{i,j,k} \quad (17)$$

where $\overline{\text{WTD}}_{i,j}$ is the average WTD at grid point (i,j) , N is the total number of years, and $\text{WTD}_{i,j,k}$ is the WTD at grid point (i,j) in year k .

To analyse how groundwater resources have changed in the past and may change in the future, we calculated the slope of change for water table depth using the Theil-Sen estimate (Sen, 1968; Theil, 1950) so that the slope for each cell was calculated along the time dimension as the median of all pairwise slopes between distinct time points (Eq. 18). To assess statistical significance, we tested the slopes for each grid cell using the Mann-Kendall test (Kendall, 1975). Slopes less than 0.01 m year^{-1} , not statistically significant at the 95 % confidence level, or both were masked. For visualization purposes, only the bottom layer of the model is presented, which represents fully unconfined aquifers or confined aquifers. To complement the trends, we calculated the mean (weighted by grid cell area) water table depth change anomalies over the simulation period aggregated by continent.

$$m_{i,j} = \text{median} \left\{ \frac{h(t_l, i, j) - h(t_k, i, j)}{t_l - t_k} \mid \forall k < l \right\} \quad (18)$$

Where $m_{i,j}$ is the median slope at grid point (i,j) , $h(t_k, i, j)$ is the value of h at time t_k and grid point (i,j) , The median is taken over all pairs of time indices $k < l$.

3 Results

3.1 Calibration, Spin-up and Initial Conditions

3.1.1 Calibration

GLOBGM proved sensitive to adjustments in hydraulic conductivity, anisotropy, and entrance resistance (Table 4 & Fig. 2e). We selected the optimal calibration set by prioritizing the accuracy of shallower estimates ($<60 \text{ m}$), as these reserves are more



Table 4. Comparison of the top 3 calibration candidates (ordered by mean depth-weighted bias) against the reference GLOBGM v1.0 benchmark. Parameter settings are dimensionless multipliers applied to the default values.

Model	Parameter Settings					Bias	Bias by Depth Class (m)						
	Ks (Fine+Coarse)		Carbonate	Anisotropy	Entry Res.	Mean	< 0	0–5	5–10	10–20	20–60	> 60	Mean
	Uncon.	Consol.											
Candidate 1	0.1	0.1	0.1	0.1	0.1	3.6	37.3	9.6	12.0	15.6	27.2	93.3	32.5
Candidate 2	0.1	0.1	0.1	1.0	0.1	3.3	37.7	9.7	11.9	15.4	27.0	93.1	32.5
Candidate 3	0.1	1.0	0.1	0.1	0.1	1.1	37.6	9.8	12.3	15.9	27.2	92.7	32.6
Uncalibrated	—	—	—	—	—	-1.6	40.1	11.8	14.7	18.3	26.5	85.7	32.9
v1.0	—	—	—	—	—	-4.8	41.3	13.6	17.1	20.4	27.3	85.6	34.2

relevant for human use and for sustaining ecosystems through baseflow contributions to surface water bodies and groundwater-dependent ecosystems (Reinecke et al., 2023; de Graaf et al., 2019; Reinecke et al., 2024; Fan et al., 2013). Average bias per depth class was calculated, and the average of these classes was used to obtain an overall depth weighted bias for each calibration permutation. Calibration permutations were rank ordered by average depth weighted bias, and the top ten were identified. Among these, one permutation displayed the best accuracy for the 0–5 m and 10–20 m depth categories, whilst still displaying the second best accuracy for the 5–10 m depth category (Table 4). Consequently, a prefactor of 0.1 for all parameters was chosen as the best set. The selection of these prefactors corresponding to the minimum values tested suggests that smaller prefactors may have provided better performance. However, the range of calibration factors was intentionally constrained within two orders of magnitude to avoid significant deviations from original parameterization and prevent over-fitting.

Greater variance among permutations for locations where simulated water table depth are deeper than observed was evident. This suggests that calibration influenced over-estimated depths more significantly than under-estimated ones (Fig. 2e). Varying calibration prefactors tended to shift entire distributions rather than simultaneously correcting outliers at both ends (Fig. 2e). As a result, the chosen prefactor set improved the accuracy of over-estimated depths but increased the error for under-estimated depths. However, this set maintained the best overall bias, as gains at the negative end of the distribution were offset by penalties at the positive end of the distribution (Table 4 & Fig. 2e). Compared to the previous un-calibrated GLOBGM v1.0 (Verkaik et al., 2024), the chosen set of calibration settings display marked reductions in overall depth weighted bias (Table 4). This is especially prevalent for shallower water table depths (<0–20 m; Table 4). For wells between 20 and 60 m calibration has little effect, whereas for wells deeper than 60 m calibration increased bias (Table 4). In comparison to GLOBGM v1.0 unweighted bias shifts from a tendency to overestimate water table depths (-4.8 m) to underestimation when calibrated (3.6 m) (Table 4 & Fig. 2e). Comparing the uncalibrated GLOBGM v1.1 with v1.0 reveals how the saturated area fraction correction and groundwater recharge downscaling affect the model's accuracy. The uncalibrated predictions were closer to the ideal bias score of zero, by a magnitude of 1.6 m, again centered around improvements in the wells that were deeper than observed which indicates model improvement (Table 4 & Fig. 2e).

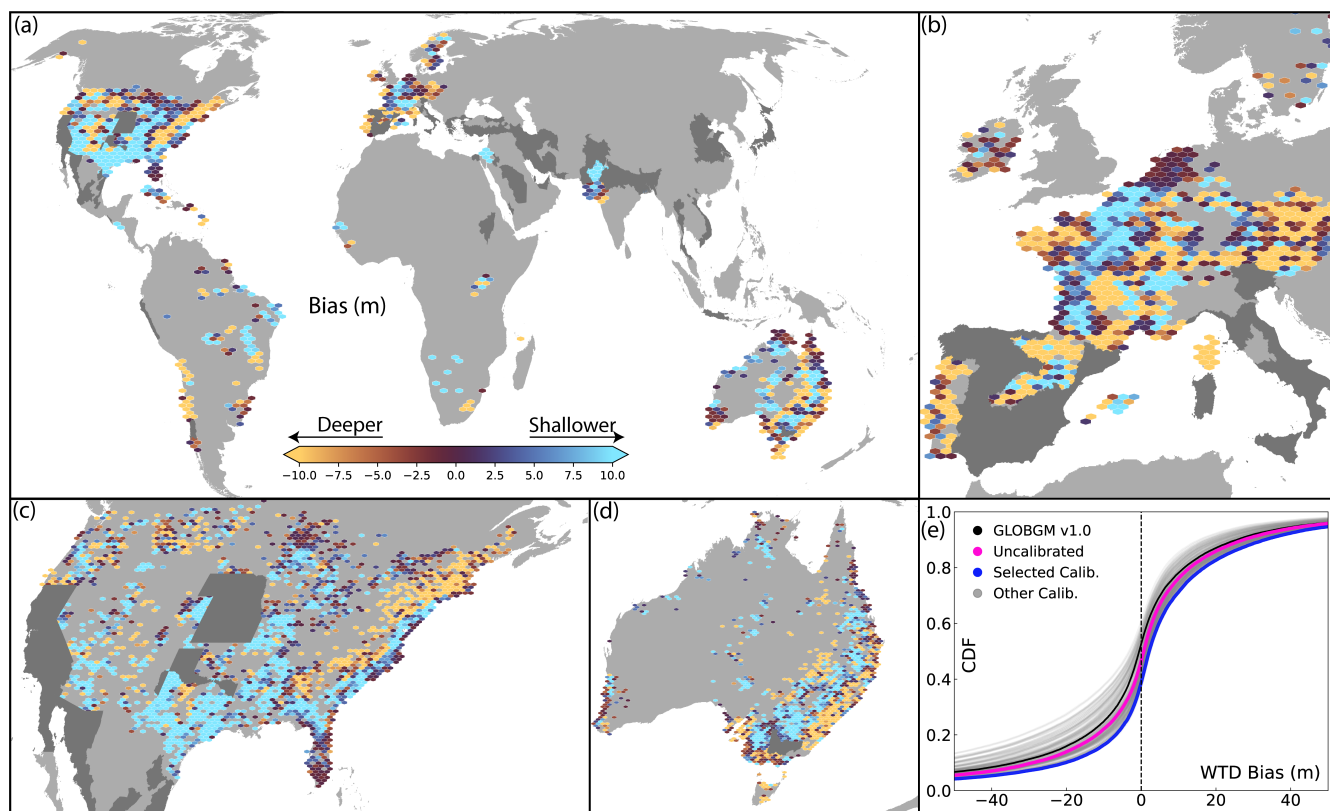


Figure 2. (a) Geographical variation of water table depth bias, for the chosen calibration setting, from 34 800 observation wells that were used to assess different calibration settings. Dark grey polygons indicate regions of known anthropogenic influence on the water system and were excluded from the calibration. With zoomed insets of the (b) western Europe, (c) United States, (d) Australia. (e) Cumulative distribution function of the water table depth bias of the 163 calibration permutations.

3.1.2 Spin-up and Initial Conditions

The objective of the first transient spin-up step was to achieve dynamic equilibrium for the year 1960, whereas the second step incorporated abstraction to reproduce expected water table depths. The key assumption is that initiating the simulation from a naturalized steady state solution and simulating 1960 iteratively, without abstraction, would yield water table depths which are shallower than observed; therefore by activating abstraction in the second step, water table depths would approach observed values. During the first step, the change in average groundwater heads between successive iterations decreased — suggesting the model approached dynamic equilibrium. The correlation based analysis aimed at detecting regions where the model was not properly spun up reveals that the majority of grid cells were correctly initiated at the start of the simulation. Regions that were incorrectly initiated were sparse and geographically confined. Most regions reached adequate spinup within the first 5 years, with 99.7 % of the domain deemed valid after the first year, 99.8 % after the second year, 99.9 % thereafter, and 99.99 % by 1965 (Fig. 3b).

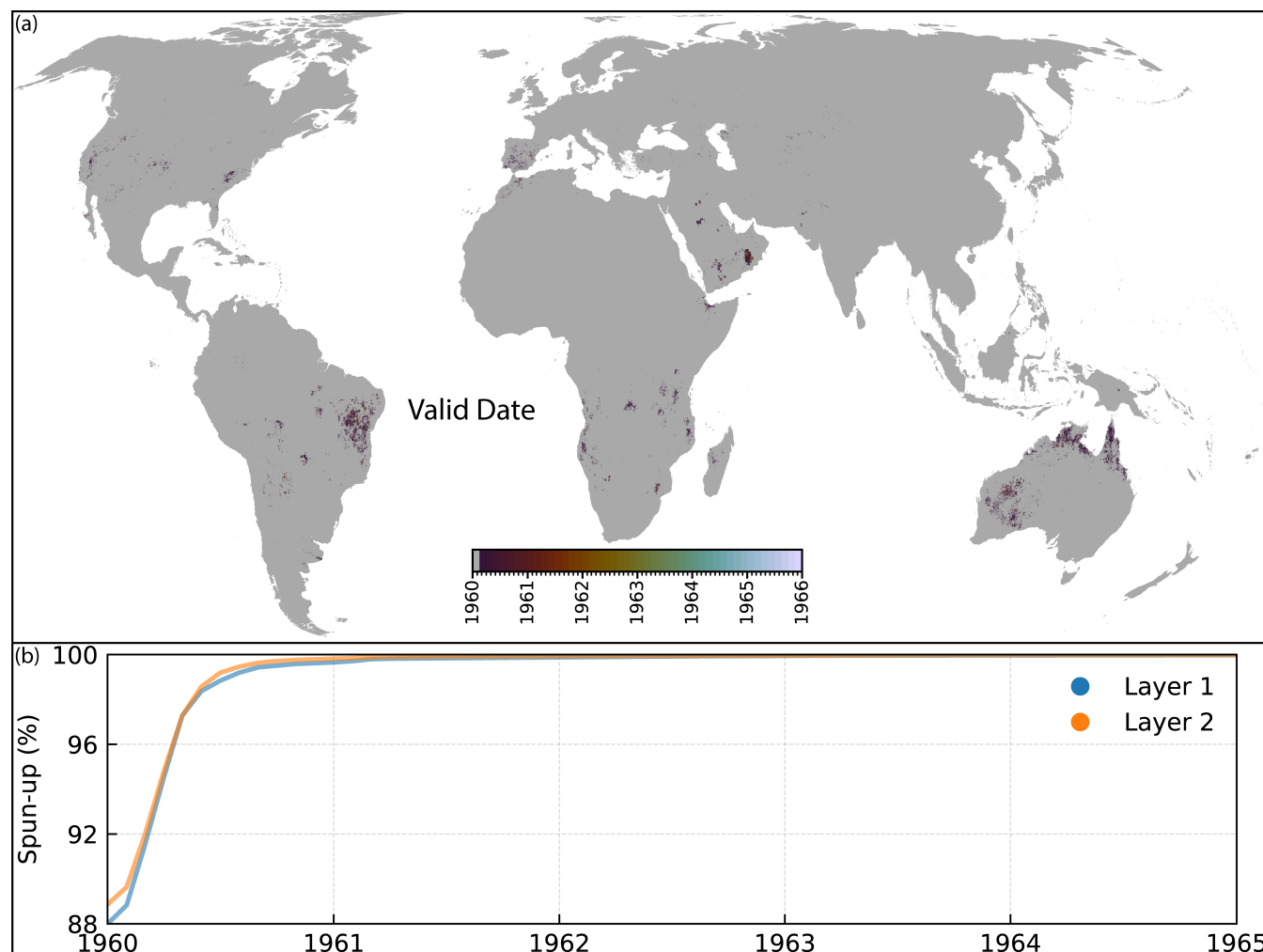


Figure 3. (a) Map showing the first year of valid simulation data per grid cell. (b) Percentage of grid cells in Layer 1 and Layer 2 that are free of spin-up artifacts between 1960 and 1965.

3.2 Quality Assurance, Validation and Bias Correction

370 3.2.1 Quality Assurance

Previous evaluations on the 5 arc-minute predecessor and prior version of GLOBGM have revealed that GLOBGM varies in performance for regions where geology is dominated by karst aquifers, permafrost, or mountainous regions (Fig. 4). GLOBGM has been shown to be more accurate in sedimentary basins compared to karst, presumably due to the underlying, questionable assumption that groundwater flow follows the Darcy approximation in karst (de Graaf et al., 2017; Verkaik et al., 2024).

375 Regarding the accuracy of predictions in mountainous regions, GLOBGM models groundwater heads within the mountain



blocks but does not parametrize for water bodies in hill slopes and smaller alluvial mountain valleys, since the hydrogeological schematization is too coarse in these regions. Moreover, it is assumed that the secondary permeability of fractured hard rock in mountainous regions can also be represented by the principles of Darcian groundwater flow (Verkaik et al., 2024). Research into the applicability of PCR-GLOBWB2 to simulate the hydrology of permafrost regions shows that it does not capture the dynamics of permafrost with high fidelity due to its simplistic representation of ice melt (i.e., degree day factor) and lack of dynamic permafrost schemes leads to underestimated discharge and by proxy groundwater recharge (Gädeke et al., 2020). Moreover, it is challenging to define groundwater resources in areas where soil is frozen up to many tens of meters deep (Diak et al., 2023; Cochand et al., 2019). These masks are provided, together with the simulations, to inform users of regions where GLOBGM v1.1 predictions are likely less reliable.

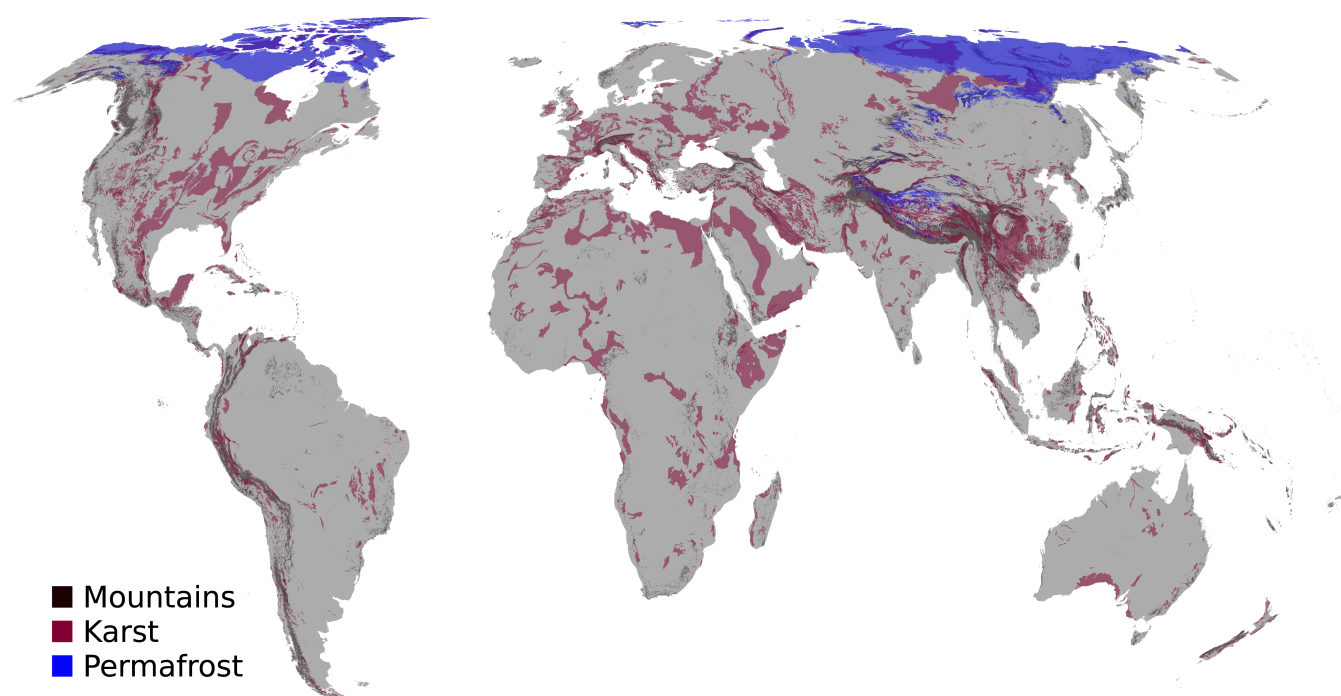


Figure 4. Regions where GLOBGM has been shown to provide less accurate results due to practical and theoretical limitations; namely, karst aquifers, permafrost and mountainous regions.

3.2.2 Validation

The ability of GLOBGM to simulate groundwater heads varies predictably with water table depth. KGE-NP scores grouped according to the depth categories reveal that the deeper simulations (>60 m & 20–60 m) are the worst performing, with approximately 75 % displaying a KGE-NP greater than -0.41, which is indicative of skillful predictions that improve upon the mean flow benchmark (Fig. 5b,f; Knoben et al. (2019)). The remaining observations between 0–20 m were more accurate



390 with approximately 90 % of locations showing skillful predictions (Fig. 5b,f). The ability to reproduce correlations shows a clear pattern with depth, where shallower values are more accurately predicted than deep and above ground water table depths (Fig. 5c,g). Whereas, the ability to reproduce variability shows no clear correlation with depth of observed groundwater in terms of average variance ratio or the bias of the variance ratio (Fig. 5d,h). Groundwater heads tend to be underestimated for observations between 0–5 m (Fig. 5e,i). Whereas groundwater heads are over estimated for deeper wells (Fig. 5e,i).

395 The spatial variation in KGE-NP score corroborates some of the conclusions derived from the cumulative distribution functions (Fig. 5a), for instance, lower KGE-NP values are evident along regions where water table depths are expected to be shallower such as arid regions and mountainous terrain (Fig. 5a). Furthermore, the effect of depth becomes evident when comparing relatively lower KGE-NP scores in the higher elevations of the Appalachian, Pyrenees, and Great Divide mountains to higher scores in the adjacent low lying regions (Fig. 5a). Similarly, the higher regions of the Indo-Gangetic plain display lower
 400 KGE-NP values when compared to more coastal points (Fig. 5a). Comparing the monthly to annual KGE-NP scores reveals that annual scores are generally higher than monthly scores (Fig. 5a, e). This trend is also evident in the constituent components of KGE-NP, where correlations are generally higher when validated at annual scales (Fig. 5c, g). Variability ratio and bias ratio show no clear trends when comparing monthly to annual values (Fig. 5d,h,r,f).

3.2.3 Machine Learning Bias Correction

405 The machine learning based bias correction demonstrated adequate performance on the holdout set, achieving a coefficient of determination of 0.6, RMSE of 32.45, and MAE of 14.4 m (Table. A1). Hydraulic conductivity (k_h) emerged as the most important feature (0.10), followed by Alpha Earth embeddings A42 (0.05) and A39 (0.04) (Fig. A1a) and elevation (0.03). The ML-corrected GLOBGM estimates show a predominantly negative bias for central North America and Europe (Fig. 6a). Whereas, the Indian subcontinent and semi-arid regions of southern Africa displayed a positive bias (Fig. 6a).

410 When grouped by depth classes, the original model showed a consistent pattern of over-estimation for shallower wells (0–10 m) and under-estimation for deeper wells (>20 m) (Fig. 6d,e). The ML correction reduced the median bias across all depth classes and improved the bias for the majority of the locations (Fig. 6d,e & Fig. A1b,c). After ML correction, the median bias shifted closer to zero across all categories, and the interquartile range contracted (Fig. 6d,e). This reduction was particularly notable for the >60 m class, where the original model displayed the largest biases (Fig. 6d,e). The bias correction was more
 415 effective at the annual time scale compared to the monthly time scale, with a more pronounced reduction in bias (Fig. 6d,e). Post correction bias reductions translated to improved KGE-NP scores. For both the annual and monthly time scales, the ML-corrected simulations displayed higher median KGE-NP scores and reduced inter-quartile ranges across all depth classes (Fig. 6b,c). These improvements led to ~4 % and ~0.6 % increase in the proportion of observations with skillful predictions ($KGE-NP > -0.41$) at annual and monthly resolutions, respectively (Fig. A1d).

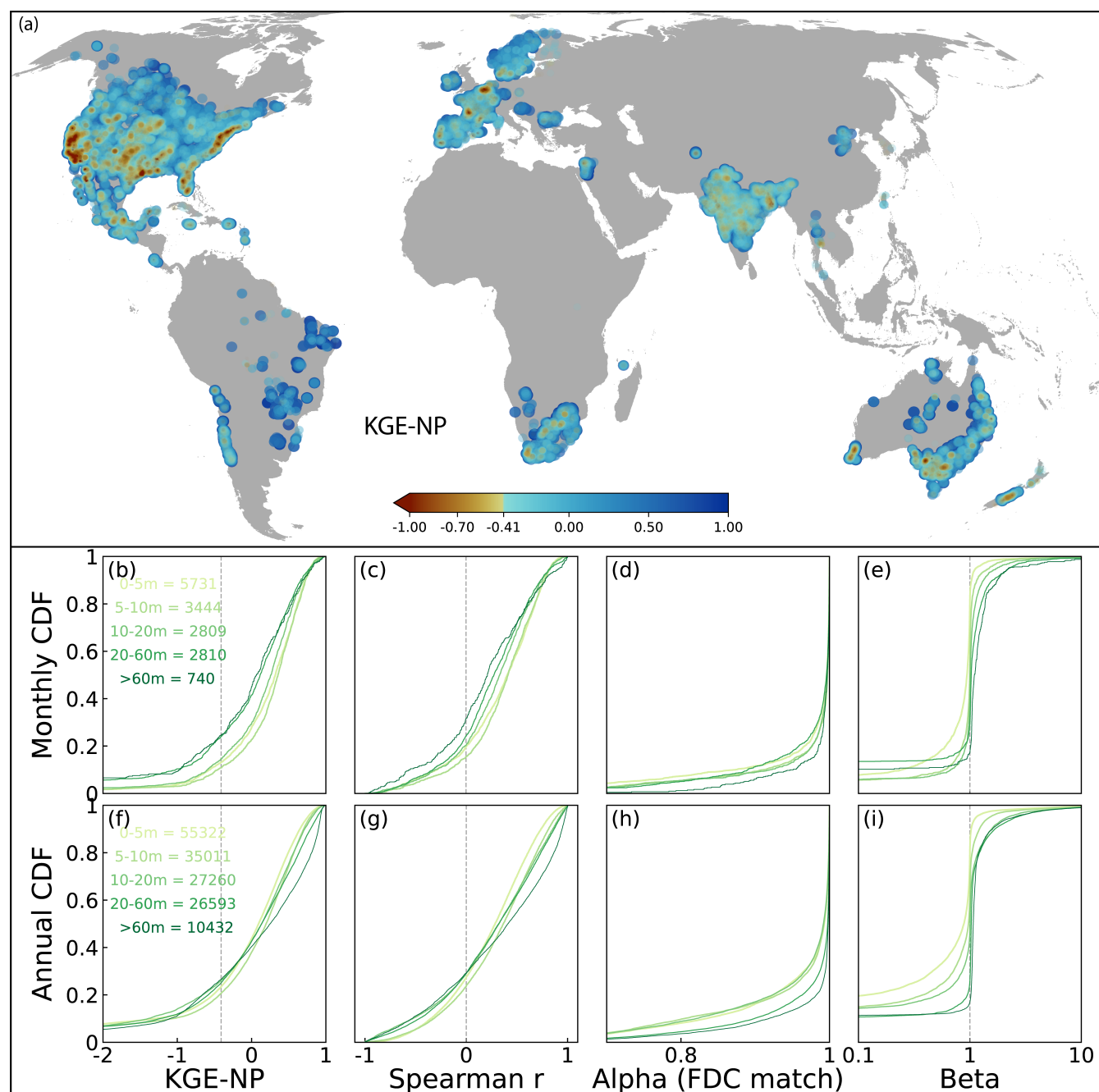


Figure 5. KGE-NP spatial distribution (a) and CDFs (b, f) and its constituent components (c, g) correlation, (d, h) beta, and (e, f) alpha calculated for monthly and annual water table depth simulated by GLOBGM from 1965–2019. Gray lines in (b, f) symbolises a KGE-NP of -0.41 which is indicative of predictions that improve upon the mean flow benchmark. Gray lines in (c, g) indicate the point at which correlations positive to the right and negative to the left, in (d, h, e, i) the point at which separates overestimation to the right and underestimation to the left.

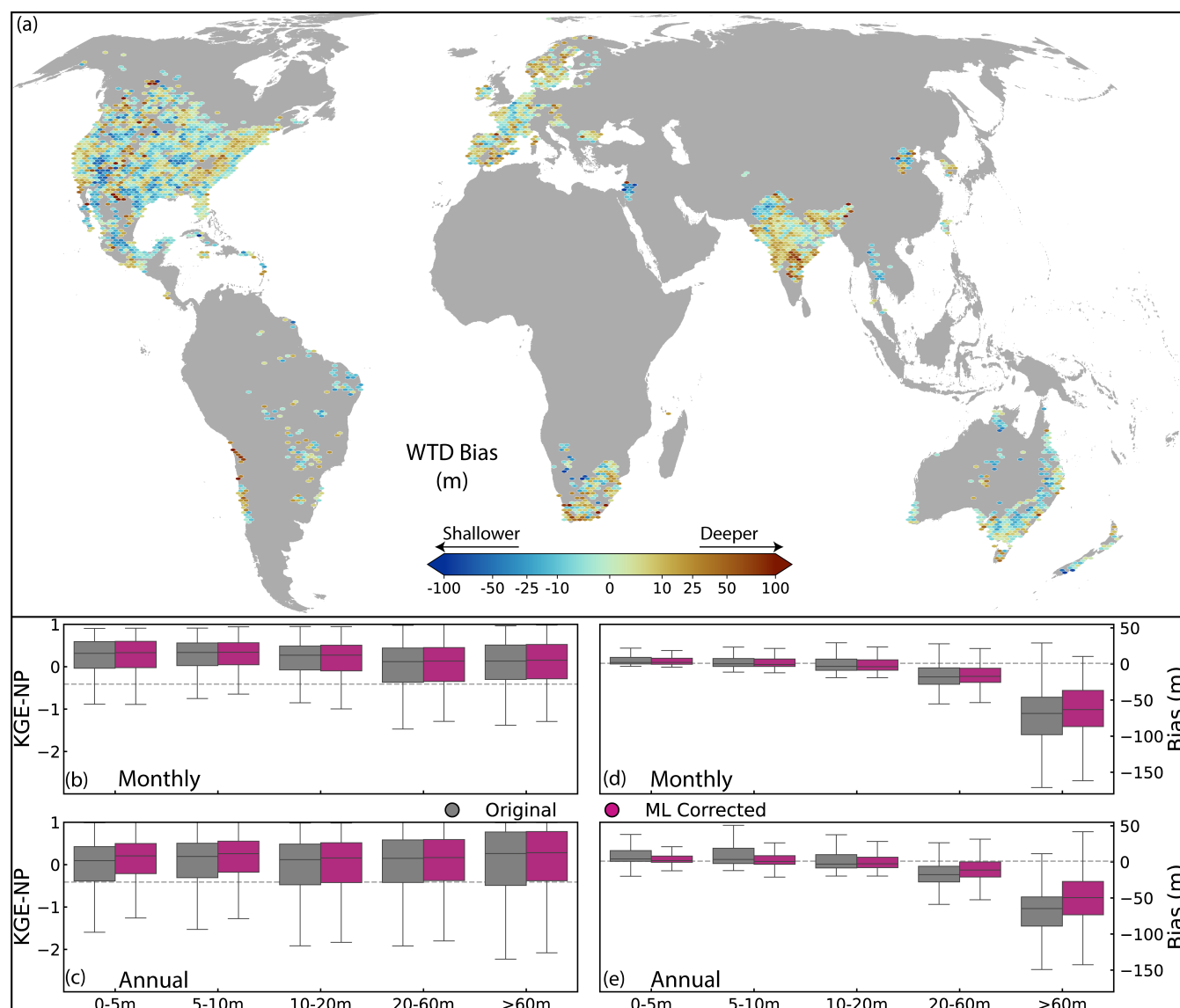


Figure 6. (a) Spatial distribution of ML-corrected mean water table depth bias (m). Original and ML-corrected (b) monthly and (c) annual non-parametric Kling-Gupta efficiency (KGE-NP) grouped by mean mean water table depth classes. Original and ML-corrected (d) monthly and (e) annual mean water table depth bias (m) across the same WTD depth classes, comparing original and ML-corrected simulations.

420 3.3 Historical Reference Simulation

Mean annual water table depth, from 1965–2019, as portrayed by the hyper-resolution groundwater model highlights the added detail in representing the spatial variability of global groundwater depths (Fig. 7a). Finer scale interactions between river systems and groundwater are evident at this hyper-resolution.

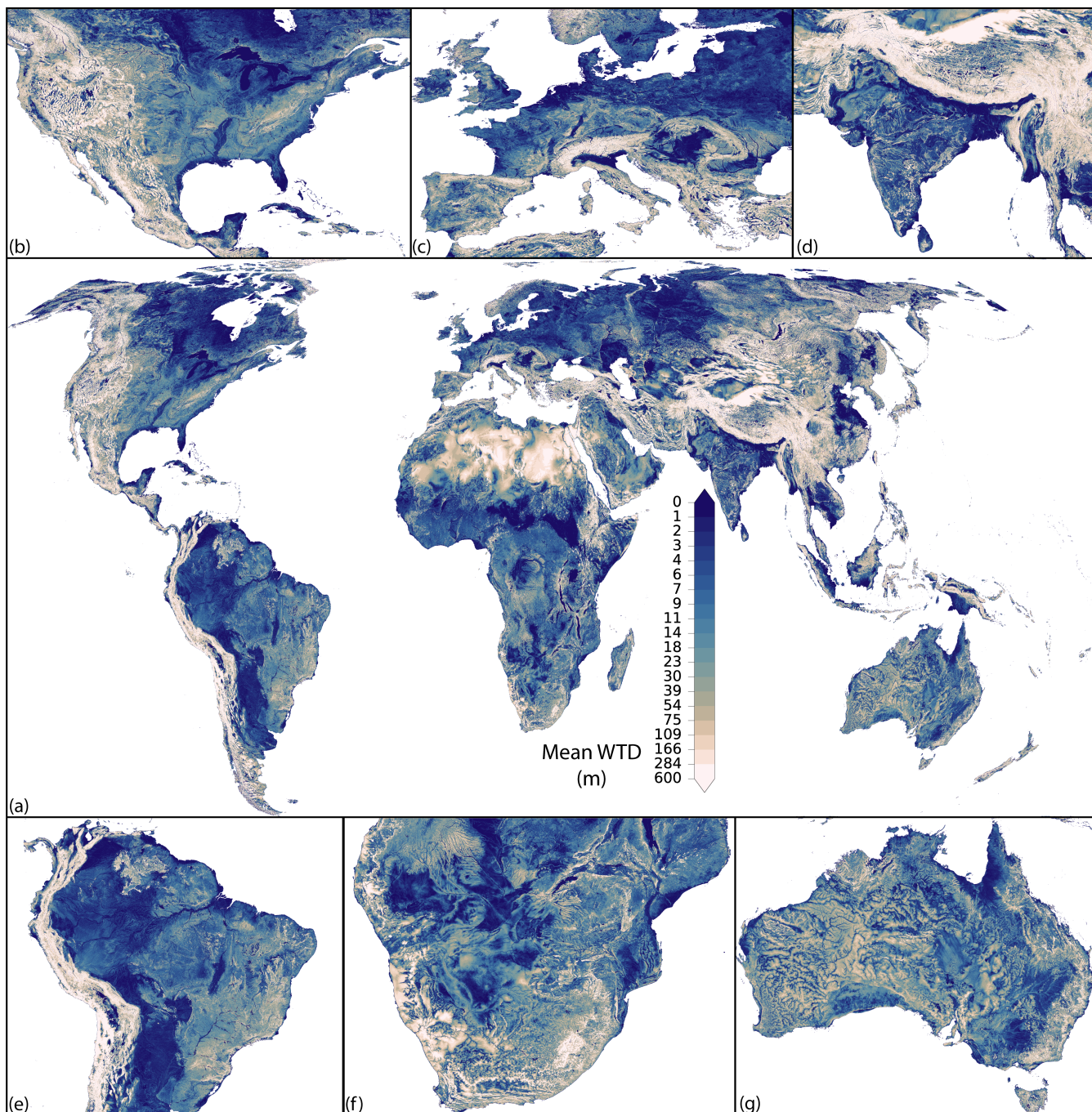


Figure 7. (a) Mean annual water table depth (1965–2019) simulated by GLOBGM. With zoomed insets providing more detail over (b) the United States, (c) western Europe, (d) Indo-Gangetic Plain, and (e) the Amazon basin, (f) southern Africa and (g) Australia. In areas where a confining layer is present, only the uppermost model layer is depicted, thus providing a from the ground surface downwards perspective



Such as the shallow reaches of the Amazon and Mississippi rivers, with their respective deltas, are visible through shallow
 425 water tables (Fig. 7e,b). Other major deltas, like the Indus and Ganges-Brahmaputra (Fig. 7d), are also well resolved at this
 resolution and shallower water table depths along the coast compared to more interior regions are present (Fig. 7b-g). Average
 water table depths largely resemble our current broad understandings of groundwater dynamics and reflect the influences of
 precipitation, topography and geology on driving the distribution of groundwater reserves.

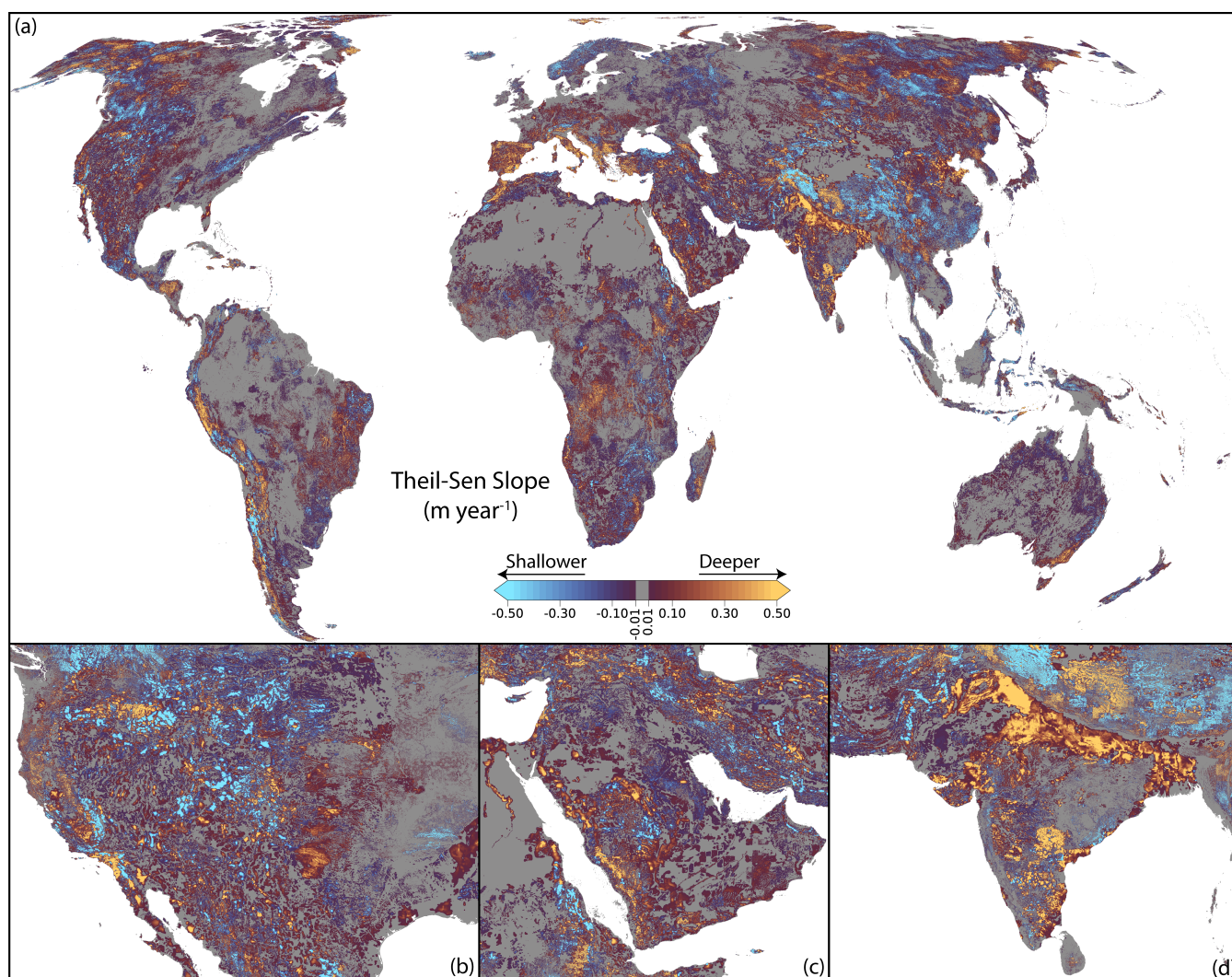


Figure 8. Theil-Sen slope of annual water table depth (1965–2019) simulated by GLOBGM. Gray pixels indicate regions where slopes are either insignificant, less than 0.01 m year^{-1} , or both. Note that this figure represents the lower model layer which is representative of unconfined and confined aquifers. Zoomed insets provide more detail over (b) United States, (c) the Arabian Peninsula, and (d) the Indo-Gangetic Plain



Water table depths are seen to be deeper in the arid regions and shallower in humid areas; reflecting the difference in
 430 groundwater recharge rates which is driven by large-scale precipitation patterns (Reinecke et al., 2024). More mountainous
 regions exhibit deeper water table depths which is expected given the relatively lower infiltration rates, lower permeability,
 and tendency for water to move down slope by streams towards lower lying ground under the influence of gravity (Somers
 and McKenzie, 2020). GLOBGM reproduces declining water table depths in regions where groundwater depletion is known
 to occur, such as the southern and central high plains in the United States (Fig. 8b), Iran (Fig. 8c), Indo-Gangetic plain (Fig.
 435 8d; (Jasechko et al., 2024)). Interestingly, GLOBGM identifies a number of regions where water table depths have risen, the
 majority of which occur in the northern latitudes and Arctic Circle, such as Northern Scandinavia and the Russian Far North.
 This could conceivably be due to climate change enhancing precipitation and groundwater recharge dynamics (McKenzie et al.,
 2021). In addition, simulated rising water table depths match regions that have been shown to display increases in groundwater
 reserves such as the Guarani aquifer in South America, the northern Great Plains in North America, and the Karoo Basin in
 440 South Africa (Fig. 8a; Jasechko et al. (2024); Taylor et al. (2013)).

3.4 Ensemble-mean Scenario Simulations

The trends in water table depths, between 2014 and 2100 shows that the different SSP-RCP scenarios do indeed suggest
 alternate futures for groundwater reserves (Fig. 9). At the global scale, more regions exhibit rising than declining water table
 depths (Fig. 9). The intensity of change and proportion of areas affected projected to under go change differs among the three
 445 SSP-RCP scenarios. More specifically, SSP1-RCP2.6 shows the lowest projected rate of change in water table depths, followed
 by SSP3-RCP7.0, whilst SSP5-RCP8.5 displays the largest projected rate of change (Fig. 9). Many regions show consistent
 directions of change across all three scenarios; however, some differences are evident. The regions in which water table depths
 are projected to rise are the Sahel, Horn of Africa, North America's Pacific Coast, Tibetan Plateau, Northern Polar regions
 and Arabian Peninsula (Fig. 9). In contrast, water table depths are projected to decline in the Mediterranean, south eastern
 450 Australia, and the African and South American tropics (Fig. 9). The Alps and the Amazon are some of the few regions that
 show disagreement in the direction of change. A declining water table depth is projected for the Amazon under SSP3-RCP7.0
 and SSP5-RCP8.5, in contrast to a projected stable water table depth SSP1-RCP2.6. Water table depth in the Alps are projected
 to rise SSP1-RCP2.6 and decline under SSP3-RCP7.0 and SSP5-RCP8.5 scenarios (Fig. 9).

Projected trends indicate that water tables will rise across the majority of continents in the coming century (Fig. 10). This
 455 rise is persistent in Africa and North America, accelerating under higher emission scenarios (Fig. 10b,d). In Asia, the water
 table remains relatively stable until the 2010s before beginning to rise, again intensifying with higher emissions (Fig. 10c).
 Australasia follows a similar upward trajectory before stabilising around 2050, with minimal variation between scenarios (Fig.
 10g). Conversely, Central and South America exhibit early increases followed by a divergence: SSP1-RCP2.6 leads to stability,
 while the other scenarios result in deepening water tables (Fig. 10g). Small islands show an initial rise until 2010 followed by
 460 a decline, with no discernible difference between the scenarios (Fig. 10f). Europe is the distinct outlier; it is the only region
 where the water table markedly deepens, particularly under higher emission scenarios, while SSP1-RCP2.6 maintains relative
 stability (Fig. 10a).

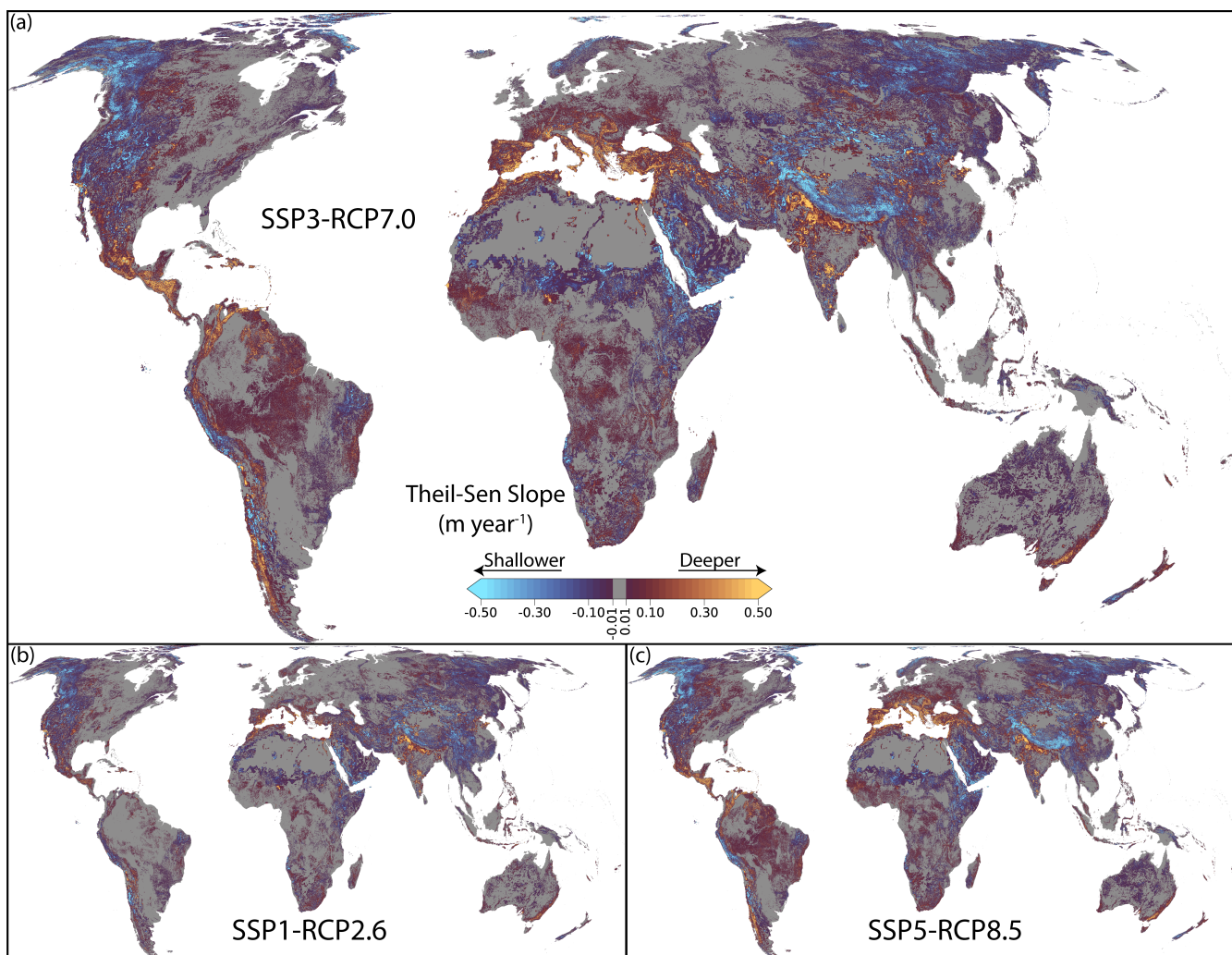


Figure 9. Theil-Sen slope of annual water table depth (2014–2100) simulated by GLOBGM under three SSP-RCP scenario. Gray pixels indicate regions where slopes are either insignificant, less than 0.01 m year^{-1} , or both. Note that this figure represents the lower model layer which is representative of unconfined and confined aquifers.

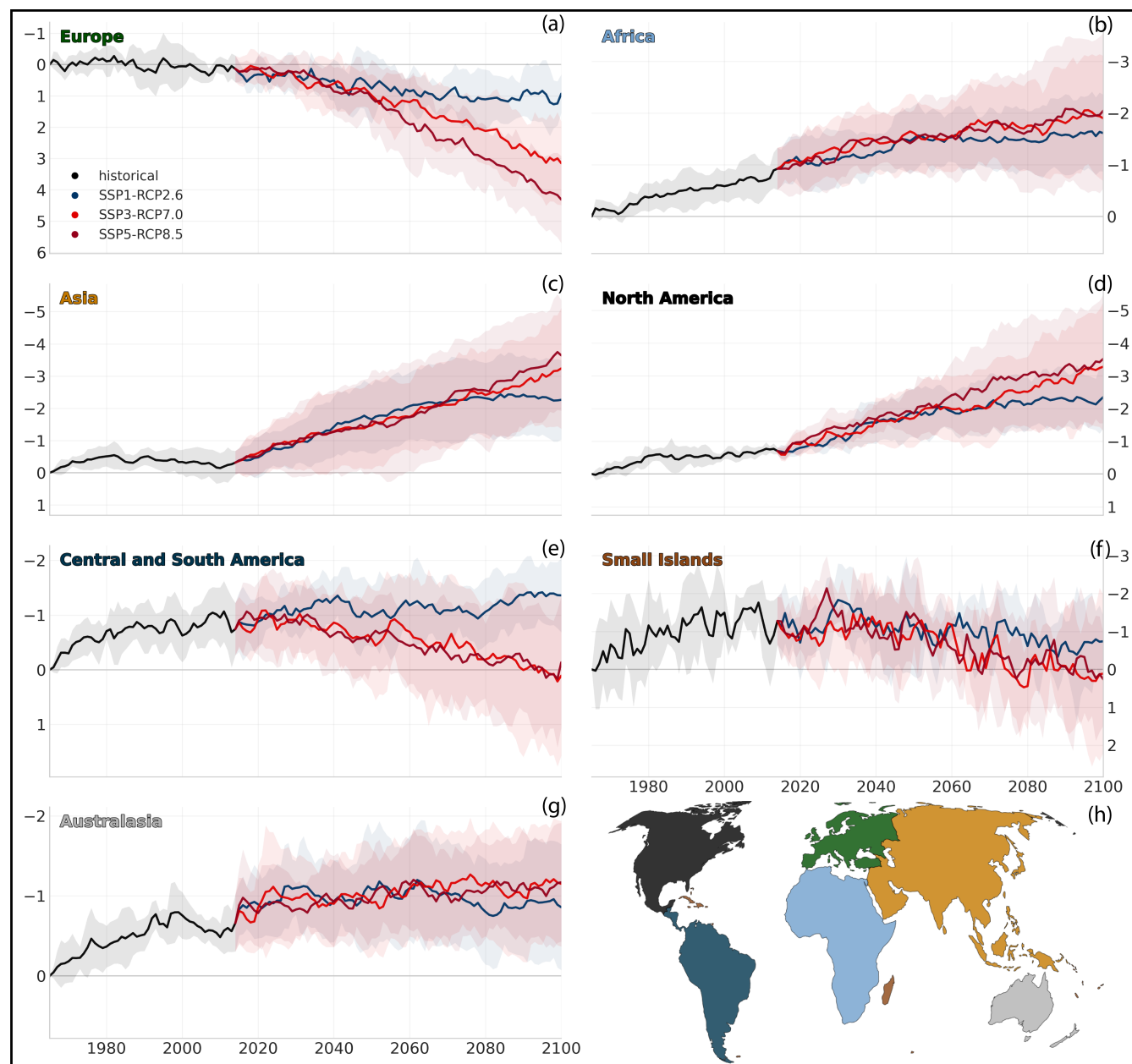


Figure 10. Time series of represent regional average WTD for the historical simulation and three Shared Socioeconomic Pathways: SSP1-RCP2.6, SSP3-RCP7.0, and SSP5-RCP8.5. Shaded bands denote the uncertainty range defined by the standard deviation. The map in (h) defines the aggregation regions



4 Discussion

In this study, we simulate and analyse water table depths and groundwater heads using GLOBGM, a hyper-resolution (~1 km at the equator) global groundwater flow model, for the past and future. These estimates are derived from a calibrated and bias corrected version of GLOBGM, which incorporates a novel methodology for the estimation of initial groundwater conditions, recharge downscaling and machine learning based bias correction.

We show that calibration of hyper-resolution groundwater models provides benefits at the global-scale, which aligns with previous studies that also present improved regional and global groundwater simulations post calibration (Fan et al., 2013). Evaluation of historical water table depths indicates that GLOBGM effectively simulates groundwater dynamics for the majority of locations where observations are present and captures expected patterns of groundwater reserves. The validation results demonstrate that GLOBGM v1.1 achieves skillful predictions ($KGE-NP > -0.41$) of groundwater heads at approximately 75–90 % of sites depending on depth and temporal resolution. Monthly validation reveals depth-dependent performance, with 90% skillful predictions for shallow-intermediate wells (0–20 m) versus 75 % for deep wells (>60 m). In contrast, annual validation shows depth-independent performance (~80 % skillful), suggesting that seasonal dynamics are more challenging to capture than inter-annual trends. Notably, declining water table depth trends in the historical reference simulation align well with known regions of excessive groundwater abstraction and depletion such as the United States, the Arabian Peninsula and the Indo-Gangetic plain. In addition, regions of recovery are captured; such as Guarani Aquifer in South America, the northern Great Plains in North America, and the Karoo Basin in South Africa (Fig. 8d; Taylor et al. (2013); Jasechko et al. (2024)). To further improve simulations, a machine learning-based bias correction was applied. Hydraulic conductivity emerged as the most influential predictive feature, implying that residual GLOBGM errors are primarily linked to uncertainties in subsurface parameterization — a well-recognized source of improvement in large-scale groundwater modeling (Reinecke et al., 2024).

An aspect of the calibration which can be improved is that the application of prefactors at the global-scale, or even based on lithological class as done here, results in the shift of the entire distribution; whereas an idealised results from calibration done on steady state groundwater models and observed well depths would result in the upper and lower ends of the bias values that converge on the ideal bias score of zero. Using global objective functions and calibration factors that influence large regions with very different hydrological characteristics inherently prevents the incorporation of much needed regional sensitivity; this is compounded by uncertainties in the datasets used to define lithological groupings and uneven distribution of observations. Efforts focused around the collection and dissemination of observed groundwater data to broaden the scope and coverage of large sample datasets will play an ongoing role in achieving realistic parameter estimates through calibration (Jasechko and Perrone, 2021). Future studies should investigate whether alternative calibration techniques such as sequential calibration (Béraud et al., 2024) or data assimilation approaches which incorporate observed data (Béraud et al., 2024; White, 2018) may aid in realizing locally specific calibration sensitivities. Similarly, regionalization methods have been shown to improve calibration of global hydrological models (Kupzig et al., 2024).

The approach used to estimate initial groundwater heads resulted in a more accurate estimation of initial states than what would have been accomplished had the traditional approach been used (i.e., steady state simulation followed by transient sub-



model without abstraction for 20 iterations). Nonetheless, results also suggest room for improvement. Ideally, the model would display strict dynamic equilibrium for all locations; however, this would require a disproportionate amount of computational power to achieve relatively small changes in initial groundwater head estimates. The estimation of accurate and representative initial states is an often cited problem for global groundwater models (Reinecke et al., 2024; de Graaf et al., 2017) and earth system models in general (Werndl, 2019). Opportunities exist to improve this by using advances in initial condition estimation algorithms, such as those based on sequence acceleration (Khatriwala, 2024) or by making use of artificial intelligence surrogate models where repetitive iterations can be simulated at a faster rate (Asher et al., 2015; Bennett et al., 2024; Pawusch et al., 2025). Similar to the importance of regional sensitivities during calibration, future studies should aim to develop initial condition estimation frameworks that explicitly incorporate regional responses which, in principle, may be applicable to initial condition estimation frameworks.

When considering the future of global groundwater resources, there is a tendency toward intensification of change in groundwater reserves in the coming century. Overall, more regions are projected to experience an increase in groundwater reserves. This finding aligns well with CMIP6 based studies showing an intensification of the hydrological cycle and is in congruence with previous regional studies (Eccles et al., 2025; Berghuijs et al., 2024). For example, there is broad agreement in terms of direction in the Sahel, Arabian Peninsula, Indus, and Mediterranean. Several regions—such as southern Africa, the Horn of Africa, and the eastern Indian subcontinent—exhibit scenario-dependent directional disagreement, with SSP1-RCP2.6, SSP3-RCP7.0, and SSP5-RCP8.5 projections diverging not only in magnitude but also in the direction of change. These disagreements emphasize the importance of climate mitigation efforts as they can lead to divergent groundwater futures in certain regions, highlighting the need for uncertainty-based adaptation measures.

Since the data produced here were generated following the ISIMIP framework, it is logical to base the future investigation on the rationale underlying this framework. That is, for the historical reference simulation the data produced here can be used for model evaluation against observed data, as is done in this work, and for uncertainty studies which compare models to each other (Reinecke et al., 2024, 2026). Future projections can be used to support studies that aim to investigate the climate-related risks to groundwater reserves under varying levels of climate change and socioeconomic conditions. Future users are especially encouraged to use the existing earth systems data sets to relate groundwater dynamics to other earth system components, as suggested in the review by Huggins et al. (2023). Besides the advantages provided by the high spatial and temporal resolution estimates of groundwater reserves presented here, the inclusion of multiple climate models across three combined climate and socioeconomic projections allows for the incorporation of uncertainty and enhances the robustness of results derived from future studies. For example, water table depth is a critical parameter when assessing the ecological impacts of groundwater depletion (Otoo et al., 2025). By analyzing these impacts using multiple models, it becomes possible to quantify uncertainty and improve confidence in projections. A pertinent case is southern Africa, a region identified here as potentially facing divergent groundwater futures depending on the trajectory of climate change (section 3.4), highlighting the value of multi-GCM assessments in regions with high climatic and hydrological variability. As stated by Reinecke et al. (2026) in the proposed ISIMIP groundwater sector framework, multi-model groundwater heads and water table depths data may facilitate cross-sectoral research on the impact of changing groundwater reserves on social, economic, and ecological systems.



Ultimately, the results presented here may serve as a baseline for future studies aimed at exploring the role of adaptation and mitigation strategies.

5 Conclusions

535 This study provides the first hyper-resolution (~1 km) monthly estimates of global groundwater heads and water table depths for past and future periods. Utilizing the global groundwater model GLOBGM in combination with CMIP6 climate projections for baseline and future periods, the simulated groundwater levels offer a valuable resource to enhance understanding of global groundwater at unexplored spatial and temporal scales. By incorporating dynamic drainage elevation, recharge downscaling, calibration, a novel spin-up procedure, and bias correction, simulations yield predictions which are skillful for 75–90%.
 540 Moreover, this work identifies opportunities for the research community to improve future modelling efforts, including the development of more efficient spin-up procedures and refined initial condition estimations and multi-groundwater model estimates and projections. Ultimately, this study establishes a basis for assessing the impacts of global change on groundwater availability by integrating both anthropogenic and climatic influences on global groundwater reserves, thereby advancing the capacity to evaluate groundwater dynamics within the broader Earth system context.

545 *Code and data availability.* The model code used to run the simulations has been archived on Zenodo and is accessible via <https://doi.org/10.5281/zenodo.17065147>. Simulated groundwater heads and water table depths are available on YODA, the research data management service of Utrecht University. For the historical reference simulation (ISIMIP3a), the dataset includes long-term, annual, and monthly groundwater head and water table depth averages. For the future projections (ISIMIP3b), long-term and annual averages are provided for each GCM under the three SSP–RCP scenarios. Ensemble means for long-term, annual, and monthly estimates of groundwater head and water table depth are
 550 provided across the three scenarios. A dataset related to the quality assurance of the simulated outputs is also available on YODA. The following datasets can be accessed directly: historical-reference-gswp3-w5e5: <https://doi.org/10.24416/UU01-AKSHOX>, globgm-cmip6-monthly: <https://doi.org/10.24416/UU01-1BXLDP>, globgm-cmip6-annual: <https://doi.org/10.24416/UU01-V6B9YS>, globgm-cmip6-average: <https://doi.org/10.24416/UU01-SLRFI7>, globgm-cmip6-quality: <https://doi.org/10.24416/UU01-16EJ3Y>. Instructions and code for accessing the data are available at https://vanjaarsveldbarry.github.io/globgm_cmip6/.

555 *Author contributions.* **BvJ:** Data curation, Formal analysis, Software, Validation, Visualization, Writing — original draft preparation. **NW:** Conceptualization, Project administration, Supervision, Writing — review & editing. **NGO:** Conceptualization, Data curation, Software, Writing — review & editing. **ES:** Conceptualization, Data curation, Software, Writing — review & editing. **JV:** Data curation, Software, Writing — review & editing. **DZ:** Conceptualization, Data curation, Writing — review & editing. **MB:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing — review & editing.



560 *Competing interests.* The authors have no other competing interests to declare.

Acknowledgements. This work made use of the Dutch national e-infrastructure with the support of the SURF Cooperative (grant no. EINF-11758, EINF-9757 & NWO-2024.031). Huub Stoffers and Stefan Wolfsheimer are thanked for their support regarding the use of Snellius. MB acknowledges support from the ERC Advanced Grant scheme (Grant no. 101019185 — GEOWAT). Myrthe Leijnse is thanked for guidance in accessing and using the hotspot polygons. Maisam Mohammadi Dadkan and Vincent Brunst are thanked for their guidance and
565 assistance with hosting data on YODA.



Appendix A: Machine Learning Bias Correction

Metric	Train	Test (Holdout)
RMSE	31.2	32.5
R ²	0.6	0.6
MAE	13.8	14.4

Table A1. Performance metrics of the XGBoost groundwater water heads model on training and independent holdout test data, including root mean squared error (RMSE), coefficient of determination (R^2), and mean absolute error (MAE).

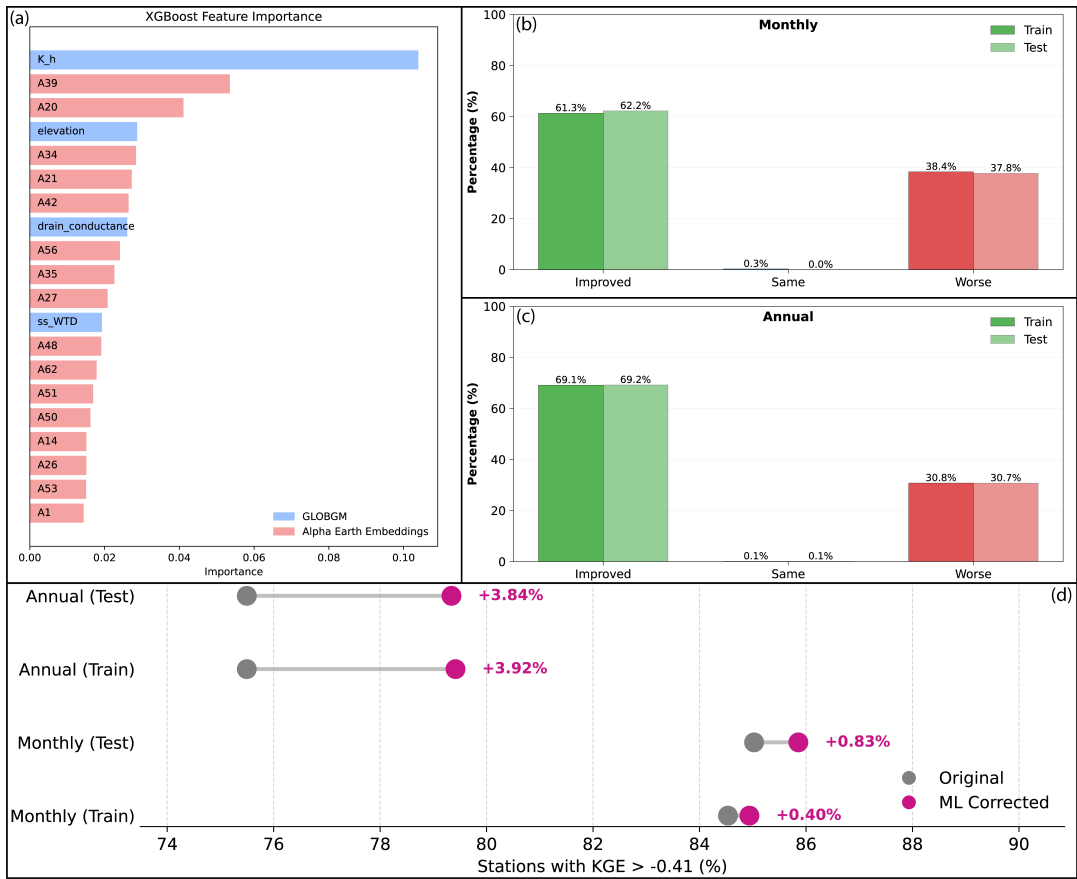


Figure A1. (a) Feature importance of predictors used in the XGBoost models. Percentage of observation wells where bias improved, stayed the same, or worsened at the monthly (b) and annual (c) resolution for the train and test sets when bias corrected. (d) Percentage of observations where KGE-NP exceeds -0.41 before and after bias correction.



References

- Asher, M. J., Croke, B. F. W., Jakeman, A. J., and Peeters, L. J. M.: A Review of Surrogate Models and Their Application to Groundwater Modeling, *Water Resources Research*, 51, 5957–5973, <https://doi.org/10.1002/2015WR016967>, 2015.
- 570 Bähge, A., Vargas, C. R., Lischeid, G., Collenteur, R., Cuthbert, M., Fleckenstein, J., Flörke, M., de Graaf, I., Gnann, S., Hartmann, A., Hugings, X., Moosdorf, N., Wada, Y., Wagener, T., and Reinecke, R.: GROW: A Global Time Series Dataset for Large-Sample Groundwater Studies, *Tech. Rep. EGU25-17713*, Copernicus Meetings, <https://doi.org/10.5194/egusphere-egu25-17713>, 2025.
- Bennett, A., Tran, H., De La Fuente, L., Triplett, A., Ma, Y., Melchior, P., Maxwell, R. M., and Condon, L. E.: Spatio-Temporal Machine Learning for Regional to Continental Scale Terrestrial Hydrology, *Journal of Advances in Modeling Earth Systems*, 16, e2023MS004095, <https://doi.org/10.1029/2023MS004095>, 2024.
- 575 Béraud, T., Claproud, M., and Gloaguen, E.: A Sequential Ensemble Smoother for Multiple Data Assimilation in Hydrogeological Modeling, *Frontiers in Water*, 6, <https://doi.org/10.3389/frwa.2024.1462914>, 2024.
- Berghuijs, W. R., Collenteur, R. A., Jasechko, S., Jaramillo, F., Luijendijk, E., Moeck, C., van der Velde, Y., and Allen, S. T.: Groundwater Recharge Is Sensitive to Changing Long-Term Aridity, *Nature Climate Change*, 14, 357–363, [https://doi.org/10.1038/s41558-024-01953-](https://doi.org/10.1038/s41558-024-01953-z)
- 580 z, 2024.
- Bierkens, M. F. P.: Global Hydrology 2015: State, Trends, and Directions, *Water Resources Research*, 51, 4923–4947, <https://doi.org/10.1002/2015WR017173>, 2015.
- Bierkens, M. F. P. and Wada, Y.: Non-Renewable Groundwater Use and Groundwater Depletion: A Review, *Environmental Research Letters*, 14, 063002, <https://doi.org/10.1088/1748-9326/ab1a5f>, 2019.
- 585 Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., De Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., Van De Giesen, N., Winsemius, H., and Wood, E. F.: Hyper-Resolution Global Hydrological Modelling: What Is next?: “Everywhere and Locally Relevant”, *Hydrological Processes*, 29, 310–320, <https://doi.org/10.1002/hyp.10391>, 2015.
- Bierkens, M. F. P., Sutanudjaja, E. H., and Wanders, N.: Large-Scale Sensitivities of Groundwater and Surface Water to Groundwater Withdrawal, *Hydrology and Earth System Sciences*, 25, 5859–5878, <https://doi.org/10.5194/hess-25-5859-2021>, 2021.
- 590 Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D’Andrea, F., Davini, P., De Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, E., L., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002010, <https://doi.org/10.1029/2019MS002010>, 2020.
- 600 Brown, C. F., Kazmierski, M. R., Pasquarella, V. J., Rucklidge, W. J., Samsikova, M., Zhang, C., Shelhamer, E., Lahera, E., Wiles, O., Ilyushchenko, S., Gorelick, N., Zhang, L. L., Alj, S., Schechter, E., Askay, S., Guinan, O., Moore, R., Boukouvalas, A., and Kohli, P.: AlphaEarth Foundations: An Embedding Field Model for Accurate and Efficient Global Mapping from Sparse Label Data, <https://doi.org/10.48550/ARXIV.2507.22291>, 2025.



- Brun, P., Zimmermann, N. E., Hari, C., Pellissier, L., and Karger, D. N.: CHELSA-BIOCLIM+ A Novel Set of Global Climate-Related
 605 Predictors at Kilometre-Resolution, <https://doi.org/10.16904/ENVIDAT.332>, 2022a.
- Brun, P., Zimmermann, N. E., Hari, C., Pellissier, L., and Karger, D. N.: Global Climate-Related Predictors at Kilometer Resolution for the
 Past and Future, *Earth System Science Data*, 14, 5573–5603, <https://doi.org/10.5194/essd-14-5573-2022>, 2022b.
- Chen, Z., Goldscheider, N., Auler, A., Bakalowicz, M., Broda, S., Drew, D., Hartmann, J., Jiang, G., Moosdorf, N., Richts, A., Ste-
 vanovic, Z., Veni, G., Dumont, A., Aureli, A., Clos, P., and Krombholz, M.: World Karst Aquifer Map (WHYMAP WOKAM),
 610 https://doi.org/10.25928/B2.21_SFKQ-R406, 2017.
- Cochand, M., Molson, J., and Lemieux, J.-M.: Groundwater Hydrogeochemistry in Permafrost Regions, *Permafrost and Periglacial Pro-
 cesses*, 30, 90–103, <https://doi.org/10.1002/ppp.1998>, 2019.
- Condon, L. E., Kollet, S., Bierkens, M. F. P., Fogg, G. E., Maxwell, R. M., Hill, M. C., Fransen, H.-J. H., Verhoef, A., Van Loon, A. F.,
 Sulis, M., and Abesser, C.: Global Groundwater Modeling and Monitoring: Opportunities and Challenges, *Water Resources Research*, 57,
 615 e2020WR029 500, <https://doi.org/10.1029/2020WR029500>, 2021.
- Costantini, M., Colin, J., and Decharme, B.: Projected Climate-Driven Changes of Water Table Depth in the World’s Major Groundwater
 Basins, *Earth’s Future*, 11, e2022EF003 068, <https://doi.org/10.1029/2022EF003068>, 2023.
- Custodio, E.: Intensive Groundwater Development: A Water Cycle Transformation, a Social Revolution, a Management Challenge, in: *Re-
 Thinking Water and Food Security*, CRC Press, ISBN 978-0-429-21299-4, 2010.
- 620 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, *Advances in Water Resources*, 102,
 53–67, <https://doi.org/10.1016/j.advwatres.2017.01.011>, 2017.
- de Graaf, I. E. M., Gleeson, T., (Rens) van Beek, L. P. H., Sutanudjaja, E. H., and Bierkens, M. F. P.: Environmental Flow Limits to Global
 Groundwater Pumping, *Nature*, 574, 90–94, <https://doi.org/10.1038/s41586-019-1594-4>, 2019.
- 625 De Graaf, I. E. M., Marinelli, B., and Liu, S.: Global Analysis of Groundwater Pumping from Increased River Capture, *Environmental
 Research Letters*, 19, 044 064, <https://doi.org/10.1088/1748-9326/ad383d>, 2024.
- Diak, M., Böttcher, M. E., Ehlert Von Ahn, C. M., Hong, W.-L., Kędra, M., Kotwicki, L., Kozirowska-Makuch, K., Kuliński, K., Lepland,
 A., Makuch, P., Sen, A., Winogradow, A., Silberberger, M. J., and Szymczycha, B.: Permafrost and Groundwater Interaction: Current
 State and Future Perspective, *Frontiers in Earth Science*, 11, 1254 309, <https://doi.org/10.3389/feart.2023.1254309>, 2023.
- 630 Döll, P., Hoffmann-Dobrev, H., Portmann, F. T., Siebert, S., Eicker, A., Rodell, M., Strassberg, G., and Scanlon, B. R.: Impact of Water
 Withdrawals from Groundwater and Surface Water on Continental Water Storage Variations, *Journal of Geodynamics*, 59–60, 143–156,
<https://doi.org/10.1016/j.jog.2011.05.001>, 2012.
- Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E.,
 Hamer, P. D., Hirschi, M., Ikonen, J., De Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N.,
 635 Parinussa, R., Pratola, C., Reimer, C., Van Der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Mois-
 ture for Improved Earth System Understanding: State-of-the Art and Future Directions, *Remote Sensing of Environment*, 203, 185–215,
<https://doi.org/10.1016/j.rse.2017.07.001>, 2017.
- Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova,
 E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M.,
 640 Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J.,
 Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton,



- M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002015, <https://doi.org/10.1029/2019MS002015>, 2020.
- 645 Eccles, R., Syktus, J., Trancoso, R., Chapman, S., Wasko, C., Evans, J. P., Thatcher, M., Di Virgilio, G., and Stassen, C.: Substantial Increases in Future Precipitation Extremes—Insights from a Large Ensemble of Downscaled CMIP6 Models, *npj Natural Hazards*, 2, 60, <https://doi.org/10.1038/s44304-025-00107-1>, 2025.
- 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.
- 650 Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., and Otero-Casal, C.: Hydrologic Regulation of Plant Rooting Depth, *Proceedings of the National Academy of Sciences*, 114, 10 572–10 577, <https://doi.org/10.1073/pnas.1712381114>, 2017.
- Ford, T. W. and Frauenfeld, O. W.: Surface–Atmosphere Moisture Interactions in the Frozen Ground Regions of Eurasia, *Scientific Reports*, 6, 19 163, <https://doi.org/10.1038/srep19163>, 2016.
- Gädeke, A., Krysanova, V., Aryal, A., Chang, J., Grillakis, M., Hanasaki, N., Koutroulis, A., Pokhrel, Y., Satoh, Y., Schaphoff, S., Müller Schmied, H., Stacked, T., Tang, Q., Wada, Y., and Thonicke, K.: Performance Evaluation of Global Hydrological Models in Six Large Pan-Arctic Watersheds, *Climatic Change*, 163, 1329–1351, <https://doi.org/10.1007/s10584-020-02892-2>, 2020.
- 655 Garrick, D. E., Hall, J. W., Dobson, A., Damania, R., Grafton, R. Q., Hope, R., Hepburn, C., Bark, R., Boltz, F., De Stefano, L., O'Donnell, E., Matthews, N., and Money, A.: Valuing Water for Sustainable Development, *Science*, 358, 1003–1005, <https://doi.org/10.1126/science.aao4942>, 2017.
- 660 Gleeson, T., Cuthbert, M., Ferguson, G., and Perrone, D.: Global Groundwater Sustainability, Resources, and Systems in the Anthropocene, *Annual Review of Earth and Planetary Sciences*, 48, 431–463, <https://doi.org/10.1146/annurev-earth-071719-055251>, 2020.
- Gnann, S., Reinecke, R., Stein, L., Wada, Y., Thiery, W., Müller Schmied, H., Satoh, Y., Pokhrel, Y., Ostberg, S., Koutroulis, A., Hanasaki, N., Grillakis, M., Gosling, S. N., Burek, P., Bierkens, M. F. P., and Wagener, T.: Functional Relationships Reveal Differences in the Water Cycle Representation of Global Water Models, *Nature Water*, 1, 1079–1090, <https://doi.org/10.1038/s44221-023-00160-y>, 2023.
- 665 Gruber, S.: Derivation and Analysis of a High-Resolution Estimate of Global Permafrost Zonation, *The Cryosphere*, 6, 221–233, <https://doi.org/10.5194/tc-6-221-2012>, 2012.
- Hagemann, S. and Gates, L. D.: Improving a Subgrid Runoff Parameterization Scheme for Climate Models by the Use of High Resolution Data Derived from Satellite Observations, *Climate Dynamics*, 21, 349–359, <https://doi.org/10.1007/s00382-003-0349-x>, 2003.
- Hartmann, J. and Moosdorf, N.: The New Global Lithological Map Database GLiM: A Representation of Rock Properties at the Earth Surface, *Geochemistry, Geophysics, Geosystems*, 13, 2012GC004 370, <https://doi.org/10.1029/2012GC004370>, 2012.
- 670 Huggins, X., Gleeson, T., Castilla-Rho, J., Holley, C., Re, V., and Famiglietti, J. S.: Groundwater Connections and Sustainability in Social-Ecological Systems, *Groundwater*, 61, 463–478, <https://doi.org/10.1111/gwat.13305>, 2023.
- IGRAC: The Global Groundwater Monitoring Network (GGMN), <https://doi.org/10.58154/6Z0Y-DA34>, 2024.
- Ilstedt, U., Bargués Tobella, A., Bazié, H. R., Bayala, J., Verbeeten, E., Nyberg, G., Sanou, J., Benegas, L., Murdiyarso, D., Laudon, H., Sheil, D., and Malmer, A.: Intermediate Tree Cover Can Maximize Groundwater Recharge in the Seasonally Dry Tropics, *Scientific Reports*, 6, 21 930, <https://doi.org/10.1038/srep21930>, 2016.
- Jasechko, S. and Perrone, D.: Global Groundwater Wells at Risk of Running Dry, *Science*, 372, 418–421, <https://doi.org/10.1126/science.abc2755>, 2021.



- Jasechko, S., Seybold, H., Perrone, D., Fan, Y., Shamsudduha, M., Taylor, R. G., Fallatah, O., and Kirchner, J. W.: Rapid Groundwater
 680 Decline and Some Cases of Recovery in Aquifers Globally, *Nature*, 625, 715–721, <https://doi.org/10.1038/s41586-023-06879-8>, 2024.
- Kendall, M. G.: Rank Correlation Methods, Griffin, London, 4. ed., 2. impr edn., ISBN 978-0-85264-199-6, 1975.
- Khatiwala, S.: Efficient Spin-up of Earth System Models Using Sequence Acceleration, *Science Advances*, 10, eadn2839,
<https://doi.org/10.1126/sciadv.adn2839>, 2024.
- Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical Note: Inherent Benchmark or Not? Comparing Nash–Sutcliffe and Kling–Gupta
 685 Efficiency Scores, *Hydrology and Earth System Sciences*, 23, 4323–4331, <https://doi.org/10.5194/hess-23-4323-2019>, 2019.
- Kollet, S., Belleflamme, A., Condon, L., Fahad, M., Goergen, K., Maxwell, R., and Naz, B.: Global Groundwater Modeling:
 Proof-of-concept of 3D Variably Saturated Flow Simulation at Kilometer Resolution, *Journal of Hydrology X*, 30, 100213,
<https://doi.org/10.1016/j.hydroa.2025.100213>, 2026.
- Kuang, X., Liu, J., Scanlon, B. R., Jiao, J. J., Jasechko, S., Lancia, M., Biskaborn, B. K., Wada, Y., Li, H., Zeng, Z., Guo, Z., Yao, Y., Gleeson,
 690 T., Nicot, J.-P., Luo, X., Zou, Y., and Zheng, C.: The Changing Nature of Groundwater in the Global Water Cycle, *Science*, 383, eadf0630,
<https://doi.org/10.1126/science.adf0630>, 2024.
- Kupzig, J., Kupzig, N., and Flörke, M.: Regionalization in Global Hydrological Models and Its Impact on Runoff Simulations: A Case Study
 Using WaterGAP3 (v 1.0.0), *Geoscientific Model Development*, 17, 6819–6846, <https://doi.org/10.5194/gmd-17-6819-2024>, 2024.
- Kuzma, S., Bierkens, M. F. P., Lakshman, S., Luo, T., Saccoccia, L., Sutanudjaja, E. H., and Beek, R. V.: Aqueduct 4.0: Updated Decision-
 695 Relevant Global Water Risk Indicators, Tech. rep., World Resources Institute, Washington, DC, 2023.
- Lange, S. and Büchner, M.: ISIMIP3b Bias-Adjusted Atmospheric Climate Input Data, <https://doi.org/10.48364/ISIMIP.842396.1>, 2021.
- Lange, S., Menz, C., Gleixner, S., Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Müller Schmied, H., Hersbach, H., Buontempo, C.,
 and Cagnazzo, C.: WFDE5 over Land Merged with ERA5 over the Ocean (W5E5 v2.0), <https://doi.org/10.48364/ISIMIP.342217>, 2021.
- Lehner, B., Verdin, K., and Jarvis, A.: New Global Hydrography Derived From Spaceborne Elevation Data, *Eos, Transactions American
 700 Geophysical Union*, 89, 93–94, <https://doi.org/10.1029/2008EO100001>, 2008.
- Leijnse, M., F P Bierkens, M., H M Gommans, K., Lin, D., Tait, A., and Wanders, N.: Key Drivers and Pressures of Global Water Scarcity
 Hotspots, *Environmental Research Letters*, 19, 054 035, <https://doi.org/10.1088/1748-9326/ad3c54>, 2024.
- Maples, S. R., Fogg, G. E., and Maxwell, R. M.: Modeling Managed Aquifer Recharge Processes in a Highly Heterogeneous, Semi-Confined
 Aquifer System, *Hydrogeology Journal*, 27, 2869–2888, <https://doi.org/10.1007/s10040-019-02033-9>, 2019.
- 705 Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler,
 S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns,
 T., Jimenez-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornbluh, L.,
 Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D.,
 Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H.,
 710 Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., Von Storch,
 J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth
 System Model Version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO₂, *Journal of Advances in Modeling Earth Systems*, 11,
 998–1038, <https://doi.org/10.1029/2018MS001400>, 2019.
- McKenzie, J. M., Kurylyk, B. L., Walvoord, M. A., Bense, V. F., Fortier, D., Spence, C., and Grenier, C.: Invited Perspective: What Lies
 715 beneath a Changing Arctic?, *The Cryosphere*, 15, 479–484, <https://doi.org/10.5194/tc-15-479-2021>, 2021.



- Moeck, C., Grech-Cumbo, N., Podgorski, J., Bretzler, A., Gurdak, J. J., Berg, M., and Schirmer, M.: Data for: A Global-Scale Dataset of Direct Natural Groundwater Recharge Rates: A Review of Variables, Processes and Relationships, <https://doi.org/10.25678/0001NG>, 2020a.
- Moeck, C., Grech-Cumbo, N., Podgorski, J., Bretzler, A., Gurdak, J. J., Berg, M., and Schirmer, M.: A Global-Scale Dataset of Direct Natural Groundwater Recharge Rates: A Review of Variables, Processes and Relationships, *Science of The Total Environment*, 717, 137 042, <https://doi.org/10.1016/j.scitotenv.2020.137042>, 2020b.
- Mohan, C., Western, A. W., Wei, Y., and Saft, M.: Predicting Groundwater Recharge for Varying Land Cover and Climate Conditions – a Global Meta-Study, *Hydrology and Earth System Sciences*, 22, 2689–2703, <https://doi.org/10.5194/hess-22-2689-2018>, 2018.
- Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., Tomkins-Tinch, C. H., Sochat, V., Forster, J., Lee, S., Twardziok, S. O., Kanitz, A., Wilm, A., Holtgrewe, M., Rahmann, S., Nahnsen, S., and Köster, J.: Sustainable Data Analysis with Snakemake, *F1000Research*, 10, 33, <https://doi.org/10.12688/f1000research.29032.2>, 2021.
- Niazi, H., Wild, T. B., Turner, S. W. D., Graham, N. T., Hejazi, M., Msangi, S., Kim, S., Lamontagne, J. R., and Zhao, M.: Global Peak Water Limit of Future Groundwater Withdrawals, *Nature Sustainability*, 7, 413–422, <https://doi.org/10.1038/s41893-024-01306-w>, 2024.
- Otoo, N. G., Sutanudjaja, E. H., van Vliet, M. T. H., Schipper, A. M., and Bierkens, M. F. P.: Mapping Groundwater-Dependent Ecosystems Using a High-Resolution Global Groundwater Model, *Hydrology and Earth System Sciences*, 29, 2153–2165, <https://doi.org/10.5194/hess-29-2153-2025>, 2025.
- Pawusch, L., Scheurer, S., Nowak, W., and Maxwell, R. M.: HydroStartML: A Combined Machine Learning and Physics-Based Approach to Reduce Hydrological Model Spin-up Time, *Advances in Water Resources*, 206, 105 124, <https://doi.org/10.1016/j.advwatres.2025.105124>, 2025.
- Perrone, D. and Jasechko, S.: Deeper Well Drilling an Unsustainable Stopgap to Groundwater Depletion, *Nature Sustainability*, 2, 773–782, <https://doi.org/10.1038/s41893-019-0325-z>, 2019.
- Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., and Rossiter, D.: SoilGrids 2.0: Producing Soil Information for the Globe with Quantified Spatial Uncertainty, *SOIL*, 7, 217–240, <https://doi.org/10.5194/soil-7-217-2021>, 2021.
- Pool, S., Vis, M., and Seibert, J.: Evaluating Model Performance: Towards a Non-Parametric Variant of the Kling-Gupta Efficiency, *Hydrological Sciences Journal*, 63, 1941–1953, <https://doi.org/10.1080/02626667.2018.1552002>, 2018.
- Reinecke, R., Foglia, L., Mehl, S., Trautmann, T., Cáceres, D., and Döll, P.: Challenges in Developing a Global Gradient-Based Groundwater Model (G³M v1.0) for the Integration into a Global Hydrological Model, *Geoscientific Model Development*, 12, 2401–2418, <https://doi.org/10.5194/gmd-12-2401-2019>, 2019.
- Reinecke, R., Gnann, S., Stein, L., Bierkens, M., De Graaf, I., Gleeson, T., OudeEssink, G., Sutanudjaja, E., Ruz-Vargas, C., Verkaik, J., and Wagener, T.: Considerable Gaps in Our Global Knowledge of Potential Groundwater Accessibility, <https://doi.org/10.31223/X5SM0R>, 2023.
- Reinecke, R., Gnann, S., Stein, L., Bierkens, M., De Graaf, I., Gleeson, T., Essink, G. O., Sutanudjaja, E. H., Ruz Vargas, C., Verkaik, J., and Wagener, T.: Uncertainty in Model Estimates of Global Groundwater Depth, *Environmental Research Letters*, 19, 114 066, <https://doi.org/10.1088/1748-9326/ad8587>, 2024.
- Reinecke, R., Akhter, T., Bähge, A., Dietrich, R., Gnann, S., Gosling, S. N., Grogan, D., Hartmann, A., Kollet, S., Kumar, R., Lammers, R., Liu, S., Liu, Y., Moosdorf, N., Naz, B., Nazari, S., Orazulike, C., Pokhrel, Y., Schewe, J., Smilovic, M., Stokal, M., Thiery, W., Wada, Y., Zuidema, S., and De Graaf, I.: The ISIMIP Groundwater Sector: A Framework for Ensemble Modeling of Global Change Impacts on Groundwater, *Geoscientific Model Development*, 19, 523–542, <https://doi.org/10.5194/gmd-19-523-2026>, 2026.



- Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., and Lo, M.-H.: Emerging Trends in Global
 755 Freshwater Availability, *Nature*, 557, 651–659, <https://doi.org/10.1038/s41586-018-0123-1>, 2018.
- Saccò, M., Mammola, S., Altermatt, F., Alther, R., Bolpagni, R., Brancelj, A., Brankovits, D., Fišer, C., Gerovasileiou, V., Griebler, C.,
 Guareschi, S., Hose, G. C., Korbel, K., Lictevout, E., Malard, F., Martínez, A., Niemiller, M. L., Robertson, A., Tanalgo, K. C., Bichuette,
 M. E., Borko, Š., Brad, T., Campbell, M. A., Cardoso, P., Celico, F., Cooper, S. J. B., Culver, D., Di Lorenzo, T., Galassi, D. M. P., Guzik,
 M. T., Hartland, A., Humphreys, W. F., Ferreira, R. L., Lunghi, E., Nizzoli, D., Perina, G., Raghavan, R., Richards, Z., Reboleira, A. S.
 760 P. S., Rohde, M. M., Fernández, D. S., Schmidt, S. I., Van Der Heyde, M., Weaver, L., White, N. E., Zagamajster, M., Hogg, I., Ruhi, A.,
 Gagnon, M. M., Allentoft, M. E., and Reinecke, R.: Groundwater Is a Hidden Global Keystone Ecosystem, *Global Change Biology*, 30,
 e17066, <https://doi.org/10.1111/gcb.17066>, 2024.
- Sellar, A. A., Walton, J., Jones, C. G., Wood, R., Abraham, N. L., Andrejczuk, M., Andrews, M. B., Andrews, T., Archibald, A. T., De Mora,
 L., Dyson, H., Elkington, M., Ellis, R., Florek, P., Good, P., Gohar, L., Haddad, S., Hardiman, S. C., Hogan, E., Iwi, A., Jones, C. D.,
 765 Johnson, B., Kelley, D. I., Kettleborough, J., Knight, J. R., Köhler, M. O., Kuhlbrodt, T., Liddicoat, S., Linova-Pavlova, I., Mizieliński,
 M. S., Morgenstern, O., Mulcahy, J., Neininger, E., O'Connor, F. M., Petrie, R., Ridley, J., Rioual, J.-C., Roberts, M., Robertson,
 E., Rumbold, S., Seddon, J., Shepherd, H., Shim, S., Stephens, A., Teixeira, J. C., Tang, Y., Williams, J., Wiltshire, A., and Griffiths,
 P. T.: Implementation of U.K. Earth System Models for CMIP6, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001946,
<https://doi.org/10.1029/2019MS001946>, 2020.
- 770 Sen, P. K.: Estimates of the Regression Coefficient Based on Kendall's Tau, *Journal of the American Statistical Association*, 63, 1379–1389,
<https://doi.org/10.1080/01621459.1968.10480934>, 1968.
- Servén, D., Brummitt, C., Abedi, H., and hlink: Dswah/pyGAM: V0.8.0, <https://doi.org/10.5281/zenodo.1476122>, 2018.
- Somers, L. D. and McKenzie, J. M.: A Review of Groundwater in High Mountain Environments, *WIREs Water*, 7, e1475,
<https://doi.org/10.1002/wat2.1475>, 2020.
- 775 Sutanudjaja, E. H.: HYPFLOWSCI6: HYdrological Projection of Future gLObal Water States with CMIP6, <https://doi.org/10.24416/UU01-YM7A5H>, 2024.
- Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., Drost, N., van der Ent, R. J., de Graaf, I. E. M., Hoch, J. M., de
 Jong, K., Karssenber, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M. W., Vannamettee, E., Wissler, D., and Bierkens, M.
 F. P.: PCR-GLOBWB 2: A 5 Arcmin Global Hydrological and Water Resources Model, *Geoscientific Model Development*, 11,
 780 2429–2453, <https://doi.org/10.5194/gmd-11-2429-2018>, 2018.
- Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M.,
 Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., Gurdak,
 J. J., Allen, D. M., Shamsudduha, M., Hiscock, K., Yeh, P. J.-F., Holman, I., and Treidel, H.: Ground Water and Climate Change, *Nature
 Climate Change*, 3, 322–329, <https://doi.org/10.1038/nclimate1744>, 2013.
- 785 Theil, H.: A Rank-Invariant Method of Linear and Polynomial Regression Analysis, *Indagationes Mathematicae*, 12, 386–392, 1950.
- Todini, E.: The ARNO Rainfall—Runoff Model, *Journal of Hydrology*, 175, 339–382, [https://doi.org/10.1016/S0022-1694\(96\)80016-3](https://doi.org/10.1016/S0022-1694(96)80016-3),
 1996.
- UN Water, ed.: Groundwater Making the Invisible Visible, no. 2022 in *The United Nations World Water Development Report*, UNESCO,
 Paris, ISBN 978-92-3-100507-7, 2022.
- 790 Vega-Briones, J., Sutanudjaja, E. H., De Jong, S., and Wanders, N.: Modelling Groundwater Hydrological Drought and Its Recovery Given
 Natural and Anthropogenic Scenarios in South America, *Hydrological Processes*, 38, e15340, <https://doi.org/10.1002/hyp.15340>, 2024.



- Verkaik, J., Sutanudjaja, E. H., Oude Essink, G. H. P., Lin, H. X., and Bierkens, M. F. P.: GLOBGM v1.0: A Parallel Implementation of a 30 Arcsec PCR-GLOBWB-MODFLOW Global-Scale Groundwater Model, *Geoscientific Model Development*, 17, 275–300, <https://doi.org/10.5194/gmd-17-275-2024>, 2024.
- 795 Vörösmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B.: Global Water Resources: Vulnerability from Climate Change and Population Growth, *Science*, 289, 284–288, <https://doi.org/10.1126/science.289.5477.284>, 2000.
- Wada, Y. and Bierkens, M. F. P.: Sustainability of Global Water Use: Past Reconstruction and Future Projections, *Environmental Research Letters*, 9, 104 003, <https://doi.org/10.1088/1748-9326/9/10/104003>, 2014.
- Wada, Y., Van Beek, L. P. H., Wanders, N., and Bierkens, M. F. P.: Human Water Consumption Intensifies Hydrological Drought Worldwide, *Environmental Research Letters*, 8, 034 036, <https://doi.org/10.1088/1748-9326/8/3/034036>, 2013.
- 800 Wang, C., Wu, Q., Weimer, M., and Zhu, E.: FLAML: A Fast and Lightweight AutoML Library, <https://doi.org/10.48550/ARXIV.1911.04706>, 2019.
- Werndl, C.: Initial-Condition Dependence and Initial-Condition Uncertainty in Climate Science, *The British Journal for the Philosophy of Science*, 70, 953–976, <https://doi.org/10.1093/bjps/axy021>, 2019.
- 805 West, C., Rosolem, R., MacDonald, A. M., Cuthbert, M. O., and Wagener, T.: Understanding Process Controls on Groundwater Recharge Variability across Africa through Recharge Landscapes, *Journal of Hydrology*, 612, 127 967, <https://doi.org/10.1016/j.jhydrol.2022.127967>, 2022.
- White, J. T.: A Model-Independent Iterative Ensemble Smoother for Efficient History-Matching and Uncertainty Quantification in Very High Dimensions, *Environmental Modelling & Software*, 109, 191–201, <https://doi.org/10.1016/j.envsoft.2018.06.009>, 2018.
- 810 Wood, E. F., Roundy, J. K., Troy, T. J., Van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., De Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., Van De Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution Global Land Surface Modeling: Meeting a Grand Challenge for Monitoring Earth’s Terrestrial Water, *Water Resources Research*, 47, 2010WR010 090, <https://doi.org/10.1029/2010WR010090>, 2011.
- Xie, J., Liu, X., Jasechko, S., Berghuijs, W. R., Wang, K., Liu, C., Reichstein, M., Jung, M., and Koirala, S.: Majority of Global River Flow Sustained by Groundwater, *Nature Geoscience*, 17, 770–777, <https://doi.org/10.1038/s41561-024-01483-5>, 2024.
- 815 Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., and Ishii, M.: The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component, *Journal of the Meteorological Society of Japan*. Ser. II, 97, 931–965, <https://doi.org/10.2151/jmsj.2019-051>, 2019.
- 820 Zamrsky, D., Oude Essink, G. H. P., and Bierkens, M. F. P.: Global Impact of Sea Level Rise on Coastal Fresh Groundwater Resources, *Earth’s Future*, 12, e2023EF003 581, <https://doi.org/10.1029/2023EF003581>, 2024.
- Zamrsky, D., Ruzzante, S., Compare, K., Kretschmer, D., Zipper, S., Befus, K. M., Reinecke, R., Pasner, Y., Gleeson, T., Jordan, K., Cuthbert, M., Castronova, A. M., Wagener, T., and Bierkens, M. F. P.: Current Trends and Biases in Groundwater Modelling Using the Community-Driven Groundwater Model Portal (GroMoPo), *Hydrogeology Journal*, 33, 355–366, <https://doi.org/10.1007/s10040-025-02882-7>, 2025.
- 825 Zell, W. O. and Sanford, W. E.: Calibrated Simulation of the Long-Term Average Surficial Groundwater System and Derived Spatial Distributions of Its Characteristics for the Contiguous United States, *Water Resources Research*, 56, e2019WR026 724, <https://doi.org/10.1029/2019WR026724>, 2020.