Coupling a large-scale hydrological model (CWatM v1.1) with a high-resolution groundwater flow model (MODFLOW 6) to assess the impact of irrigation at regional scale

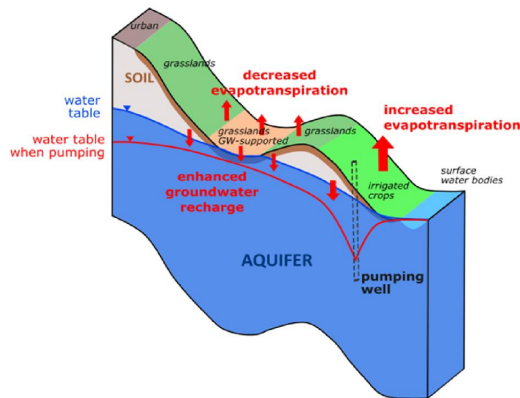
Luca Guillaumot[1], Mikhail Smilovic[1], Peter Burek[1], Jens de Bruijn[1,2], Peter Greve[1], Taher Kahil[1], Yoshihide Wada[1]

[1] Water Security Research Group, Biodiversity and Natural Resources Program, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
[2] Institute for Environmental Studies, VU University, De Boelelaan 1087, 1081HV, Amsterdam, The Netherlands

Correspondence to: Luca Guillaumot (guillaumot@iiasa.ac.at)(guillaumot@iiasa.ac.at)

Short summary. We apply fordevelop and test the first time a large-scale hydrological model at regional scale with a very high spatial resolution and includingthat includes a water management and groundwater flow model. The model is used to inferThis study infers the impact of surface and groundwater-based irrigation on groundwater recharge and on evapotranspiration and groundwater recharge in both irrigated and non-irrigated areas. WaterWe argue that water table time-fluctuations recorded in boreholes are can be used as validation data if water management is well reproduced; however, reproducing water table depth would require finerimplemented and spatial resolution (<is ≤ 100 m)..

Abstract. In the context of changing climate and increasing water demand, large-scale hydrological models are helpful for understanding and projecting future water resources across scales. Groundwater is a critical freshwater resource and strongly controls river flow throughout the year. It is also essential for ecosystems and contributes to evapotranspiration, resulting in climate feedback. However, groundwater systems worldwide are quite diverse, including thick multi-layer aquifers and thin heterogeneous aquifers. Recently, efforts have been made to improve the representation of groundwater systems in large-scale hydrological models. The evaluation of the accuracy of these model outputs is challenging because: (1) they are applied at much coarser resolutions than hillslope scale, (2) they simplify geological structures generally known at local scale, and (3) they do not adequately include local water management practices (mainly groundwater pumping). Here, we apply a large-scale hydrological model (CWatM), coupled with the groundwater flow model MODFLOW, in two different climatic, geological, and socioeconomic regions, the Seewinkel area (Austria) and the Bhima basin (India). The coupled model enables simulation of the impact of the water table on groundwater-soils and groundwater-rivers exchanges, groundwater recharge through leaking canals, and groundwater pumping. This regional scale analysis enables assessment of the model's ability to simulate water tables at fine spatial resolutions (1 km for CWatM, 100–250 m for MODFLOW) and when groundwater pumping is well

estimated. Evaluating large-scale models remains challenging, but the results show that the reproduction of (1) average water table fluctuations and (2) water table depths without bias can be a benchmark objective of such models. We found that ~~subgrid~~grid resolution ~~(river incision)~~ is the main factor that affects water table depth bias because it smooths river incision, while pumping affects time fluctuations. Finally, we use the model to assess the impact of groundwater-based irrigation pumping on evapotranspiration, groundwater recharge, and water table observations from boreholes.

1 Introduction

Regional and large-scale hydrological models are often used to assess water resource trajectories under different scenarios for climate change, socioeconomic development, and water management. Despite extensive work, at least three challenges persist: appropriately representing groundwater dynamics and flow (Gleeson et al., 2021; Kollet and Maxwell, 2008; Reinecke et al., 2020; Sutanudjaja et al., 2011; Vergnes et al., 2014), including human impact (Hanasaki et al., 2017; Wada et al., 2017, 2014), and improving spatial resolution (Bierkens, 2015; Fan et al., 2019; Wood et al., 2011). The last point ~~is of interest for locally relevant applications, but it could~~can be replaced in part by ~~"~~'representing water flows driven by small-scale topography', and is of interest for processes ~~driven by topography"~~at large-scale and locally relevant applications. Several large-scale hydrological models include representations of groundwater flow between grid cells and interactions among groundwater, soils, and surface water bodies, such as CWatM (Burek et al., 2020), LISFLOOD (Trichakis et al., 2017), ORCHIDEE (Verbeke et al., 2019), ParFlow-CLM (Keune et al., 2016; Maxwell et al., 2015), PCR-GLOBWB (de Graaf et al., 2017; Sutanudjaja et al., 2018), ISBA-CTRIP (Decharme et al., 2019), VIC (Scheidegger et al., 2021), LEAFHYDRO (Martínez-de la Torre and Miguez-Macho, 2019), and WaterGAP (Reinecke et al., 2019a). These models differ somewhat in their implementation, including the physical representation and parametrization of the groundwater. Currently, developments are

oriented towards hyper-resolution models (less than or around a 1 km grid) and representing water management (Hanasaki et al., 2022). However, large-scale model resolutions (~10–50 km) remain much coarser than the hillslope-scale controlling hydrologic processes, as hypothesized by Fan et al. (2019) and Swenson et al. (2019). In addition, coarse resolutions of groundwater representation tend to smooth hydraulic gradients, leading to unrealistic aquifer properties (Shrestha et al., 2018) and potentially ~~underestimating~~to underestimate water table depth drawdowns due to groundwater pumping because withdrawals are applied to entire grid cells instead to be applied to punctual boreholes.

A proper representation of groundwater is essential to consider lateral groundwater exchanges between grid cells; otherwise, they remain connected only through the river or drainage network. Darcy's law links hydraulic head gradient, hydraulic conductivity, and aquifer thickness and can describe this lateral groundwater flow. Some studies have already highlighted the contribution of lateral groundwater flow to regions and basins (Krakauer et al., 2014; Schaller and Fan, 2009). Groundwater flows redistribute water from hillslope to regional scale leading to evapotranspiration variability between upstream and downstream areas (Condon et al., 2020; Condon and Maxwell, 2019; Famiglietti and Wood, 1994; Fan and Miguez-Macho, 2010; Keune et al., 2016; Miguez-Macho and Fan, 2012; Swenson et al., 2019). These studies highlight that evapotranspiration is impacted to a greater degree during dry seasons and where the water table is shallow. Oversimplifying lateral water redistribution and exchanges between soils and groundwater may also induce nonnegligible evapotranspiration biases (Koirala et al., 2019; Martínez-de la Torre and Miguez-Macho, 2019; Rouholahnejad Freund and Kirchner, 2017; Wang et al., 2018). The impact of the depth of the groundwater level on evapotranspiration and consequently on net groundwater recharge (recharge minus capillary flux) has also been demonstrated by Szilagyi et al. (2013) based on observed depth to groundwater and evapotranspiration estimated from MODIS data and by Koirala et al. (2017) based on simulated depth to groundwater and remote sensing data. Finally, reproducing soil moisture drainage, capillary rise, and baseflow more accurately depends on properly representing groundwater depth and time fluctuation.

Recent studies based on different models (de Graaf et al., 2017; Fan et al., 2013; Martínez-de la Torre and Miguez-Macho, 2019; Maxwell et al., 2015; Sutanudjaja et al., 2014, 2011; Vergnes et al., 2020) have compared simulated and observed water tables at continental and regional scales; these studies have argued that the main spatial trends were well reproduced. However, some of these studies have acknowledged that water table depth is not well reproduced, given the coarse spatial resolution (including problems of the spatial representativity of the boreholes and potential bias sampling) and the lack of representation of water management within the models (Fan et al., 2013; Maxwell et al., 2015; Reinecke et al., 2020). This raises the question of the reliability of such models in terms of parametrization and application (Gleeson et al., 2021). For example, regional and continental scale models often assume very simple groundwater pumping schemes or no pumping at all, and they consider that these withdrawals simply leave the system (Vergnes et al., 2020; Surinaidu et al., 2013; Martínez-de la Torre and Miguez-Macho, 2019). Hanasaki et al. (2022) demonstrated the difficulty of introducing regional water management schemes into large-scale hydrological models. Therefore, the interaction between human water management and water availability is a critical aspect of large-scale hydrological models. Humans strongly impact natural water fluxes, affecting available water resources (Keune et al., 2018; Taylor et al., 2012). Irrigation sustained by groundwater significantly impacts the water cycle

in several regions (Cao et al., 2016; Dalin et al., 2017; Gleeson et al., 2012; Keune et al., 2018; Siebert et al., 2010; Taylor et al., 2012). To predict future water resources, it is necessary to decipher climate and human contributions contained within the space and time variability of hydrological signals. Large-scale hydrological models have been developed to estimate water availability and water use from surface water bodies and groundwater at a global scale (Döll et al., 2014; Wada et al., 2014).

90 Recently, Hanasaki et al. (2017) improved the representation of human interventions in H08 model and showed more realistic river discharges and terrestrial water storage anomalies for several huge basins. Sadki et al. (2022) explored the possibility of improving dams representation in ISBA-CTRIP model applied over Spain. Long et al. (2020) studied the impact of the south-to-north water diversion in China on groundwater pumping using CWatM. Using a global scale hydrological model coupled with MODFLOW, de Graaf et al. (2019) reproduced the water table drawdown dynamic caused by pumping and its impact on

95 rivers. ~~They identified that rivers reach their environmental flow limit~~They identified that rivers reach their environmental flow limit (meaning that groundwater baseflow supporting rivers fails below its 10th percentile as suggested by Gleeson and Richter, 2018) before substantial groundwater depletion occurs.

Studying and simulating regions with significant use of ground and surface waters requires accounting for the management and linkages between both sources of water. Surface and groundwater management are fundamentally connected. Water

100 demand that is satisfied with surface water stored in reservoirs and delivered through pipes or canals may be supplemented with groundwater when the timing or volume of delivery does not coincide with the need. Distribution networks, including urban pipes and agricultural canals, may encourage groundwater recharge through aging and leaking infrastructure. Further, it is necessary to understand who has access to surface water to know where groundwater is the only source that can satisfy the demand. Some farmers will withdraw river water in the rainy season and otherwise depend exclusively on groundwater or

105 leave fields fallow in the dry season. In addition, to understand groundwater's spatial and temporal use it is necessary to appreciate the spatial and temporal distribution of crop water needs.

In this paper, we present refinements on both fronts relative to previous studies, by applying the model at finer resolutions and accounting for water management and comparing simulations with observed water table fluctuations and water table depths. Our new water management representation includes estimating irrigation demand and automatic supply from canals (with

110 potential leakage recharging groundwater), reservoirs, and/or groundwater (Smilovic et al., 2022). To this end, we coupled a high-resolution version of CWatM (~1 km resolution) with MODFLOW implemented at high resolutions of 100 and 250 m. These model versions are used in the Seewinkel region (Austria) and the Bhima basin (India), extending over 573 and 46000 km^2, respectively. Studying groundwater processes at regional scale allows for better calibration and validation of models, at a high resolution, based on observed groundwater levels and estimated groundwater pumping. The Seewinkel area is much

115 smaller and less complex than the Bhima basin regarding the aquifer, and less anthropized regarding water management. Comparing the two regions is of interest to evaluate how geological (as well as geomorphological) complexity constrains the model's ability to reproduce water tables. The comparison should also help clarify the impact of two levels of water management on the water cycle.

The remainder of this paper is organized as follows. We first describe the CWatM model (sections 2.1 and 2.2). Section 3 describes the two study regions. Model performance (section 4) is evaluated based on the water table observed in the monitoring well networks. Then we describe the experiment to assess the impact of groundwater-based irrigation on water tables and the water cycle (section 5). The results show the model's ability to reproduce temporal and spatial variability of water table depth (section 6.1). Then the experiment shows how different water cycle components are impacted by irrigation (sections 6.2 and 6.3). Finally, we discuss the most critical factors affecting the model's accuracy (section 7).

2 Method

2.1 The CWatM model

CWatM (Community Water Model) is a distributed hydrological model that can be implemented at regional to global scales (Burek et al., 2020). The model is developed with the Python programming language and is open-access (https://cwatm.iiasa.ac.at/). CWatM aims to reproduce the main hydrological processes, including water management. In this study, CWatM is applied in two regions and coupled with a finer-resolution groundwater flow model MODFLOW (Harbaugh, 2005). CWatM is used at a spatial resolution of 1 km in Seewinkel and ~0.9 km (30 arcsec) in Bhima. Because the Bhima basin covers a large area, CWatM is applied in a geographic projection system (WGS84), while square cells are used for Seewinkel (Lambert Azimuthal Equal Area projection system ETRS89). MODFLOW is applied at 100 m in Seewinkel and 250 m in Bhima, using square cells in both regions.

The following sections describe more extensively how groundwater processes and pumping are modeled. The coupled CWatM-MODFLOW uses a computationally efficient approach combining MODFLOW 6's Basic Model Interface (BMI) (Hughes et al., 2017) and Python packages Flopy (Bakker et al., 2016) and xmipy (Russcher et al., 2020). The coupling of CWatM and groundwater is illustrated in Figure 1. In particular, it is necessary to use high-resolution groundwater models, as topography controls the variability in the water table. For example, Figure 1 contains an illustration of 1 CWatM cell containing 100 MODFLOW cells (10 shown within the cross-section), and it highlights that different land cover types are considered in each CWatM cell. Soil processes that occur within each land cover part are modeled independently before being averaged according to their respective coverage area within the CWatM cell. This approach allows consideration of the subgrid variability without reducing resolution.

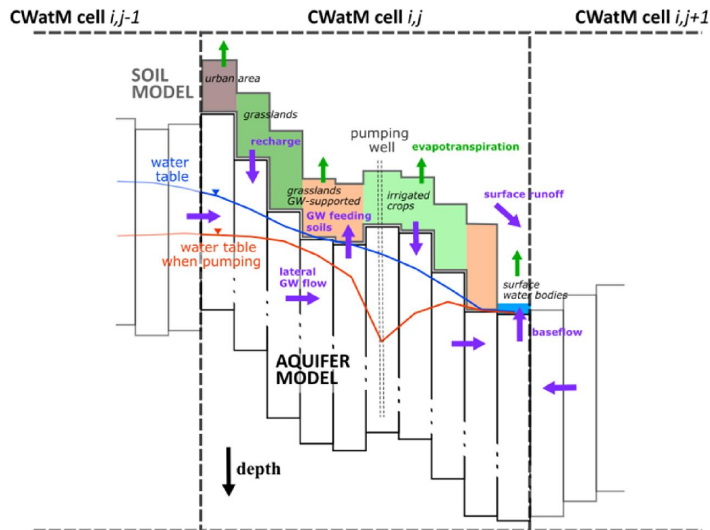



Figure 1: CWatM vertical and lateral sections focusing on coupling with MODFLOW. The CWatM cells are composed of five land cover types: grasslands, grasslands experiencing groundwater (GW) capillary rise, irrigated crops, urban areas, and surface water bodies. The figure illustrates the impact of aquifer pumping on the water table depth and consequently on baseflow and groundwater capillary rise toward soils.

2.2 Including groundwater within CWatM

2.2.1 Representation of soil and groundwater flows

CWatM-MODFLOW simulates subdaily hydrological processes that occur in soil and surface water bodies. Soils are represented in three layers. Unsaturated soil moisture redistribution and capillary rise effects are calculated using the van Genuchten–Mualem equations and soil hydraulic properties (Wösten and van Genuchten, 1988). At every daily timestep and within each CWatM cell, the upper soil layer receives infiltration from precipitation and irrigation. If the underlying simulated water table is below the bottom soil layer, water percolates from the deepest soil layer as a function of saturation. This water reaches the groundwater layer of the model as groundwater recharge. If the water table is above the lower bound of the soil, downward percolation does not occur, and groundwater feeds the soil (i.e., capillary flux). Finally, evapotranspiration depletes soil moisture, given the potential evapotranspiration demand and soil water availability.

The groundwater model is implemented using MODFLOW 6. Each day, the coupling is performed based on the following steps Figure 2: (1) CWatM initializes the timestep and begins processing the surface hydrological components, (2) CWatM simulates the groundwater recharge and extraction, which are converted in memory to MODFLOW inputs, (3) MODFLOW

inputs are passed to the MODFLOW model using the BMI, (4) MODFLOW runs the timestep using CWatM outputs, (5) MODFLOW outputs (baseflow and groundwater capillary rise) are read into CWatM using the BMI and are converted in memory into CWatM inputs, and (6) CWatM simulates other surface hydrological components. The MODFLOW model can be used at different space and time resolutions.

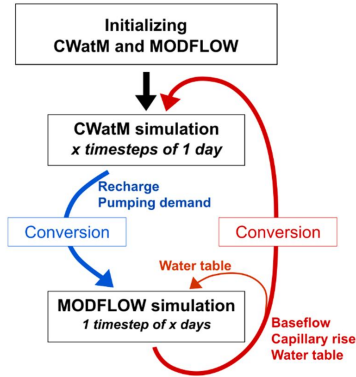



Figure 2: Scheme of the CWatM-MODFLOW coupling. *Conversion* refers to changing spatial resolution and units and passing to MODFLOW model using BMI (blue case) or reading MODFLOW output (red case). Note MODFLOW model can be used at different space and time resolutions

CWatM provides a recharge rate to each MODFLOW cell for each time step. However, when the previous simulated water table reaches the top of the aquifer (equal to the bottom of the soil layer), the aquifer cell is saturated. In this case, the recharge rate is set to zero. Recharge occurs once the water table drops below the top of the aquifer. Hydraulic heads, also referred to as a water table here, are computed in MODFLOW by numerically solving the groundwater flow equation (combining Darcy's law and the equation of continuity), considering no anisotropy and with one unconfined layer on the Dupuit–Forchheimer assumption (Eq.1):

$$\frac{\partial}{\partial x}(T\frac{\partial h}{\partial x}) + \frac{\partial}{\partial y}(T\frac{\partial h}{\partial y}) = \theta\frac{\partial h}{\partial t} - R(t) + Q(t)$$

where T is the transmissivity [L^2/T], equal to the hydraulic conductivity (K) times the saturated thickness of the aquifer (e_{sat}), θ is the porosity, R is the recharge rate [L/T], and Q is the pumping rate [L/T]. Hydraulic head lateral propagation is controlled by diffusivity $D = K.e_{sat} / \theta$. Upward flow is computed using the DRAIN MODFLOW package when the simulated water table reaches the top of the aquifer. This flow is equal to the difference between the hydraulic head in the cell and the top of the aquifer multiplied by a conductance parameter [L^2/T]. This upward flow is partitioned between the baseflow feeding the river's network directly and the groundwater capillary rise, feeding the soils. This ratio depends on the percentage of rivers attributed to each MODFLOW cell. Suppose a MODFLOW cell is highly saturated and no river is identified above this cell. In that case,

groundwater can indirectly feed the river network through capillary rise. In this area, the soil in CWatM becomes oversaturated,
reproducing an area of groundwater overflow.

Moreover, exchanges between groundwater and surface water bodies (lakes, rivers, and channels) are implemented within
CWatM-MODFLOW. Below the surface water bodies, if the simulated water table reaches the top of the aquifer (equal to the
bottom of the surface water body), an upward flow is sent from the aquifer toward the surface water body. On the contrary,
leakage toward the aquifer occurs if the simulated water table drops below the top of the aquifer. Lakes and reservoirs are
identified inside the CWatM grid using an identifier and the fraction of the 'water' land cover type (as illustrated by the 'surface
water bodies' fraction receiving baseflow in Figure 1).

2.2.2 Subgrid representation of hydrological processes

2.2.2 Land cover variability within the CWatM-MODFLOW grid cell is considered using a subgrid approach. Each CWatM
cell includes five land cover fractions: grasslands (including non-irrigated crops), irrigated crops, groundwater-supported
grasslands (Figure 1), urban areas, and water body areas. For each CWatM cell, the sum of each land cover fraction equals 1.
Here, groundwater-supported grasslands have the same properties as grasslands but correspond to land fractions where
groundwater often feeds soils. Setting up a groundwater-supported grasslands fraction allows the groundwater flow to be
concentrated toward the fraction of the soil that usually receives groundwater instead of distributing this flow homogeneously
over all CWatM grid cells (Figure 1). By this means, we can better take advantage of the finer resolution of the groundwater
layer (100 or 250 m). Consequently, soil moisture, surface runoff, and evapotranspiration are more pronounced in groundwater-
supported grasslands than in regular grasslands. Technically, the groundwater-supported grasslands fractions need to be
estimated after an initialization simulation to infer the groundwater-supported areas. Thus, CWatM-MODFLOW is run for a
first time during the simulation period to initialize the water table and to define the groundwater-supported fraction for each
CWatM grid cell. From this first simulation, MODFLOW cells are defined as groundwater-supported areas if groundwater
feeds the soil for more than 4 months out of 12. Then, groundwater-supported fractions are computed at CWatM resolution
and are used during simulations. Because this land cover fraction is constant over time, the groundwater-saturated land fraction
may exceed some grid cells' predefined fractions during extreme events. In this case, groundwater is also sent to other land
cover types (grasslands and irrigated areas). In other cases, when the simulated groundwater-saturated area is smaller than the
predefined groundwater-supported area, groundwater is distributed equally within the groundwater-supported fraction.

2.2.3 Groundwater pumping

CWatM-MODFLOW simulates water management and demand. Irrigation depends on the fraction of the irrigated crops land
cover type on each CWatM cell. Irrigation need is estimated as a function of potential crop evapotranspiration and available
soil moisture within the land cover type (Burek et al., 2020). Next, CWatM imposes pumping to the groundwater model for
each day. The water demand from all sectors (irrigation, livestock, industry, and households) can be attributed to surface water
bodies and to different groundwater model cells depending on where boreholes have been defined in the groundwater model

(Figure 1). Pumping in the aquifer is simulated with the WELL MODFLOW package. Groundwater abstraction can be limited if the water table in the pumping wells drops below a certain depth. This limit can correspond to physical or economic constraints and is adapted to the study areas. Finally, once the water table recovers above this limit, pumping can begin again. Note that here, we benefit from a new CWatM development, including surface water and agricultural management (discussed

220   further by Smilovic et al., 2022). To appreciate the impacts of surface water management on groundwater, more than 40 reservoirs simulated in the Bhima basin were outfitted with daily reservoir-specific operations and connected to specific spatial distribution areas (command areas) and canal networks. Reservoirs in the model distribute water based on daily command area demand and according to daily maximums with preference given to non-irrigation requests (domestic, industrial, and livestock). A fraction of this water leaks through the canal network. Further, rivers and lakes are set to satisfy some agricultural

225   demands. Water demand may be supplemented with groundwater when surface water volume does not coincide with the need. For irrigation demand, this study includes more than a dozen spatially distributed crops specific to the given region, each represented by four crop-specific growth stages affecting water use and planting and harvest dates. Regions within command areas have access to both surface and groundwater, while irrigated crops outside command areas only have access to groundwater. Irrigation is applied depending on crop water needs, soil moisture deficits, and surface and groundwater

230   availability. Several agricultural practices encourage groundwater recharge, such as non-precision irrigation, bare and freshly seeded fields at the beginning of the rainy season (low actual evapotranspiration), and leaky surface water distribution networks.

3 Two study regions with contrasted groundwater dynamics

The two study areas offer the possibility to refine CWatM-MODFLOW spatial resolution and water management, and to

235   examine the ability of the model to reproduce groundwater tables for decades, due to the existence of monitoring borehole networks

3.1 Seewinkel region, Austria

The Seewinkel region is a subpart of the Burgenland region in Austria (Karner et al., 2019; Magyar et al., 2014; Mitter and Schmid, 2021). The study area, of around 573 km^2, is limited in the west by Lake Neusiedl (Neusiedlersee in German) and in

240   the south and in the east by the border between Austria and Hungary. The Seewinkel region has been classified as semi-arid (Karner et al., 2019; Magyar et al., 2021), even if the aridity index, defined as the ratio between mean annual precipitation and potential evapotranspiration, reaches 1.0. Evapotranspiration shows a strong seasonal pattern, while precipitation is relatively homogeneous throughout the year (Table 1). The region is an important agricultural area, and around 16% of it is allocated to irrigated crops (maize, potatoes, vineyards, vegetables, and wheat). Only groundwater is used to satisfy irrigation demands.

245   The study area hosts several natural ponds of environmental interest. These ponds mainly depend on the water table in the

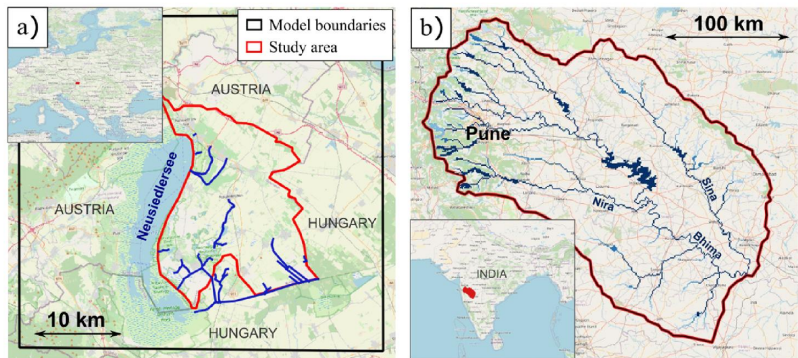
aquifer. To keep the ponds safe, pumping from irrigation wells is limited when the water table reaches a threshold depth (Magyar et al., 2021). The Seewinkel region and associated groundwater are modestly impacted by human water management.

© OpenStreetMap contributors 2022. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

Figure 3: Location of the two study regions. a) The Seewinkel region in Austria. Blue lines represent canals included in the model. CWatM and ModFlow models are set up respectively at 1 km and 100 m resolutions. b) The Upper Bhima basin in India. Blue lines represent the main rivers. CWatM and ModFlow models are set up at ~0.9 km and 250 m resolutions, respectively.

The CWatM-MODFLOW model (shown in black in Figure 3a) is applied at 1 km resolution, associated with a MODFLOW resolution of 100 m from a 10 m resolution DEM (Comprehensive digital field model (DGM) of the state of Burgenland. Produced from Airborne Laserscan data 2019.). To capture the influence of regional groundwater flow, the modeled area is larger than the study area (marked in red in Figure 3a). Therefore, small ponds are included in the study area, but Neusiedlersee is not. Outside the study area, the percentage of river within each MODFLOW cell is defined from a river network computed with the finer 10 m resolution DEM assuming that rivers are created when the drainage area exceeds 1 km^2. There is no river inside the study area, but there are several canals draining groundwater. Note that the MODFLOW DRAIN boundary condition on these canals is set at 1 m below the altitude inferred from the 100 m DEM. Groundwater leaving the aquifer below canals and lakes directly reaches the surface water network (see the baseflow in Figure 1). Using field data, the Seewinkel aquifer is modeled as one aquifer layer with a thickness of 20 m. Regional characteristics and model properties are synthesized in Table 1.

One pumping well is set up every 1 km^2 at the center of each CWatM cell. Irrigation demand and associated pumping rates depend on the deficit between potential evapotranspiration demand and current soil water contents within the CWatM cell (see Supplementary Material about the influence of irrigation efficiency and spatial density of pumping wells). Finally, the simulated CWatM-MODFLOW pumping rate at the Seewinkel scale for 2015 is 48% smaller than the abstraction limits imposed for the year 2015 in Seewinkel. Therefore, the model seems to appropriately represent groundwater use.

Table 1: Main characteristics and model properties of Seewinkel and Bhima. Annual averages are given for 1983–2016 and 2000–2009 periods for Seewinkel and Bhima, respectively. Aquifer permeability and porosity are obtained after calibration.

	Seewinkel region	Bhima basin
Precipitation [mm/yr]	582	805
Potential evapotranspiration [mm/yr]	578	1077
Fraction of irrigated area [%]	16	30
CWatM resolution [m]	1000	~900 (30'')
MODFLOW resolution [m]	100	250
Average soil thickness (3 layers) [m]	0.05, 0.25, 1.70	0.05, 0.26, 1.16
Aquifer permeability [m/s]	5.0×10^{-5}	1.2×10^{-5}
Aquifer porosity	0.07	0.018
Aquifer thickness [m]	20	50
Groundwater pumping simulated with CWatM-MODFLOW [mm/yr]	31	107 (+115 from surface water bodies)

270

3.2 Bhima basin, India

The Bhima basin is located in India and hosts 19 million people. The basin area is 46.000 km². The average aridity index is 0.75 (Table 1); however, precipitation shows a strong seasonality due to the monsoon. Both surface and groundwater satisfy domestic, industrial, and agricultural water demand. Irrigation is by far the most important groundwater use. Around 30% of

275 the basin area is allocated to irrigated crops. The river discharge recorded at the most downstream gauging stations (Takli) indicates a flow of 120 mm/year on average, but this flows for only 2 months on average.

The CWatM model is applied at 30 arcsec resolution (~900 m), associated with a MODFLOW resolution of 250 m from a 90 m resolution DEM (Yamazaki et al., 2019). Lakes and reservoirs are marked using HydroLAKES (Messager et al., 2016) and are refined and expanded upon with local data made available by the Pune Irrigation Circle. The percentage of river within

280 each MODFLOW cell is defined from a river network computed with a 90 m resolution DEM, assuming that rivers are created when the draining area exceeds 1 km². In Seewinkel, the aquifer upper limit under the rivers is set at 1 m below the altitude inferred from the 250 m DEM. This correction allows us to better take into account the river incision, which is strongly smoothed when resolution is upscaled to 100 or 250 m. When the water table reaches the top of the aquifer, groundwater is sent either to the soil or to the surface water network (see the groundwater feeding soils and baseflow indicated in Figure 1),

285 based on the river percentage map. In Bhima, canal networks are added to bring water from reservoirs toward cultivated areas

where water is required for irrigation (see section 2.2.3). Leakage is considered to occur below these canals through a water conveyance efficiency of 70%.

The Bhima basin aquifer was modeled with MODFLOW in Surinaidu et al. (2013) at 1 km resolution, assuming spatially uniform recharge and pumping, deriving each as a linear relationship with precipitation, incorporating different data sources.

290 The ranges for hydraulic conductivity and specific yield used for calibration were also used for this study, and we refer the reader to Surinaidu et al. (2013) for a more detailed description of the region. The model presented here expands on this by including surface water hydrology and management, spatially distributed groundwater recharge and outflow (pumping, capillary rise, and baseflow), spatially distributed non-irrigation demand, and by modelling crop-specific irrigation demand and increasing the spatial resolution. While the Seewinkel alluvial aquifer was identified, the Bhima basin is much larger and

295 is not composed of one homogeneous aquifer alone. The area hosts crystalline rocks, where weathering may potentially enhance permeability and porosity. We considered one homogenous aquifer layer with a thickness of 50 m (Surinaidu et al., 2013). We set up one pumping well in every MODFLOW cell (see Supplemental Material regarding the influence of irrigation efficiency and the spatial density of pumping wells). As for the Seewinkel model, pumping rates depend on the deficit between potential evapotranspiration demand and soil water content in each CWatM cell. Finally, the imposed pumping rate depends

300 on the fraction of irrigated areas above each pumping well. Pumping is prevented when the water table falls below a depth of 15 m (Surinaidu et al., 2013).

Table 1 provides mean simulated water withdrawals. Annual surface and groundwater withdrawals within the Bhima basin are estimated for different sectors for 2013 in the Upper Bhima Subbasin draft report. The model simulates total waterthese withdrawals closely (5% higher), with higher groundwater use and lower surface water use (Table 1)... Through discussions

305 with local water managers and engineers, including through stakeholder workshops (Karutz et al., 2022), it was found that groundwater use is generally underreported and underestimated for both irrigation and non-irrigation purposes. Further, canal leakage is not included in the report, producing an underestimation of groundwater availability. This model, therefore,In agreement with the uncertainty of the reported values and following experts opinion mentioned above, this model appropriately represents the region's annual surface and ground water use.

310 **4 Validation of the hydrological models**

4.1 Available observed data

Water table data from 81 and 373 monitoring boreholes were gathered for the Seewinkel and Bhima areas, respectively. These observed datasets come from the eHYD https://ehyd.gv.at/ database in Austria (BMLRT, 2020) and the Central Groundwater Board in India. They have been preprocessed to make them compatible with the simulated water table. First, boreholes close

315 to the model's boundary limits were removed, given that the simplified no-flow boundary condition in the groundwater model could lead to an unrealistic water table in these parts. Second, boreholes were removed where time data covered less than 50% of the simulation period. Third, we did not consider 5% of the remaining boreholes with the greatest discrepancy between

simulation and observation because they are impacted by specific local conditions that are not well reproduced by the model. This prevents them from influencing the calibration. Indeed, because model assumptions such as homogeneous aquifer or pumping wells location, it can be inferred that some boreholes could not be well represented by the model due to the vicinity of pumping wells, rivers or local heterogeneities. Finally, we kept 62 and 351 boreholes for the Seewinkel and Bhima areas (Figure 4a and Figure 5a), respectively. In the Bhima basin, we also used daily discharge data aggregated at a weekly scale at five gauging stations from the National Hydrology Project in India.

4.2 Comparison between observed and simulated water table

To fully benefit from the informative content of the observed water table, we separately evaluated time-averaged water tables (static part) and water table time fluctuations (transient part). This was done to better understand how each parameter is sensitive to both static and transient parts of observed data. As described above, the calibration relies on 62 and 351 boreholes for the Seewinkel and Bhima, respectively.

First, we assessed the model's ability to reproduce mean spatial variability of the water table driving the lateral groundwater flow. Observed and simulated water tables were averaged along the calibration period at each borehole. However, a comparison between observed and simulated water tables is not relevant and not sensitive to parameters, as the water table mimics the topography, which extends to several orders of magnitude, as noted by (Gleeson et al., 2021) and (Reinecke et al., 2020). Moreover, model topography can differ from actual topography. On the other hand, water table depths contain more discriminant information and are essential for the interaction between groundwater, soil, and surface. Therefore, we compared the water table depth (WTD) instead of the water tables. As a criterion, we used the "normalized mean water table depth difference" [%], defined by the following equation (Eq. 2):

$$C_{mean} = 100 \times \frac{1}{n} \times \sum_{1}^{n} |\frac{WTD_{obs,mean} - WTD_{sim,mean}}{WTD_{obs,mean}}|$$

First, we assessed the model's ability to reproduce mean spatial variability of the water table driving the lateral groundwater flow. Observed and simulated water tables were averaged along the calibration period at each borehole. However, a comparison between observed and simulated water tables is not relevant and not sensitive to parameters, as the water table mimics the surface elevation, which extends to several orders of magnitude, as noted by Gleeson et al. (2021) and Reinecke et al. (2020). Indeed, observed water table varies from 430 to 1000 m in Bhima while observed water table depth varies from 1 to 20 m. Moreover, model topography can differ from actual topography. On the other hand, water table depths contain more discriminant information and are essential for the interaction between groundwater, soil, and surface. Therefore, we compared the water table depth (WTD) instead of the water tables. As a criterion, we used the "normalized mean water table depth difference" [%], defined by the following equation (Eq. 2):

$$C_{mean} = 100 \times \frac{1}{n} \times \sum_{i=1}^{n} |\frac{\overline{WTD}_{obs,i} - \overline{WTD}_{sim,i}}{\overline{WTD}_{obs,i}}|$$

where n is the number of monitoring boreholes; *obs* and *sim* refer to observed and simulated depths, respectively; and $mean\overline{WTD}$ refers to the time ~~average. A C_{mean} value of 50% means~~-averaged WTD. Note that ~~the model reproduces~~ $mean\overline{WTD}_{obs,i}$ ranges from 0.25 to 12 m and from 1 to 20 m for Seewinkel and Bhima, respectively. Thus, shallow water table ~~depth with a discrepancy of 50%~~depths do not have too much weight on ~~average. A value below 50% would emphasize good model accuracy~~C_{mean}.

The second criterion focuses on time fluctuations in water tables observed in boreholes, whatever the mismatch value between observed and simulated time-averaged water table depth. Water table fluctuations correspond to the initial signal after the time average is removed. To simplify, we averaged observed and simulated water table fluctuations from all boreholes instead of comparing the fluctuations at each borehole. This approach smooths local behaviors and informs the water table fluctuations (*WTF*) at the basin scale at the first order. As a criterion, we used the normalized root mean square error [*%*] on average time fluctuations, defined by (Eq. 3):

$$nRMSE = \frac{100}{std(WTF_{obs})} \sqrt{\frac{1}{n_{obs}} \times \sum_{t=1}^{n_{obs}} (WTF_{obs,t} - WTF_{sim,t})^2}$$

where n_{obs} refers to the number of observations in time~~.~~ *(t)*. Thus, the *nRMSE* corresponds to the *RMSE* expressed as a percentage of the standard deviation (*std*) of the observed data.

~~The~~ CWatM-MODFLOW simulation ~~begins~~starts in 1981 and ends in 2016 for Seewinkel, ~~and~~while it begins in 1993 and ends in 2009 for Bhima. For both regions, ~~the~~ models are run ~~for the~~a first time during the ~~entire~~whole period to initialize the water table. ~~Groundwater parameters, such as~~For Bhima, hydraulic conductivity and ~~porosity, are calibrated manually by comparing simulated and observed water table recorded monthly~~specific yield variables were included first in ~~boreholes from 1983 to 2016 and twice a year from 1997 to 2009 for the Seewinkel and Bhima basins, respectively. Note that the initial CWatM-MODFLOW parameters had already been calibrated~~a calibration (Fortin et al., 2012) ~~at coarser resolution (MODFLOW at 500 m resolution) for Bhima using both daily river discharge (five gauges with 2000–2009) and water table. Standard CWatM parameters were used for Seewinkel. Therefore, only groundwater parameters are calibrated in this study.~~using five daily discharge stations (2000-2009), and ranges for calibration were derived from Surinaidu et al. (2013). The top 20% from the discharge calibration were further analyzed for water table fluctuation and depths. Groundwater parameters hydraulic conductivity and porosity are further calibrated by comparing simulated and observed water table recorded monthly in boreholes from 1983 to 2016 and twice a year from 1997 to 2009, respectively, for the Seewinkel and Bhima. Standard CWatM parameters were used in Seewinkel. Therefore, only groundwater parameters are calibrated in this study

5 Experiments to infer the impact of irrigation

The effects of irrigation are evaluated by comparing simulations with and without irrigation. To infer the effects of irrigation during the simulation period, we focus on the annual average of several variables (Figure 1): groundwater store, soil water content, evapotranspiration, groundwater recharge, and the fraction of humid areas (areas where groundwater often feeds soils). We focus on the impact of irrigation on actual evapotranspiration rate [L/m^2] and groundwater recharge in two land cover types, namely, groundwater-supported grasslands and irrigated areas. We expect that irrigation will increase evapotranspiration in irrigation areas and reduce evapotranspiration (and increase recharge) in non-irrigated areas. In Seewinkel, we also assess the impact of irrigation during the very dry summer of 2003, as we expect a more substantial effect of irrigation during that period. We also adjusted the irrigation efficiency parameter, which also impacts groundwater pumping, as well as the spatial density of pumping wells to test the sensitivity of the results to these settings (Supplement Material).

Land cover variability within the CWatM-MODFLOW grid cell is considered using a subgrid approach. Each CWatM cell includes five land cover fractions: grasslands (including non-irrigated crops), irrigated crops, groundwater-supported grasslands (Figure 1), urban areas, and water body areas. For each CWatM cell, the sum of each land cover fraction equals 1. Here, groundwater-supported grasslands have the same properties as grasslands but correspond to land fractions where groundwater often feeds soils. Setting up a groundwater-supported grasslands fraction allows the groundwater flow to be concentrated toward the fraction of the soil that usually receives groundwater instead of distributing this flow homogeneously over all CWatM grid cells (Figure 1). By this means, we can better take advantage of the finer resolution of the groundwater layer (100 or 250 m). Consequently, soil moisture, surface runoff, and evapotranspiration are more pronounced in groundwater-supported grasslands than in regular grasslands. Technically, the groundwater-supported grasslands fractions need to be estimated after an initialization simulation to infer the groundwater-supported areas. Thus, the calibrated CWatM-MODFLOW model is run for a first time during the simulation period to initialize the water table and to define the groundwater-supported fraction for each CWatM grid cell. From this first simulation, MODFLOW cells are defined as groundwater-supported areas if groundwater feeds the soil for more than 4 months out of 12. Then, groundwater-supported fractions are computed at CWatM resolution and are used during simulations. Because this land cover fraction is constant over time, the groundwater-saturated land fraction may exceed some grid cells' predefined fractions during extreme events. In this case, groundwater is also sent to other land cover types (grasslands and irrigated areas). In other cases, when the simulated groundwater-saturated area is smaller than the predefined groundwater-supported area, groundwater is distributed equally within the groundwater-supported fraction. Therefore, choosing a threshold of "4 months out of 12" is a compromise allowing to focus on areas significantly supported by groundwater. This illustrates the importance of simulating well water table depths.

6 Results

6.1 Groundwater parameters

Optimal aquifer permeability and porosity (Table 1) are compared to different datasets as described below. The permeability and porosity in Seewinkel are 5×10^{-5} m/s and 0.07, respectively. ~~Aquifer thickness was fixed at 20 m. The~~These values are respectively 3.10^{-4} m/s and 0.19 in the global GLHYMPS database (Gleeson et al., 2014; Huscroft et al., 2018) ~~overestimates these values to 3.10^{-4} m/s and 0.19, respectively.~~. The imposed aquifer thickness (20 m) agrees with the global depth-to-bedrock map (Shangguan et al., 2016), where the average value for the study area is around 22 m.

In Bhima, the permeability and porosity are 1.2×10^{-5} m/s and 0.018, respectively. ~~The aquifer thickness was fixed at 50 m. The GLHYMPS database underestimates the permeability (~~These values are respectively 3.16×10^{-6} m/s) and ~~overestimates the porosity (~~0.09~~)~~. in the global GLHYMPS database. The global depth-to-bedrock values consider thinner aquifers in this region (from zero to a few meters~~). For~~) compared to the imposed aquifer thickness (50 m). Note that for similar permeability (1×10^{-5} m/s) and thickness (50 m), Surinaidu et al. (2013) calibrated a porosity of 0.01–0.03 in a regional model. This suggests that permeability is underestimated and porosity overestimated in GLHYMPS over this basin.

Mis en forme : Exposant

6.2 Validation

The model reproduces the time-averaged water table depths recorded in boreholes in Seewinkel (Figure 4b), with a C_{mean} value of 38%. Therefore, we can infer that the hydraulic gradient and water table depth are reasonably well reproduced. The results for the Bhima basin show more contrast (Figure 5b), as the C_{mean} is 69% and simulated water table depth is too shallow on average. Indeed, the simulated mean water table averaged over all monitoring wells is 1.7 m higher than the observed mean in Bhima (Figure 6b), while the model shows a slight bias of 0.5 m in Seewinkel (Figure 6a). The reason for this bias is examined in the Discussion section.

In Seewinkel, the observed mean water table lies above a 1 m (2 m) depth for 17% (47%) of monitoring wells while this fraction is 29% (56%) in the model (Figure 4b). Thus, the model overestimates very shallow conditions (<1 m depth). The total land fraction where simulated mean water table is shallower than a 1 m (2 m) depth is 25% (46%). Thus, it can be inferred that a nonnegligible land fraction hosts shallow groundwater in Seewinkel.

In Bhima, the observed mean water table lies above a 1 m (2 m) depth for 0% (2%) of monitoring wells while this fraction is 17% (32%) in the model (Figure 5b). Thus, we conclude that the model overestimates shallow conditions in Bhima. Moreover, the total land fraction where simulated mean water table is shallower than a 1 m (2 m) depth is 6% (10%), indicating that monitoring wells are preferentially located in shallow areas.

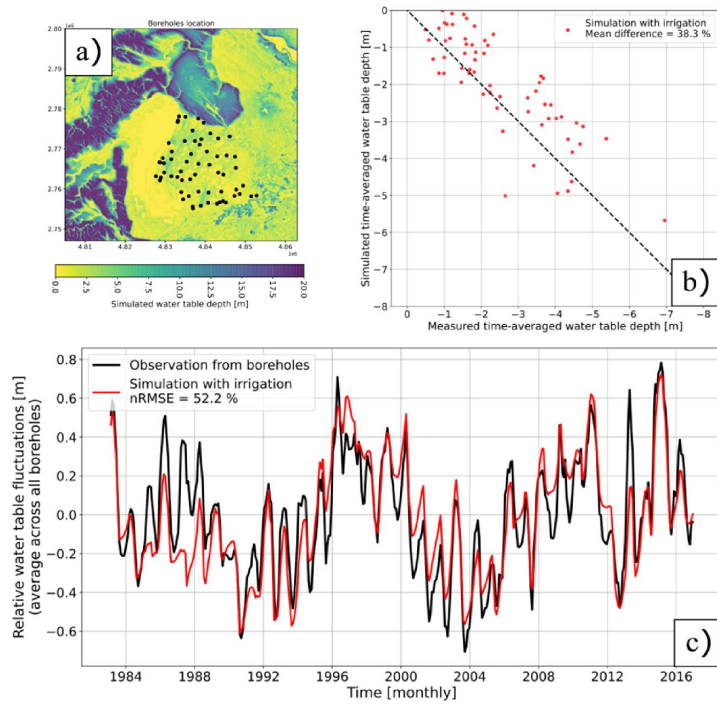
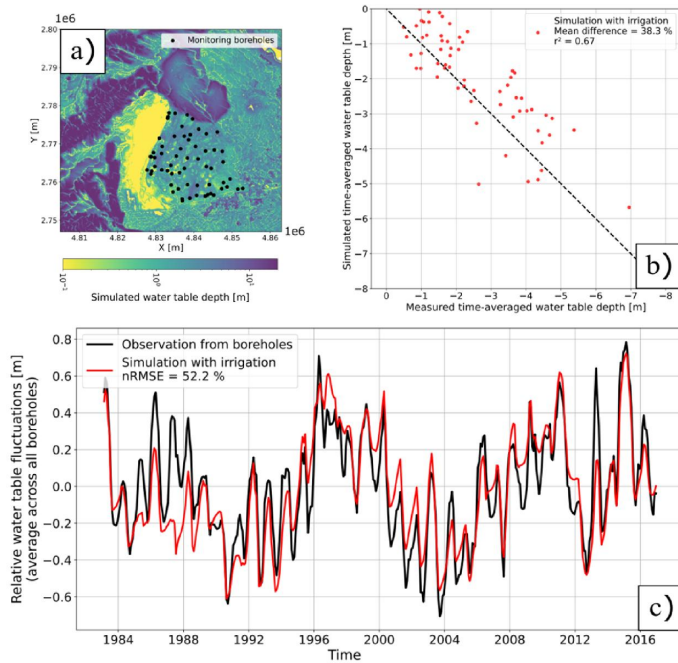



Figure 4: a) Map of the monitoring boreholes in the Seewinkel region. b) Comparison between observed and simulated time-averaged water table depth. c) Comparison between monthly observed and simulated water table time fluctuations averaged across all monitoring boreholes and expressed as anomalies (relative to the time average).

Both models successfully reproduce water table fluctuations averaged across all monitoring boreholes (Figure 4c) and Figure 5c). *nRMSE* reaches 52% and 41% for the Seewinkel and Bhima models, respectively. While the primary seasonal and inter-annual behavior is well reproduced in both areas, some events or periods are not well captured. For example, the Seewinkel model underestimates simulated fluctuations from 1986 to 1988. In addition, fluctuations are overestimated from 2000 to 2005, but the seasonal amplitude is slightly underestimated. In the Bhima basin, it also appears difficult to reproduce several inter-annual fluctuations and seasonal behavior at the same time. For example, changing the hydraulic parameters to increase seasonal amplitudes from 2003 to 2005 would decrease the base level even more during this period (Figure 5c). Likewise, increasing or decreasing pumping rates would improve the mean water table bias but would increase the nRMSE focused on water table fluctuations (Supplemental Material).

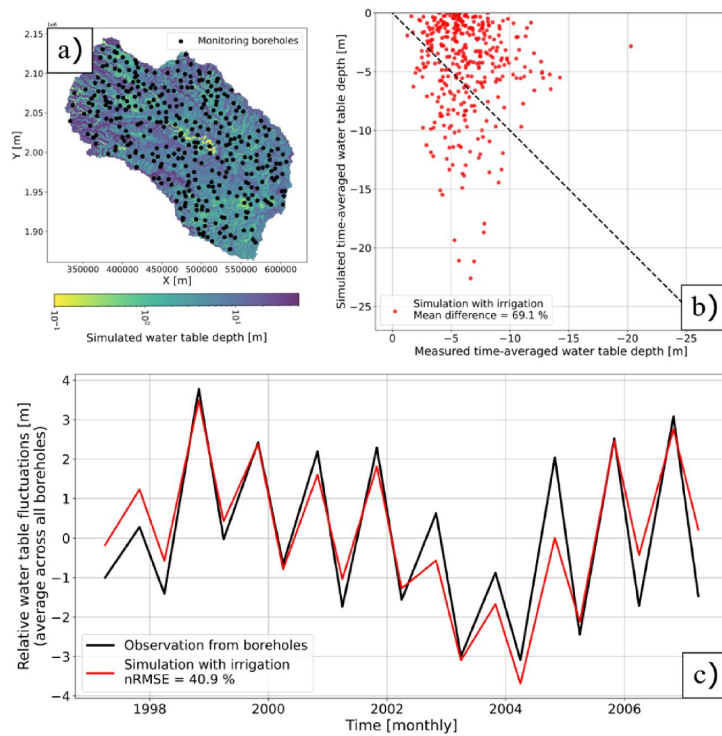
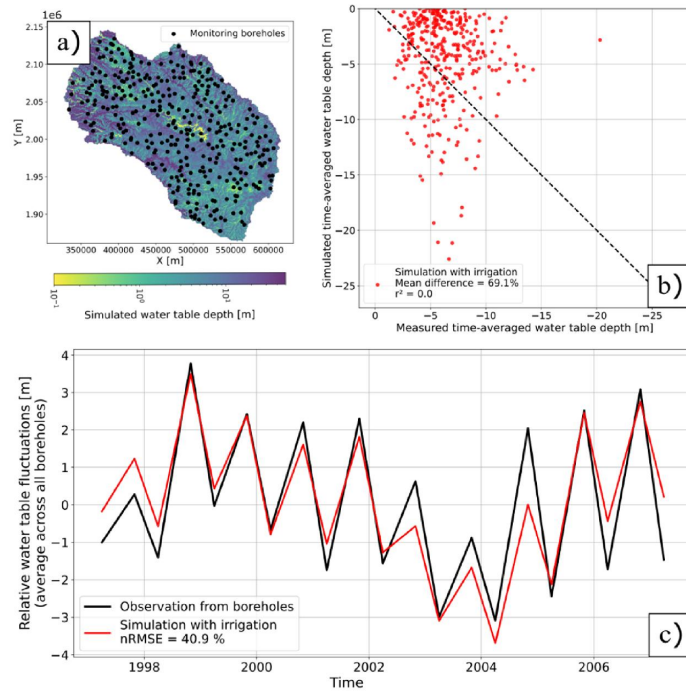



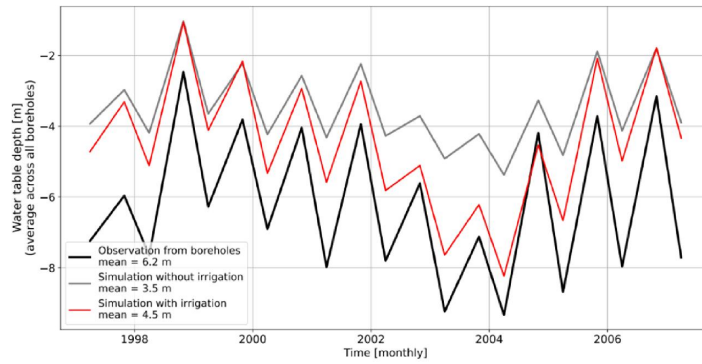
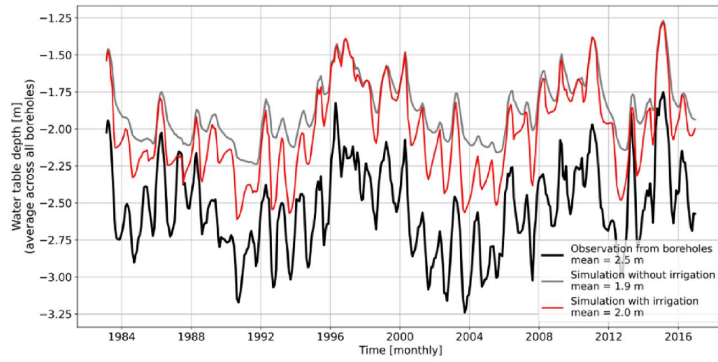
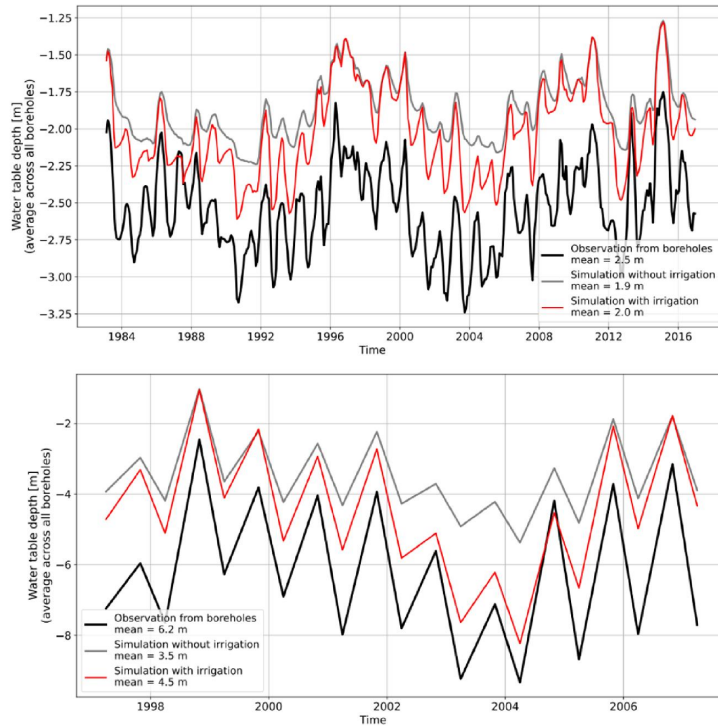
Figure 5: a) Map of the monitoring boreholes on the Bhima basin. **b)** Comparison between observed and simulated time-averaged water table depth. **c)** Comparison between observed and simulated water table time fluctuations averaged across all monitoring boreholes and expressed as anomalies (relative to the time average). **Based on the daily simulation, simulated water table is compared at the same days than observed water table, twice a year, before and after the monsoon.**

A comparison between observed and simulated weekly discharge in the Bhima basin is provided in the ~~Appendix A.1. The KGE efficiency values are 0.75, 0.68, 0.54, 0.76, and 0.66 from upstream to downstream stations.~~Supplementary Material. The Kling-Gupta Efficiency (*KGE*) (Gupta et al., 2009) values are 0.75, 0.68, 0.54, 0.76, and 0.66 from upstream to downstream stations. These results and the associated good criterion on water table fluctuations (*nRMSE*) make us confident that the model simulates rainfall partitioning between discharges and evapotranspiration wells.

6.3 Impact of pumping on the water table

The impact of pumping on the water table is a function of mean pumping rate and aquifer properties (permeability, saturated thickness, and porosity). Groundwater pumping theoretically decreases the water table within and around irrigated areas. The

results show that groundwater pumping enhances the seasonal water table fluctuations by intensifying water table recession during the dry season in both regions (Figure 6). It is also shown in Figure 6 that irrigation accentuates water table depletion during the driest years. Interestingly, during the recharge seasons, the water table within the borehole network is not strongly influenced by pumping in the two study regions, as the water tables with and without pumping are very close (Figure 6).

465   Because the water table is a proxy of groundwater storage, this result indicates that groundwater storage recovers rapidly and/or irrigation enhances net groundwater inputs. However, this is not the case for the driest periods (i.e., 2004–2005 in Seewinkel and 2003–2005 in Bhima).

470   **Figure 6: Comparison between absolute water table depth fluctuations obtained from CWatM with (red lines) and without (gray lines) irrigation in the Seewinkel (upper graph) and Bhima (lower graph) areas. Black lines represent observed data. Water table depth fluctuations are aggregated from 62 and 351 boreholes for Seewinkel and Bhima, respectively.**

Comparing models with and without irrigation allows us to infer the impact of pumping on the water table. However, groundwater recharge is impacted by irrigation, as we elaborate in the next section. Pumping reduces the average water table
475   at boreholes from 0 during the wet season to around 30 cm and 1–2 m during the growing season in Seewinkel and Bhima, respectively (Figure 6). The impact of pumping also varies spatially, depending on whether the monitoring borehole is located near the irrigation areas where pumping occurs. Drawdown due to pumping obtained on the time-averaged water table in monitoring boreholes varies between 0 and 50 cm and between zero and a few meters for Seewinkel and Bhima, respectively. The impact of pumping on the water table is more pronounced in Bhima than in Seewinkel (Figure 6). This can be explained
480   in two ways. First, aquifer permeability, thickness, and porosity are higher in Seewinkel (5×10^{-5} m/s, 20 m, and 0.07,

respectively) than in Bhima (1.2×10^{-5} m/s, 50 m, and 0.018). Therefore, water table drawdowns around pumping wells are less pronounced in Seewinkel. Second, the mean pumping rate is smaller in Seewinkel (31 mm/yr) than in Bhima (107 mm/yr).

6.4 Impact of irrigation

As expected, irrigation increases the total actual evapotranspiration in the two study regions. In Seewinkel, evapotranspiration in irrigated areas increases from 77 to 87 mm/yr. In Bhima, evapotranspiration in irrigated areas increases significantly, from 102 to 209 mm/yr. However, these changes in evapotranspiration are substantially smaller than the allocated irrigation withdrawals due to conveyance and application losses and the fact that some of the irrigated water reaching the soil percolates toward the groundwater without being used by crops. Indeed, groundwater pumping amounts to an annual average of 31 mm/yr in Seewinkel, while annual groundwater and surface water withdrawal amounts to 107 and 115 mm/yr in Bhima, respectively.

The increase in evapotranspiration is associated with a decrease in river discharge. In Seewinkel, irrigation reduces streamflow from 158 to 152 mm/yr. However, pumping increases the net lateral groundwater inflow to the study area by around 2 mm/yr. In Bhima, irrigation reduces streamflow from 257 to 167 mm/yr. In all, 3 and 14 mm/yr of the groundwater pumping reaches the river network due to irrigation losses in Seewinkel and Bhima, respectively).

Irrigation also significantly increases groundwater recharge in both areas. Recharge increases by 24% and 50% in Seewinkel and Bhima, respectively. This is counterbalanced by groundwater pumping such that irrigation causes a reduction of net groundwater input of 12 and 5 mm/yr in Seewinkel and Bhima, respectively. However, the net recharge reduction is more pronounced without irrigation from the surface water bodies in Bhima. Indeed, around 40 mm/yr of the recharge rise is attributed to irrigation from surface water bodies in Bhima. Note that in both models, irrigation increases recharge in irrigated lands, not only during irrigation season but also during the wet season (Appendix A.21).

Recharge increases not only in irrigated areas (92% and 94% for Seewinkel and Bhima, respectively) but also in grasslands (6% and 5%) and in groundwater-supported grasslands (2% and 1%). These results can be explained by the influence of groundwater pumping, which reduces the water table outside the irrigated areas. When irrigation is not applied, recharge is limited in some periods, due to the high water table in groundwater-supported grasslands and episodically in other land covers. The water table is deeper when irrigation pumping is used, thus enhancing water percolation from the soils to the aquifer. Consequently, soil humidity and evapotranspiration rates decrease in groundwater-supported areas due to irrigation, while recharge increases. This mechanism is illustrated by Figure 1, which shows a 1D section of a water table with and without irrigation.

In Seewinkel, the mean soil water content in groundwater-supported areas decreases from 85 to 79 cm due to pumping (Figure 7a). The associated evapotranspiration rate diminishes from 595 to 573 mm/yr/m² per unit area of groundwater-supported areas. In Bhima, the mean soil water content in groundwater-supported areas decreases from 48 to 46 cm when pumping is applied, leading to a drop in the associated evapotranspiration rate from 920 to 752 mm/yr/m² per unit area of groundwater-supported areas (Figure 7b).

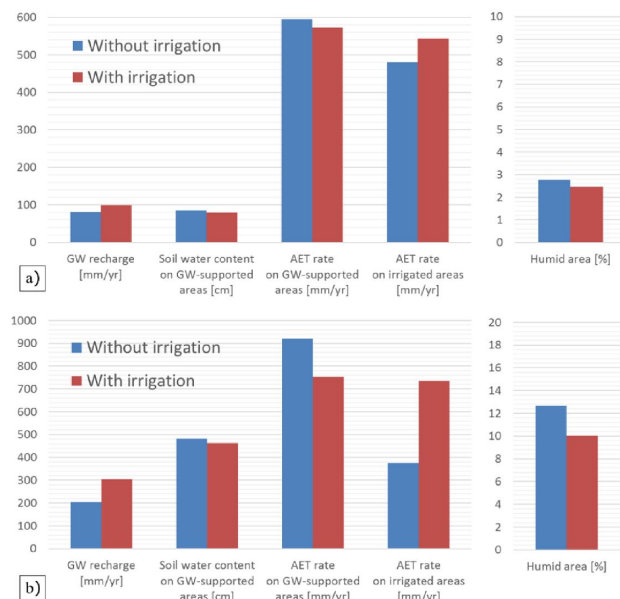



Figure 7: Comparison of some indicators obtained from CWatM-MODFLOW simulation with and without irrigation in annual average values in 1983–2016 for Seewinkel (a) and in 1997-2009 for Bhima (b). Note that the units are different for each variable. Areas were considered humid when the groundwater supported soils for at least 4 months of 12.

The water cycle is slightly impacted by irrigation at an annual scale in Seewinkel, and we can expect a more significant impact through the dry season when irrigation withdrawals occur. Focusing on the dry summer of 2003, the soil water content in groundwater-supported areas decreases from 70 to 65 cm due to pumping (Appendix A.3).Supplementary Material). Consequently, the evapotranspiration rate in groundwater-supported areas decreases from 292 to 274 mm/m² and increases by 40% in irrigated areas.

7 Discussion

7.1 Impact of irrigation on the hydrological cycle

Here, we tested the implication of irrigation using the coupled CWatM-MODFLOW model at very high resolutions and including a fine representation of water management practices. In Seewinkel, groundwater is the only source of irrigation, while surface water in Bhima satisfies half of the irrigation demand. Our experiment to infer the impact of the irrigation

highlights that irrigation does not significantly impact the average annual water balance in Seewinkel. However, the impact of irrigation is more pronounced during dry periods (e.g., the dry summer of 2003). The low average pumping rate can explain the small effect of pumping in Seewinkel (31 mm/yr) compared to the recharge rate (80 mm/yr). Moreover, the net water balance of the aquifer only drops by 12 mm/yr because recharge increases by 19 mm/yr due to irrigation. In addition, permeability and porosity in Seewinkel are high enough to limit the water table drawdowns induced by pumping. Finally, the soil storage is larger in Seewinkel than in Bhima (average soil water content in groundwater-supported areas is 85 and 48 cm, respectively), involving less dependence on groundwater (less irrigation and less groundwater recharge).

In Bhima, the experiment reveals that irrigation strongly increases evapotranspiration from irrigated areas and slightly reduces available water in groundwater-supported areas. Consequently, river discharge is strongly diminished. Like Seewinkel, groundwater recharge increases both in irrigated and non-irrigated lands. The inclusion of water transfer within the unsaturated zone, thickened due to pumping, could reduce effective recharge, as pointed out by Cao et al. (2016). However, due to the pumping rate and the thickness of the aquifer, the water table drawdown was much more significant in their study.

Assuming negligible groundwater interaction within irrigated lands, 70% and 47% of applied irrigation (without accounting for evaporation and runoff losses) percolates below soils and increases groundwater recharge in Seewinkel and Bhima, respectively. Similar values are identified in several other studies. Estimates of recharge from excess irrigation are variable and range from 0% to 80% (Groundwater Resource Committe, 2009; Le Coz et al., 2013; Meixner et al., 2016; Ochoa et al., 2013; Roark and Healy, 1998; Scanlon et al., 2005; Willis and Black, 1996), and depend on the efficiency of irrigation techniques (Grafton et al., 2018; Meredith and Blais, 2019).

In Seewinkel and Bhima, we find that groundwater storage is replenished each year during the wet season in spite of pumping during dry seasons because water tables with and without pumping are very similar during wet seasons (Figure 6). We infer that this result depends on the studied regions, as explained below. Indeed, the Seewinkel and Bhima aquifers are shallow, and pumping is limited by water table depth thresholds. Consequently, groundwater resources can be replenished during wet seasons. In Seewinkel, pumping is limited to protect the natural ponds sustained by groundwater. In Bhima, pumping is limited due to economic constraints and decreases in aquifer permeability with depth.

Recharge is enhanced where the water table is deep, while soil water content and evapotranspiration increase where it is shallow (Figure 1) (Martínez-de la Torre and Miguez-Macho, 2019; Swenson et al., 2019). Therefore, the spatial distribution of the water table depth is critical when focusing on the interaction between groundwater and soils. Our results suggest that simulated water tables are overestimated. This point is discussed in the next section. In addition, the model could be improved by allowing crops to draw water directly from groundwater systems where their roots exceed the soil depth and where the water table is shallow. Finally, water table depth representation is also relevant from an irrigation perspective. First, pumping redistributes groundwater, potentially impacting baseflow or neighboring areas. Second, water table depth is a proxy for available groundwater resources. Finally, it affects pumping costs (Turner et al., 2019).

7.2 Difficulty of reproducing water table variability

560 The primary result of this study is that the coupled CWatM-MODFLOW model managed to reproduce the average water table time fluctuations (Figure 4c and Figure 5c) very well. This result is in line with other large-scale models (de Graaf et al., 2017; Martínez-de la Torre and Miguez-Macho, 2019; Sutanudjaja et al., 2014, 2011; Vergnes et al., 2020). Groundwater pumping appears to be necessary to reproduce time-fluctuation anomalies in the water table (fluctuations are amplified by pumping in Figure 6). We infer that, by calibrating water table fluctuations separately, we simulated a good combination between

565 groundwater recharge, pumping, and aquifer lateral flow (driven by hydraulic diffusivity). Therefore, water table time fluctuations contain important information (Houben et al., 2022), even when reproducing the absolute water table depth remains challenging, particularly in Bhima, as also noted by Vergnes et al. (2020). Note that here, we initially averaged water table time-fluctuations from all boreholes located within the basins, similarly to Long et al. (2020).

By contrast, the spatial distribution of water table depth can be challenging to reproduce and may be less informative because

570 it is subject to spatial resolution and geological heterogeneity. The comparison between observed and simulated water table depth is satisfactory in Seewinkel (Figure 4b) but only provides first-order satisfaction in Bhima (Figure 5b). A first explanation for this may be that the Seewinkel aquifer is more homogenous and well-known. Consequently, the model can better describe its properties, such as thickness, permeability, and porosity. In contrast, the Bhima basin is much larger and hosts mainly crystalline rocks with potential variable weathering and fracture density, making its aquifer properties heterogeneous. Beyond

575 the aquifer parameters, the ability to reproduce time-averaged water table depth is constrained by two critical factors in both areas: the pumping rates and the spatial resolution of the groundwater model. These two factors are more acute in Bhima, where pumping is more pronounced and the resolution is coarser (250 m). In both study areas, simulated mean annual pumping rates at the basin scale were validated. However, we acknowledge that pumping estimates could remain uncertain regarding annual averages, seasonal behavior, and pumping well locations.

580 Groundwater model resolution is critical because degrading resolution smoothens topographical variability. The altitude of local low points in valleys due to river incision, often corresponding to aquifer resurgences, is increased more than it is in neighboring cells when degrading resolution. Consequently, the water table is shallower than expected because the low points, locally controlling the water table, are higher than expected (see topography and water table comparisons at 100 and 250 m resolution in the Appendix A.42). This explains why very shallow conditions (<1 m depth) close to rivers are overestimated

585 in Seewinkel despite the 100 m resolution employed. A similar bias was obtained by Maxwell et al. (2015), who used a spatial resolution of 1 km over the continental US.

The weakness of large-scale models for reproducing the spatial distribution of water table depths has been mentioned but not adequately explained in several previous studies (Fan et al., 2013; Maxwell et al., 2015; Reinecke et al., 2020). These studies indicate that mismatch between models and observed data could be due to: (i) well representativity within a too coarse mesh

590 (see also Martínez-de la Torre and Miguez-Macho, 2019), (ii) well sampling bias, or (iii) groundwater pumping. However, the ability of large-scale models to reproduce water tables remains primordial, and the means of evaluating them should be

improved (Gleeson et al., 2021). Here, for the first time, we used a large-scale hydrological model, including groundwater flow and pumping, at a very fine resolution. We note that, as expected, regional models show a better aptitude to reproduce water table depth. For instance, Vergnes et al. (2020) reproduced a mean water table depth on 639 piezometers across several French regions with a median bias of around 3 m. Similarly, Soltani et al. (2021) simulated water tables with an *RMSE* of 3.8 m based on around 28,000 groundwater-head stations in Denmark. Jing et al. (2018) reproduced water table depths with an *RMSE* of 6.3 m. In this study, we simulated mean water table depths with *RMSEs* of 1.0 and 5.0 m in Seewinkel and Bhima, respectively. Surinaidu et al. (2013) obtained a similar *RMSE* in Bhima with a simpler model at a coarser resolution and a uniform recharge/discharge. However, a comparison with water table fluctuations, depth bias, and surface water flows, including discharge and reservoir levels, was not performed.

We argue that simulated water tables tend to be shallower than the observed water table due to the coarse spatial resolution, as explained above. However, increasing the groundwater model resolution would significantly increase computing time, which makes it challenging to perform sensitivity analysis. Indeed, the same simulation performed in the Bhima basin at a 100 m resolution lasts around 2 days, versus 9 h at a 250 m resolution. Figure 8 illustrates how increasing the resolution from 250 to 100 m slightly reduces water table depth bias. While the *nRMSE* criterion is not significantly better, highlighting the role of heterogeneity rather than spatial resolution, the mean water table depth is now overestimated by 0.5 m instead of 1.7 m (Figure 8). Reinecke et al. (2019b) focused on the reason for water table depth mismatch and its sensitivity. They highlighted the importance of hydraulic conductivity, recharge, and surface water body elevation. The surface water body elevation can be assimilated to the top of the aquifer in the two CWatM-MODFLOW models described here. Therefore, this critical factor is linked to spatial resolution. By comparison, Reinecke et al. (2019a) improved their RMSE from 54 to 27 m using a resolution of ~9 km and 900 m, respectively. Note that their experiment was located in a steep domain (New Zealand).

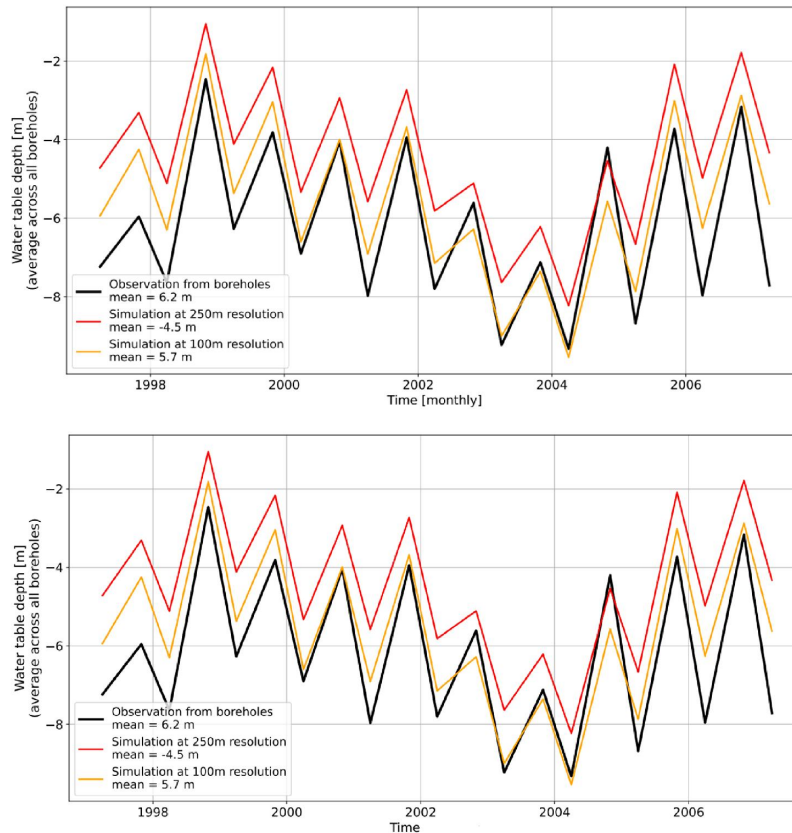



Figure 8: Comparison between simulated water table depth using a groundwater model resolution of 250 and 100 m in the Bhima basin. Black lines represent observed data. Water table depth fluctuations are aggregated from 351 boreholes. Note how improving the resolution reduces the water table.

7.3 Importance of implementing a subgrid approach in large-scale models

CWatM uses a mosaic approach to represent the subgrid variability of the land cover type. In addition, we developed a subgrid approach to improve model efficiency, specifically for groundwater exchanges, to benefit from the finer MODFLOW resolution. Indeed, each CWatM grid cell (1 × 1 km) contains a land cover fraction where soil is supported by groundwater. Therefore, we found that the impact of pumping on soil-groundwater exchanges was more significant when groundwater

capillary rise focused on a fraction of each grid cell instead of distributing the flow homogenously over the whole cell. This approach enables better simulation of the link between the water table, evapotranspiration, and recharge without refining the CWatM grid.

625 As pointed out previously, surface water body elevation, linked to the spatial resolution here, is important for simulating absolute water table depth (Reinecke et al., 2019b). This is important, as the river network elevation strongly controls water table depth. Moreover, rivers represent groundwater outlets, as CWatM-MODFLOW generates baseflow based on the simulated water table below rivers. However, a whole cell cannot be entirely assimilated to a river due to the coarse resolution of large-scale models; thus, two critical issues can be noted. First, large-scale hydrological modelers never have actually been

630 observed in river networks, including the finest low-order and intermittent rivers. They usually generate a river network based on a digital elevation model (Yamazaki et al., 2019), so that partitioning between baseflow and groundwater capillary rise is predefined before the simulation. Second, groundwater reaching the top of the aquifer can feed either rivers (as baseflow) or soils (as capillary rise). Our modeling approach relies on a finer digital elevation model used to compute a river network at a finer resolution than the groundwater model (sections 2.2.1 and 3). However, we acknowledge that small natural or artificial

635 drains are not included in the model. Then, partitioning between baseflow and capillary rise is defined at the resolution of the groundwater model. This partitioning impacts the depth of the water table. Indeed, increasing the baseflow fraction decreases the water table depth and reduces the soil water availability. Our approach partly deals with this issue as the resolution is very fine (100 and 250 m) and because cells where groundwater oversaturates soils generate overland runoff, feeding rivers.
Other subgrid approaches exist. The TOPMODEL approach (Beven and Kirkby, 1979), for instance, is used in many regional-

640 scale models. It infers subsurface flow convergence based on a finer digital elevation model. Vergnes et al. (2014) also proposed an alternative approach, in which a groundwater flow equation simulates the water table dynamic but capillary rise only occurs in a fraction of the cells based on the distribution of the topography at the subgrid level. Similar approaches are used within CWatM and PCR-GLOBWB when no lateral groundwater flow is considered. Another option is to represent coarse cells or basins with equivalent (representative) hillslopes (Fan et al., 2019; Loritz et al., 2017; Troch et al., 2003).

645 Recently, new hillslope models have been developed for use at a large scale (Hazenberg et al., 2015; Swenson et al., 2019). However, this conceptual modeling approach cannot include regional lateral flow convergences or groundwater pumping. Finally, another approach would be to aim for hyper-resolution in each compartment (surface, soil, and groundwater). However, this would require huge computing capacity to be achievable at a continental scale.

8 Conclusion

650 Parametrization, spatial resolution, and human water management are significant challenges for implementing physically based large-scale hydrological models. As in Hanasaki et al. (2022), where resolution was ~2 km, we applied a large-scale model to two regions, including water management practices with a hyper-resolution of 1 km for CWatM and of 100–250 m for the groundwater model. Benedict et al. (2019) also recently addressed the resolution issue in large-scale models. They found no

significant improvement in discharge predictions even after improving the spatial resolution. However, their hydrological
model did not include groundwater lateral flow, and resolution was still coarse (~5 km). Reinecke et al. (2020) found that
improving resolution alone is not enough to reproduce water table depth distribution. Their finest resolution was ~900 m. After
all, we found that improving resolution and model parameters and including pumping data are necessary to improve large-
scale models. At the same time, calibrating the model at a regional scale with a finer resolution than 100 m appears challenging.
Supporting the results of several previous studies, we found that relative water table fluctuations can be sufficiently well
simulated to infer groundwater model parameters at the aquifer scale (Houben et al., 2022) and to simulate discharge well
(Soltani et al., 2021; Sutanudjaja et al., 2014). Calibrating large-scale hydrological models on discharge data and the water
table would ensure that the main water cycle is reproduced (evapotranspiration and discharge), as well as inner processes such
as groundwater recharge, baseflow, and groundwater support to vegetation. The evaluation of large-scale hydrological models
remains a challenge (Gleeson et al., 2021), as water table depth patterns are difficult to reproduce at large scales. This pattern
is only acceptable at the first-order in the Bhima basin, even at 100–250 m resolution. This result is explained by the difficulty
of including the heterogeneity of groundwater bodies and pumping within the model. After all, improving resolution to 100 m
reduced the bias of the depth distribution of the water table. We argue that reproducing mean water table fluctuations and depth
distributions without bias constitutes a significant improvement in large-scale hydrological models. Finally, validated models
can be used to predict responses to climate and human interventions such as land cover changes and irrigation management.

Appendix

A.1 Model validation on discharge stations in Bhima

In Bhima, CWatM-MODFLOW successfully reproduces weekly streamflow in five discharge stations. The KGE efficiency
values are 0.75, 0.68, 0.54, 0.76, and 0.66 from upstream to downstream. Simulated and observed streamflow data are
presented on Figure A.1.

Code de champ modifié

675

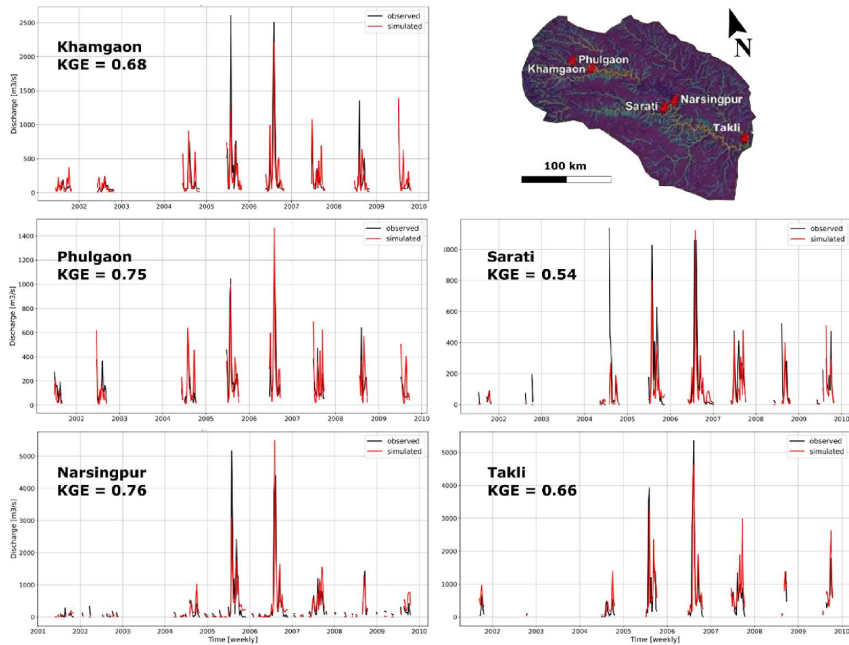



Figure A1: Comparison between observed and simulated weekly discharge at five discharge stations located in the Bhima basin. The figure in the upper left corner shows the stations locations.

A.2A.1 Impact of irrigation on groundwater recharge

Irrigation increases groundwater recharge in Burgenland and Bhima. Figure A.2 shows that irrigation increases recharge during
680 irrigation periods but also during humid seasons when there is no irrigation. During dry seasons, note that recharge is equal or close to 0 mm/yr when no irrigation is applied (red lines). During wet seasons, recharge peaks due to rainfall events are amplified because of water applied during the dry season (orange line).

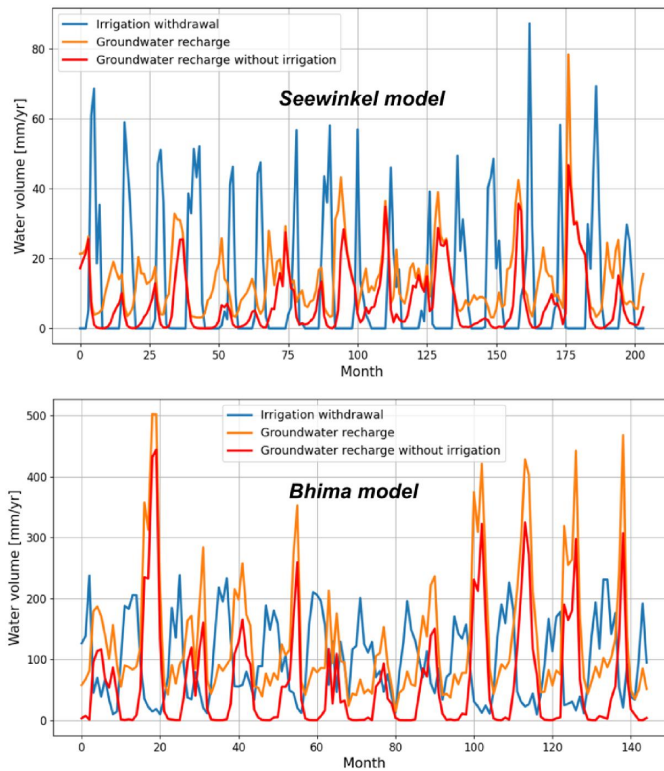



Figure A2: Comparison between average monthly recharge with and without irrigation in irrigated areas in Seewinkel (upper graph) and Bhima (lower graph). Irrigation withdrawal is also represented in blue.

685

~~A.3 Impact of irrigation in summer 2003 in Seewinkel~~

~~The impact of irrigation during summer 2003 is illustrated in Figure A.3. As expected, the impact of irrigation is more pronounced during summer 2003 than the annual average because pumping and irrigation occur during this season. Moreover, 2003 is very dry, so pumping demand is higher than usual.~~

690

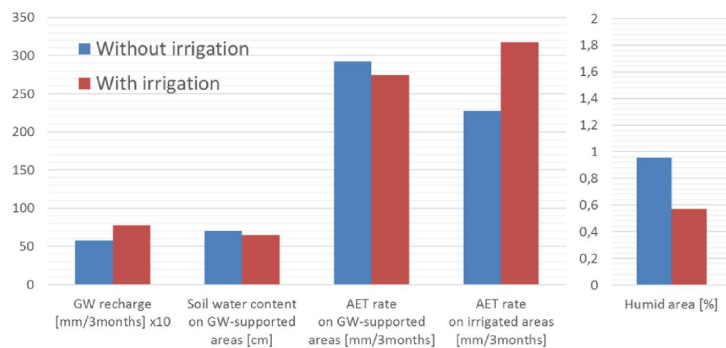



Figure A3: Comparison of some indicators obtained from CWatM-MODFLOW simulation in Seewinkel with and without irrigation during the dry summer 2003 (June, July, and August). Note that units are different for each variable. Areas were considered humid when groundwater supported soils for at least 1 month.

A.4

695 **A.2 Impact of the spatial resolution on topography and simulated water table**

The spatial resolution of the groundwater flow model MODFLOW implemented in CWatM-MODFLOW affects absolute water table depths. This is due to the fact that degrading the spatial resolution smooths topography, as highlighted by a 1D section of CWatM-MODFLOW in Bhima (Figure A.4). The resulting simulated water table is affected by the smoothing of the topography, as low points, representing groundwater resurgences, are higher at 250 than at 100 m resolution.

700

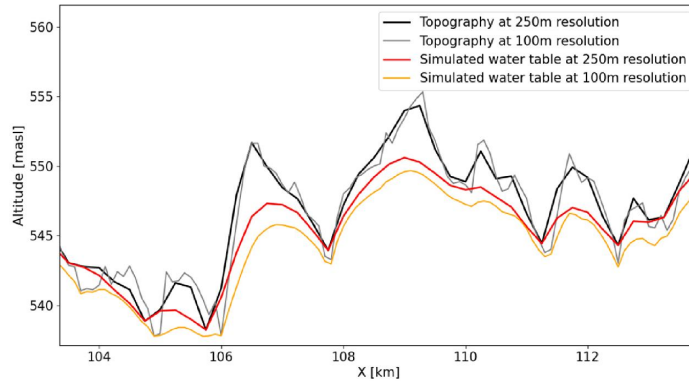



Figure A4: 1D section of topography and the simulated water table in Bhima. Comparison between a groundwater model resolution of 100 and 250 m. Note how topography at 100 m resolution is below topography at 250 m resolution in valleys. The resulting water table appears lower at a 100 than at 250 m resolution over the whole domain.

Code and data availability

705   CWatM-MODFLOW model and data used in this study are available on Zenodo (10.5281/zenodo.6609071). CWatM codes, tutorials and documentation are available on https://cwatm.iiasa.ac.at/ and on https://github.com/iiasa/CWatM.

Acknowledgments

This study was partially supported by the WaterstressAT project (KR19AC0K17504), funded by the Austrian Climate and Energy Fund as part of the Austrian Climate Research Programme 12th call (2019). This study and the model development

710   were also conducted as part of the Belmont Forum Sustainable Urbanisation Global Initiative (SUGI)/Food–Water–Energy Nexus theme for which coordination was supported by the US National Science Foundation under grant ICER/EAR-1829999 to Stanford University. We thank Christian Wawra for his feedback concerning the groundwater model in Seewinkel.

Competing Interests

The authors declare that they have no conflict of interest.

715 **Authors contribution**

LG, PB, MS and JdB developed the model code (CWatM-MODFLOW coupling). MS, PB, LG prepared model input and observation datasets. LG performed the simulations. LG prepared the manuscript with contributions from all co-authors.

References

Comprehensive digital field model (DGM) of the state of Burgenland. Produced from Airborne Laserscan data 2019.:
720 https://www.burgenland.at/service/open-government-data/bestandsliste/detail/service/ogddigitales-gelaendemodell-10-meter-burgenland/.

Bakker, M., Post, V., Langevin, C. D., Hughes, J. D., White, J. T., Starn, J. J., and Fienen, M. N.: Scripting MODFLOW Model Development Using Python and FloPy, 54, 733–739, https://doi.org/10.1111/gwat.12413, 2016.

Benedict, I., Van Heerwaarden, C. C., Weerts, A. H., and Hazeleger, W.: The benefits of spatial resolution increase in global
725 simulations of the hydrological cycle evaluated for the Rhine and Mississippi basins, Hydrol. Earth Syst. Sci., 23, 1779–1800, https://doi.org/10.5194/hess-23-1779-2019, 2019.

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrol. Sci. Bull., 24, 43–69, https://doi.org/10.1080/02626667909491834, 1979.

Bierkens, M. F. P.: Global hydrology 2015: State, trends, and directions, Water Resour. Res., 51, 4923–4947,
730 https://doi.org/10.1002/2015WR017173.Received, 2015.

BMLRT: Wassergüte in Österreich, Jahresbericht 2016-2018, 2020.

Burek, P., Satoh, Y., Kahil, T., Tang, T., Greve, P., Smilovic, M., Guillaumot, L., Zhao, F., and Wada, Y.: Development of the Community Water Model (CWatM v1.04) - A high-resolution hydrological model for global and regional assessment of integrated water resources management, Geosci. Model Dev., 13, 3267–3298, https://doi.org/10.5194/gmd-13-3267-2020,
735 2020.

Cao, G., Scanlon, B. R., Han, D., and Zheng, C.: Impacts of thickening unsaturated zone on groundwater recharge in the North China Plain, J. Hydrol., 537, 260–270, https://doi.org/10.1016/j.jhydrol.2016.03.049, 2016.

Condon, L. E. and Maxwell, R. M.: Simulating the sensitivity of evapotranspiration and streamflow to large-scale groundwater depletion, Sci. Adv., 5, https://doi.org/10.1126/sciadv.aav4574, 2019.

740 Condon, L. E., Atchley, A. L., and Maxwell, R. M.: Evapotranspiration depletes groundwater under warming over the contiguous United States, Nat. Commun., 11, https://doi.org/10.1038/s41467-020-14688-0, 2020.

Le Coz, M., Favreau, G., and Ousmane, S. D.: Modeling Increased Groundwater Recharge due to Change from Rainfed to Irrigated Cropping in a Semiarid Region, Vadose Zo. J., 12, vzj2012.0148, https://doi.org/10.2136/vzj2012.0148, 2013.

Dalin, C., Wada, Y., Kastner, T., and Puma, M. J.: Groundwater depletion embedded in international food trade, Nature, 543,
745 700–704, https://doi.org/10.1038/nature21403, 2017.

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J. P., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., and

Voldoire, A.: Recent Changes in the ISBA-CTRIP Land Surface System for Use in the CNRM-CM6 Climate Model and in Global Off-Line Hydrological Applications, J. Adv. Model. Earth Syst., 11, 1207–1252, https://doi.org/10.1029/2018MS001545, 2019.

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

Famiglietti, J. S. and Wood, E. F.: Multiscale modeling of spatially variable water and energy balance processes, Water Resour. Res., 30, 3061–3078, https://doi.org/10.1029/94WR01498, 1994.

755 Fan, Y. and Miguez-Macho, G.: Potential groundwater contribution to Amazon evapotranspiration, Hydrol. Earth Syst. Sci., 14, 2039–2056, https://doi.org/10.5194/hess-14-2039-2010, 2010.

Fan, Y., Li, H., and Miguez-Macho, G.: Global patterns of groundwater table depth, Science (80-.)., 339, 940–943, https://doi.org/10.1126/science.1229881, 2013.

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A.,
760 Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B., Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., van Verseveld, W., Volk, J., and Yamazaki, D.: Hillslope Hydrology in Global Change Research and Earth System Modeling, Water Resour. Res., 1737–1772, https://doi.org/10.1029/2018WR023903, 2019.

765 Fortin, F. A., De Rainville, F. M., Gardner, M. A., Parizeau, M., and Gagńe, C.: DEAP: Evolutionary algorithms made easy, J. Mach. Learn. Res., 13, 2171–2175, 2012.

Gleeson, T. and Richter, B.: How much groundwater can we pump and protect environmental flows through time? Presumptive standards for conjunctive management of aquifers and rivers, River Res. Appl., 34, 83–92, https://doi.org/10.1002/rra.3185, 2018.

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

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

775 Gleeson, T., Wagener, T., Döll, P., Zipper, S. C., West, C., Wada, Y., Taylor, R., Scanlon, B., Rosolem, R., Rahman, S., Oshinlaja, N., Maxwell, R., Lo, M. H., Kim, H., Hill, M., Hartmann, A., Fogg, G., Famiglietti, J. S., Ducharne, A., De Graaf, I., Cuthbert, M., Condon, L., Bresciani, E., and Bierkens, M. F. P.: GMD perspective: The quest to improve the evaluation of groundwater representation in continental-to global-scale models, Geosci. Model Dev., 14, 7545–7571, https://doi.org/10.5194/gmd-14-7545-2021, 2021.

780 de Graaf, I. E. M., van Beek, R. L. P. H., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E. H., and Bierkens, M. F. P.:

A global-scale two-layer transient groundwater model: Development and application to groundwater depletion, Adv. Water Resour., 102, 53–67, https://doi.org/10.1016/j.advwatres.2017.01.011, 2017.

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.

785   Grafton, R. Q., Williams, J., Perry, C. J., Molle, F., Ringler, C., Steduto, P., Udall, B., Wheeler, S. A., Wang, Y., Garrick, D., and Allen, R. G.: The paradox of irrigation efficiency, Science (80-.)., 361, 748–750, https://doi.org/10.1126/science.aat9314, 2018.

Groundwater Resource Committe: Report of the ground water resource estimation commitee, Ministry of Water Resources - Government of India, 113 pp., 2009.

790   Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hanasaki, N., Yoshikawa, S., Pokhrel, Y., and Kanae, S.: A global hydrological simulation to specify the sources of water used by humans, Hydrol. Earth Syst. Sci. Discuss., 08, 1–53, https://doi.org/10.5194/hess-2017-280, 2017.

795   Hanasaki, N., Matsuda, H., Fujiwara, M., Hirabayashi, Y., Seto, S., Kanae, S., and Oki, T.: Toward hyper-resolution global hydrological models including human activities: application to Kyushu Island, Japan, Hydrol. Earth Syst. Sci. Discuss., 1–33, 2022.

Harbaugh, A. W.: MODFLOW 2005 , The U . S . Geological Survey Modular GroundWater Model: the GroundWater Flow Process, U.S. Geol. Surv. Tech. Methods, 253, 2005.

800   Hazenberg, P., Fang, Y., Broxton, P., Gochis, D., Niu, G.-Y., Pelletier, J. D., Troch, P. A., and X, Z.: A hybrid-3D hillslope hydrological model for use in Earth system models, Water Resour. Res., 51, 8218–8239, https://doi.org/10.1002/2014WR016842, 2015.

Houben, T., Pujades, E., Kalbacher, T., Dietrich, P., and Attinger, S.: From dynamic groundwater level measurements to regional aquifer parameters--Assessing the power of spectral analysis, Water Resour. Res., 2022.

805   Hughes, J. D., Langevin, C. D., and Banta, E. .: Documentation for the MODFLOW 6 Framework, USGS Tech. Methods 6-A57, 42, 2017.

Huscroft, J., Gleeson, T., Hartmann, J., and Börker, J.: Compiling and Mapping Global Permeability of the Unconsolidated and Consolidated Earth: GLobal HYdrogeology MaPS 2.0 (GLHYMPS 2.0), Geophys. Res. Lett., 45, 1897–1904, https://doi.org/10.1002/2017GL075860, 2018.

810   Jing, M., Heße, F., Kumar, R., Wang, W., Fischer, T., Walther, M., Zink, M., Zech, A., Samaniego, L., Kolditz, O., and Attinger, S.: Improved regional-scale groundwater representation by the coupling of the mesoscale Hydrologic Model (mHM v5.7) to the groundwater model OpenGeoSys (OGS), Geosci. Model Dev., 11, 1989–2007, https://doi.org/10.5194/gmd-11-1989-2018, 2018.

Karner, K., Mitter, H., and Schmid, E.: The economic value of stochastic climate information for agricultural adaptation in a

815 semi-arid region in Austria, J. Environ. Manage., 249, 109431, https://doi.org/10.1016/j.jenvman.2019.109431, 2019.

Karutz, R., Omann, I., Gorelick, S. M., Klassert, C. J. A., Zozmann, H., Zhu, Y., Kabisch, S., Kindler, A., Figueroa, A. J., Wang, A., and Küblböck, K.: Capturing Stakeholders' Challenges of the Food–Water–Energy Nexus—A Participatory Approach for Pune and the Bhima Basin, India, Sustain., 14, 2022.

Keune, J., Gasper, F., Goergen, K., Hense, A., Shrestha, P., Sulis, M., and Kollet, S.: Studying the influence of groundwater
820 representations on land surface-atmosphere feedbacks during the European heat wave in 2003, J. Geophys. Res., 121, 13,301-13,325, https://doi.org/10.1002/2016JD025426, 2016.

Keune, J., Sulis, M., Kollet, S., Siebert, S., and Wada, Y.: Human Water Use Impacts on the Strength of the Continental Sink for Atmospheric Water, Geophys. Res. Lett., 45, 4068–4076, https://doi.org/10.1029/2018GL077621, 2018.

Koirala, S., Jung, M., Reichstein, M., de Graaf, I. E. M., Camps-Valls, G., Ichii, K., Papale, D., Ráduly, B., Schwalm, C. R.,
825 Tramontana, G., and Carvalhais, N.: Global distribution of groundwater-vegetation spatial covariation, Geophys. Res. Lett., 44, 4134–4142, https://doi.org/10.1002/2017GL072885, 2017.

Koirala, S., Kim, H., Hirabayashi, Y., Kanae, S., and Oki, T.: Sensitivity of Global Hydrological Simulations to Groundwater Capillary Flux Parameterizations, Water Resour. Res., 55, 402–425, https://doi.org/10.1029/2018WR023434, 2019.

Kollet, S. J. and Maxwell, R. M.: Capturing the influence of groundwater dynamics on land surface processes using an
830 integrated, distributed watershed model, Water Resour. Res., 44, 1–18, https://doi.org/10.1029/2007WR006004, 2008.

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

Long, D., Yang, W., Scanlon, B. R., Zhao, J., Liu, D., Burek, P., Pan, Y., You, L., and Wada, Y.: South-to-North Water Diversion stabilizing Beijing's groundwater levels, Nat. Commun., 11, https://doi.org/10.1038/s41467-020-17428-6, 2020.

835 Loritz, R., Hassler, S. K., Jackisch, C., Allroggen, N., Van Schaik, L., Wienhöfer, J., and Zehe, E.: Picturing and modeling catchments by representative hillslopes, Hydrol. Earth Syst. Sci., 21, 1225–1249, https://doi.org/10.5194/hess-21-1225-2017, 2017.

Magyar, N., Hatvani, I. G., Magyar, N., Zessner, M., and Kovács, J.: The Water Framework Directive: Can more information be extracted from groundwater data? A case study of Seewinkel, Burgenland, eastern Austria, 22,
840 https://doi.org/10.1007/s10040-013-1093-x, 2014.

Magyar, N., Hatvani, I. G., Arato, M., Trasy, B., Blaschke, A. P., and Kovacs, J.: A New Approach in Determining the Decadal Common Trends in the Groundwater Table of the Watershed of Lake Neusiedlersee., 13, https://doi.org/https://doi.org/10.3390/w13030290, 2021.

Martínez-de la Torre, A. and Miguez-Macho, G.: Groundwater influence on soil moisture memory and
845 land–atmosphere interactions in the Iberian Peninsula, Hydrol. Earth Syst. Sci. Discuss., 1–37, https://doi.org/10.5194/hess-2018-626, 2019.

Maxwell, R. M., Condon, L. E., and Kollet, S. J.: A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3, Geosci. Model Dev., 8, 923–937,

https://doi.org/10.5194/gmd-8-923-2015, 2015.

850 Meixner, T., Manning, A. H., Stonestrom, D. A., Allen, D. M., Ajami, H., Blasch, K. W., Brookfield, A. E., Castro, C. L., Clark, J. F., Gochis, D. J., Flint, A. L., Neff, K. L., Niraula, R., Rodell, M., Scanlon, B. R., Singha, K., and Walvoord, M. A.: Implications of projected climate change for groundwater recharge in the western United States, J. Hydrol., 534, 124–138, https://doi.org/10.1016/j.jhydrol.2015.12.027, 2016.

Meredith, E. and Blais, N.: Quantifying irrigation recharge sources using groundwater modeling, Agric. Water Manag., 214,
855 9–16, https://doi.org/10.1016/j.agwat.2018.12.032, 2019.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., and Schmitt, O.: Estimating the volume and age of water stored in global lakes using a geo-statistical approach, Nat. Commun., 7, 1–11, https://doi.org/10.1038/ncomms13603, 2016.

Miguez-Macho, G. and Fan, Y.: The role of groundwater in the Amazon water cycle: 2. Influence on seasonal soil moisture and evapotranspiration, J. Geophys. Res. Atmos., 117, https://doi.org/10.1029/2012JD017540, 2012.

860 Mitter, H. and Schmid, E.: Informing groundwater policies in semi-arid agricultural production regions under stochastic climate scenario impacts, Ecol. Econ., 180, 106908, https://doi.org/10.1016/j.ecolecon.2020.106908, 2021.

Ochoa, C. G., Fernald, A. G., Guldan, S. J., Tidwell, V. C., and Shukla, M. K.: Shallow Aquifer Recharge from Irrigation in a Semiarid Agricultural Valley in New Mexico, J. Hydrol. Eng., 18, 1219–1230, https://doi.org/10.1061/(asce)he.1943-5584.0000718, 2013.

865 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, Geosci. Model Dev., 12, 2401–2418, https://doi.org/10.5194/gmd-12-2401-2019, 2019a.

Reinecke, R., Foglia, L., Mehl, S., Herman, J. D., Wachholz, A., Trautmann, T., and Döll, P.: Spatially distributed sensitivity of simulated global groundwater heads and flows to hydraulic conductivity, groundwater recharge, and surface water body
870 parameterization, Hydrol. Earth Syst. Sci., 23, 4561–4582, https://doi.org/10.5194/hess-23-4561-2019, 2019b.

Reinecke, R., Wachholz, A., Mehl, S., Foglia, L., Niemann, C., and Döll, P.: Importance of Spatial Resolution in Global Groundwater Modeling, 58, 363–376, https://doi.org/10.1111/gwat.12996, 2020.

Roark, M. and Healy, D. F.: Quantification of deep percolation from two flood irrigated alfalfa fields, roswell basin, new mexico, Albuquerque, New Mexico, 32 pp., 1998.

875 Rouholahnejad Freund, E. and Kirchner, J. W.: A Budyko framework for estimating how spatial heterogeneity and lateral moisture redistribution affect average evapotranspiration rates as seen from the atmosphere, Hydrol. Earth Syst. Sci., 21, 217–233, https://doi.org/10.5194/hess-21-217-2017, 2017.

Russcher, M. J., Hofer, J., and Hughes, J. D.: xmipy--Python bindings for the eXtended Model Interface version 1.0.0 (1.0.0), https://doi.org/10.5281/zenodo.4740983, 2020.

880 Sadki, M., Munier, S., Boone, A., and Ricci, S.: Implementation and sensitivity analysis of a Dam-Reservoir OPeration model (DROP v1 . 0) over Spain, Geosci. Model Dev. Discuss. [preprint], https://doi.org/https://doi.org/10.5194/gmd-2022-108, 2022.

Scanlon, B. R., Reedy, R. C., Stonestrom, D. A., Prudic, D. E., and Dennehy, K. F.: Impact of land use and land cover change on groundwater recharge and quality in the southwestern US, Glob. Chang. Biol., 11, 1577–1593, https://doi.org/10.1111/j.1365-2486.2005.01026.x, 2005.

Schaller, M. F. and Fan, Y.: River basins as groundwater exporters and importers: Implications for water cycle and climate modeling, J. Geophys. Res. Atmos., 114, https://doi.org/10.1029/2008JD010636, 2009.

Scheidegger, J. M., Jackson, C. R., Muddu, S., Tomer, S. K., and Filgueira, R.: Integration of 2D lateral groundwater flow into the variable infiltration capacity (VIC) model and effects on simulated fluxes for different grid resolutions and aquifer diffusivities, 13, https://doi.org/10.3390/w13050663, 2021.

Shangguan, W., Hengl, T., de Jesus, J. M., Yuan, H., and Dai, Y.: Mapping the global depth to bedrock for land surface modeling, J. Adv. Model. Earth Syst., 8, 1248–1269, https://doi.org/10.1002/2016MS000689, 2016.

Shrestha, P., Sulis, M., Simmer, C., and Kollet, S.: Effects of horizontal grid resolution on evapotranspiration partitioning using TerrSysMP, J. Hydrol., 557, 910–915, https://doi.org/10.1016/j.jhydrol.2018.01.024, 2018.

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., and Portmann, F. T.: Groundwater use for irrigation - A global inventory, Hydrol. Earth Syst. Sci., 14, 1863–1880, https://doi.org/10.5194/hess-14-1863-2010, 2010.

Smilovic, M., Guillaumot, L., Burek, P., de Bruijn, J., Tang, T., and Fridman, D.: Refining human influence in CWatM., Geosci. Model Dev. [In Prep., 2022.

Soltani, M., Bjerre, E., Koch, J., and Stisen, S.: Integrating remote sensing data in optimization of a national water resources model to improve the spatial pattern performance of evapotranspiration, J. Hydrol., 603, 127026, https://doi.org/10.1016/j.jhydrol.2021.127026, 2021.

Surinaidu, L., Bacon, C. G. D., and Pavelic, P.: Agricultural groundwater management in the Upper Bhima Basin, India: Current status and future scenarios, Hydrol. Earth Syst. Sci., 17, 507–517, https://doi.org/10.5194/hess-17-507-2013, 2013.

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

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

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., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M. W., Vannametee, E., Wisser, D., and Bierkens, M. F. P.: PCR-GLOBWB 2: a 5arcmin Global Hydrological and Water Resources Model, Geosci. Model Dev., 11, 2429–2453, https://doi.org/10.5194/gmd-11-2429-2018, 2018.

Swenson, S. C., Clark, M., Fan, Y., Lawrence, D. M., and Perket, J.: Representing Intra-Hillslope Lateral Subsurface Flow in the Community Land Model, J. Adv. Model. Earth Syst., https://doi.org/10.1029/2019MS001833, 2019.

Szilagyi, J., Zlotnik, V. A., and Jozsa, J.: Net recharge vs. depth to groundwater relationship in the platte river Valley of

Nebraska, United States, 51, 945–951, https://doi.org/10.1111/gwat.12007, 2013.

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.

920   M., Yechieli, Y., Gurdak, J. J., Allen, D. M., Shamsudduha, M., Hiscock, K., Yeh, P. J.-F., Holman, I., and Treidel, H.: Ground water and climate change, Nat. Clim. Chang., 3, 322, 2012.

Trichakis, I., Burek, P., de Roo, A., and Pistocchi, A.: Towards a Pan-European Integrated Groundwater and Surface Water Model: Development and Applications, Environ. Process., 4, 81–93, https://doi.org/10.1007/s40710-017-0216-0, 2017.

Troch, P. a, Paniconi, C., and Emiel van Loon, E.: Hillslope-storage Boussinesq model for subsurface flow and variable source

925   areas along complex hillslopes: 1. Formulation and characteristic response, Water Resour. Res., 39, 1316, https://doi.org/10.1029/2002wr001728, 2003.

Turner, S. W. D., Hejazi, M., Yonkofski, C., Kim, S. H., and Kyle, P.: Influence of Groundwater Extraction Costs and Resource Depletion Limits on Simulated Global Nonrenewable Water Withdrawals Over the Twenty-First Century, Earth's Futur., 7, 123–135, https://doi.org/10.1029/2018EF001105, 2019.

930   Verbeke, T., Tootchi, A., Jost, A., Ghattas, J., Cheruy, F., and Ducharne, A.: Subgrid-scale parametrization of groundwater-soil moisture interactions in the ORCHIDEE land surface model : first results at global scale, in: EGU General Assembly 2019, 2019.

Vergnes, J. P., Decharme, B., and Habets, F.: Introduction of groundwater capillary rises using subgrid spatial variability of topography into the ISBA land surface model, J. Geophys. Res., 119, 11,065-11,086, https://doi.org/10.1002/2014JD021573,

935   2014.

Vergnes, J. P., Roux, N., Habets, F., Ackerer, P., Amraoui, N., Besson, F., Caballero, Y., Courtois, Q., De Dreuzy, J. R., Etchevers, P., Gallois, N., Leroux, D. J., Longuevergne, L., Le Moigne, P., Morel, T., Munier, S., Regimbeau, F., Thiéry, D., and Viennot, P.: The AquiFR hydrometeorological modelling platform as a tool for improving groundwater resource monitoring over France: Evaluation over a 60-year period, Hydrol. Earth Syst. Sci., 24, 633–654, https://doi.org/10.5194/hess-

940   24-633-2020, 2020.

Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth Syst. Dyn., 5, 15–40, https://doi.org/10.5194/esd-5-15-2014, 2014.

Wada, Y., Bierkens, M. F. P., De Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., Konar, M., Liu, J., and Schmied, H. M.: Human water interface in hydrological modelling: current status and future directions, Hydrol. Earth Syst. Sci, 21,

945   4169–4193, https://doi.org/10.5194/hess-21-4169-2017, 2017.

Wang, F., Ducharne, A., Cheruy, F., Lo, M. H., and Grandpeix, J. Y.: Impact of a shallow groundwater table on the global water cycle in the IPSL land–atmosphere coupled model, Clim. Dyn., 50, 3505–3522, https://doi.org/10.1007/s00382-017-3820-9, 2018.

Willis, T. M. and Black, A. S.: Irrigation increases groundwater recharge in the Macquarie Valley, Aust. J. Soil Res., 34, 837–

950   847, https://doi.org/10.1071/SR9960837, 1996.

Wood, E. F., Roundy, J. K., Troy, T. J., Beek, L. P. H. Van, Bierkens, M. F. P., Blyth, E., Roo, A. De, Döll, P., Ek, M., Famiglietti, J., Gochis, D., Giesen, N. Van De, Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Lidard, C. P., 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 Resour. Res., 47, 1–10, https://doi.org/10.1029/2010WR010090, 2011.

Wösten, J. H. M. and van Genuchten, M. T.: Using Texture and Other Soil Properties to Predict the Unsaturated Soil Hydraulic Functions, Soil Sci. Soc. Am. J., 52, 1762–1770, https://doi.org/10.2136/sssaj1988.03615995005200060045x, 1988.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., and Pavelsky, T. M.: MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset, Water Resour. Res., 55, 5053–5073, https://doi.org/10.1029/2019WR024873, 2019.