Geoscientific
Model Development

Model description paper

GEB v0.1: a large-scale agent-based socio-hydrological model – simulating 10 million individual farming households in a fully distributed hydrological model

Jens A. de Bruijn[1,2], **Mikhail Smilovic**[1], **Peter Burek**[1], **Luca Guillaumot**[1], **Yoshihide Wada**[1,3], **and Jeroen C. J. H. Aerts**[2]

[1]International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
[2]Institute for Environmental Studies, Vrije Universiteit Amsterdam (VU Amsterdam), De Boelelaan 1087,
1081 HV Amsterdam, the Netherlands
[3]Department of Physical Geography, Utrecht University, Utrecht, the Netherlands

Correspondence: Jens A. de Bruijn (jens.de.bruijn@vu.nl)

Abstract. Humans play a large role in the hydrological system, e.g. by extracting large amounts of water for irrigation, often resulting in water stress and ecosystem degradation. By implementing large-scale adaptation measures, such as the construction of irrigation reservoirs, water stress and ecosystem degradation can be reduced. Yet we know that many decisions, such as the adoption of more effective irrigation techniques or changing crop types, are made at the farm level by a heterogeneous farmer population. While these decisions are usually advantageous for an individual farmer or their community, aggregate effects of those decisions can have large effects downstream. Similarly, decisions made by other stakeholders, such as governments, often have basin-wide effects and affect each farmer differently. To fully comprehend how the human–natural water system evolves over time and space and to explore which interventions are suitable to reduce water stress, it is important to consider human behaviour and feedbacks to the hydrological system simultaneously at the local household and large basin scales. Therefore, we present the Geographical, Environmental, and Behavioural (GEB) model, a coupled agent-based hydrological model that simulates the behaviour and daily bidirectional interaction of more than 10 million individual farm households with the hydrological system on a personal laptop. Farmers exhibit autonomous heterogeneous behaviour based on their characteristics, assets, environment, management policies, and social network. Examples of behaviour are irriga-

tion, generation of income from selling crops, and investment in adaptation measures. Meanwhile, reservoir operators manage the amount of water available for irrigation and river discharge. All actions can be taken at a daily time step and influence the hydrological system directly or indirectly. GEB is dynamically linked with the spatially distributed grid-based hydrological model CWatM at $30''$ resolution (< 1 km at the Equator). Because many smallholder farm fields are much smaller than 1×1 km, CWatM was specifically adapted to implement dynamically sized hydrological response units (HRUs) at the farm level, providing each agent with an independently operated hydrological environment. While the model could be applied anywhere globally at both large and small scales, we explore its implementation in the heavily managed Krishna basin in India, which encompasses $\sim 8\%$ of India's land area and ~ 12.1 million farmers.

1 Introduction

Water stress, defined as water demand exceeding water availability, is an increasing threat to human livelihood through, for example, decreasing agricultural yields, insufficient water for drinking and sanitation, and degrading ecosystems (Ablo and Yekple, 2018; van Leeuwen et al., 2016; Porporato et al., 2001; Kummu et al., 2016). A growing number of regions are expected to experience severe water stress in the

future, largely driven by an increasing population and climate change (Kummu et al., 2016; Veldkamp et al., 2015). Effective water management can help to reduce water stress, but this requires knowledge of the current status of water resources, socio-economic development, climate change, and the effects of interventions (Ibisch et al., 2016) on upstream and downstream water availability (Veldkamp et al., 2017). Therefore, hydrological models, which simulate the hydrological system, are a widely used tool to provide an integrative vision of the system and formulate effective policies.

Humans play a large role in the hydrological system. For example, governments and other organizations construct reservoirs (Biggs et al., 2007) and channels for inter-basin transfers (Gupta and van der Zaag, 2008), disrupting natural flows. Small-scale adaptations, such as groundwater pumping (Llamas and Martínez-Santos, 2005), rainwater harvesting (Li et al., 2000), changing crop use (Kuil et al., 2018), and irrigation practices (Nouri et al., 2019b; Mollinga, 2003), are often realized at the individual or communal level. While these measures are usually beneficial for some, adverse effects can be experienced by other water users across different scales (Di Baldassarre et al., 2021). In addition, the costs and benefits of water-stress-related interventions may vary throughout a heterogeneous farmer population.

To fully comprehend how water stress develops over time and space and to explore which interventions are suitable to reduce water stress, it is important to understand feedbacks in a coupled human–natural system simultaneously with local household and large basin scales. For example, when water is extracted upstream, water availability downstream can be reduced (Veldkamp et al., 2017). Farmers at the head of a command area can have access to a larger and more reliable water supply than those at the tail end (Mollinga, 2003), which incentivizes farmers at the head to adopt water-intensive high-return crops, reducing water availability downstream (Wallach, 1984). Similarly, upstream farmers that invest in rainwater-harvesting techniques reduce the amount of water available downstream (Bouma et al., 2011). Another example is through increased groundwater use, where individual well users lower the groundwater table in a larger region (Llamas and Martínez-Santos, 2005). While some farmers have the resources to invest in deeper wells, other farmers are driven further into poverty (Batchelor et al., 2003).

While most hydrological models are well suited to simulate the hydrological system at a large scale, they treat small-scale human behaviour rather simplistically and homogeneously. In these models, humans often do not learn over time and do not change their adaptive behaviour under changing water risk (Aerts et al., 2018). In reality, agents adapt to changes in their environment and also respond to each other (Wens et al., 2020). For example, a water pricing tax by the government has a direct influence on household water use, and farmers might construct wells in response to drought events. An agent-based model (ABM) appears to be an effective tool that can be used to simulate these complex heterogeneous behaviours and feedbacks. Therefore, the research realm of "socio-hydrology" has developed models that dynamically couple hydrological and agent-based models to better simulate the hydrological system as well as the behaviour of individual heterogeneous agents. Using such a coupled model allows for tracking changes in the natural system (e.g. the effect of changes in climate) or changes in the human system (e.g. government policies or adaptation behaviour) through the entire human–natural system. For example, drought events can accelerate adaptation behaviour, making farmers more resilient to the next drought. At the same time, such adaptation behaviour can negatively influence water storage (e.g. increased groundwater extraction; Streefkerk et al., 2023). Simulating such feedbacks requires a coupled model.

In general, two approaches can be differentiated where a hydrological component is added to an agent-based model: (i) adding a hydrological component which is also agent based (e.g. river segments are represented as agents which exchange water) or (ii) adding a traditional hydrological model; thus, a fully distributed gridded model where water flows from one grid cell to another based using a routing equation. In the first agent-based approach, all the environmental components, such as river segments, are simulated as agents. For example, Becu et al. (2003) simulated farmers, irrigation behaviour, and crop and vegetation dynamics. Their model uses a simple routing scheme that considers water abstraction and water diversions by canal managers. Another example is Huber et al. (2019), who created a basin-scale coupled model where water flows downstream from a hydrological river agent to another river agent, while other agents such as farmers or water managers can abstract water from the river. In this approach, the hydrological component is usually relatively simple, largely because authors usually build the hydrological component from scratch. (ad. ii) The second – hydrological model – approach is to couple an agent-based model with a more traditional hydrological model by allowing the agents to interact with its water storage (Streefkerk et al., 2023). For example, the widely used MP-MAS (Schreinemachers and Berger, 2011; Arnold et al., 2015) is coupled to WASIM-ETH (Schulla and Jasper, 2007), a fully distributed hydrological model. Van Oel et al. (2010) published a larger coupled grid-based model at a 270 m resolution that simulates the irrigation behaviour of individual farmers in a large basin using a grid size of 270×270 m. This approach also benefits from ongoing methodological progress in hydrological modelling (Bierkens, 2015). Large-scale hydrological models are run at an increasingly higher resolution, while other advances, such as HydroBlocks (Chaney et al., 2016), allow users to effectively combine grid cells into hydrological response units (HRUs) while retaining the ability to accurately simulate the hydrological system, including dynamic routing (Chaney et al., 2021).

Some agent-based models with a hydrological component were released, which simulate water management decisions by groups of people, such as sectors or villages (e.g. Huber et al., 2019; Streefkerk et al., 2023). Other agent-based models represent single water users, such as a person or household (Schreinemachers and Berger, 2011; Wens et al., 2020; Becu et al., 2003; Arnold et al., 2015). These models are better suited to simulate individual adaptation pathways, which are often paramount in capturing the heterogeneity of the farmer population (Wens et al., 2020; e.g. Bert et al., 2011; Tamburino et al., 2020). Yet, to simulate the effect of a single agent on the hydrological system, at least one HRU per agent is required to properly represent system feedback at the individual level (Schreinemachers and Berger, 2011). Using a gridded model, this means that the grid cell size cannot be larger than the smallest farm. This requires a large computation time and computational resources, especially in regions with smallholder farms. So far, this has limited the ability of coupled models that capture the full heterogeneity of the agent population to be effectively applied on a large scale.

We propose to resolve this issue by simulating hydrological processes at the field scale using dynamically sized HRUs within a grid cell, with each HRU representing a single farm. Each farm-level HRU can be individually operated by an agent. This way, each individual crop field is simulated as a HRU in addition to other land use types. Due to their dynamic size (e.g. 1 unit for 50 % of the cell and 10 units each representing 5 % of the cell), CWatM can be run at a relatively coarse resolution, such as $30''$ (< 1 km at the Equator) to simulate a large hydrological basin while allowing simulation of small and individually operated farms. Because agents can directly interact with these units (their fields), we can, for the first time, investigate the interaction between small-scale individual behaviour and large basin-wide hydrological processes.

Therefore, we present the Geographical, Environmental, and Behavioural (GEB) model, named after Geb, the Egyptian god of the earth. The model is an agent-based model (ABM) that is dynamically linked with a specifically adapted version of the Community Water Model (CWatM; Burek et al., 2020). GEB can simulate large-scale hydrological processes as well as the individual behaviour of more than 10 million individual farming households one-to-one and their bidirectional interactions with the hydrological system. CWatM is used as a coupled model to simulate the hydrological cycle at a grid resolution of $30''$ (< 1 km at the Equator). Individual farmer households (~ 12.2 million on a normal laptop with 10 GB of free RAM) and reservoir operators are simulated as fully integrated agents that can dynamically interact with (i.e. respond to and influence) the water balance in CWatM. Through this coupling, each individual farmer can, at a daily time step, decide to irrigate from various sources (i.e. surface, reservoir, or groundwater). Furthermore, farmers can decide to plant and harvest crops based on the available water in their environment, the status of their

crops, their risk aversion, crop price, water price, weather conditions, etc. Moreover, farmers can adapt, for example, by investing in water-saving techniques, drilling boreholes, and changing crop type. All these decisions can be made at a daily time step.

In this work, we describe how the open-source model is set up, followed by an example application model in the heavily managed Krishna basin, which encompasses $\sim 257\,000$ km^2 or ~ 8 % of India's land area. Here, we simulate the adaptive patterns to water stress of ~ 12.1 million farmers and show how adaptation through irrigation efficiency and crop choice can influence both individual farmers as well as the hydrological system through various artificial storylines. All model code is extensively documented on https://jensdebruijn.github.io/GEB/ (last access: 2 May 2023).

2 Model description

The GEB model is an open-source coupled hydrological and agent-based model jointly developed at the International Institute for Applied Systems Analysis (IIASA) and the Institute for Environmental Studies (IVM, VU Amsterdam) and developed in Python. The agent-based model can simulate millions of individual farmers in addition to other agents that interact bidirectionally with the hydrological CWatM model (Burek et al., 2020). In this manner, GEB simulates the water cycle and how this interacts with the individual decision-making of farmers and water managers, such as crop management and growth, and irrigation and reservoir management in large (or small) basins at a daily time step. The model can be adapted to run various scenarios (Fig. 1), influencing the ABM (e.g. provision of subsidies to farmers or the construction of additional reservoirs) or the water cycle (e.g. varying future climate scenarios).

Figure 2 shows further detail on the main model interlinkages between the human and hydrological components. The model consists of two dynamically integrated parts, which run synchronized at a daily time step: a field-scale hydrological model (blue; see Sect. 2.1 for details) and an agent-based model that simulates crop farmers and reservoir managers (orange; see Sect. 2.2 for details).

During each time step, the model is forced by a daily set of *meteorological data*, considering the distribution of *land use* and *crops*. *Potential evapotranspiration* is determined for both cropland and non-cropland, which is followed by the determination of both the *water availability* and *potential demand*. Potential irrigation demand for non-paddy irrigation is computed as the difference between current soil moisture content and the soil moisture content at field capacity in the root zone, limited by the infiltration capacity. The potential irrigation demand for paddy-irrigated land is computed as the difference between the current water level in the paddies and the targeted water level. Here, *reservoir operators* can release
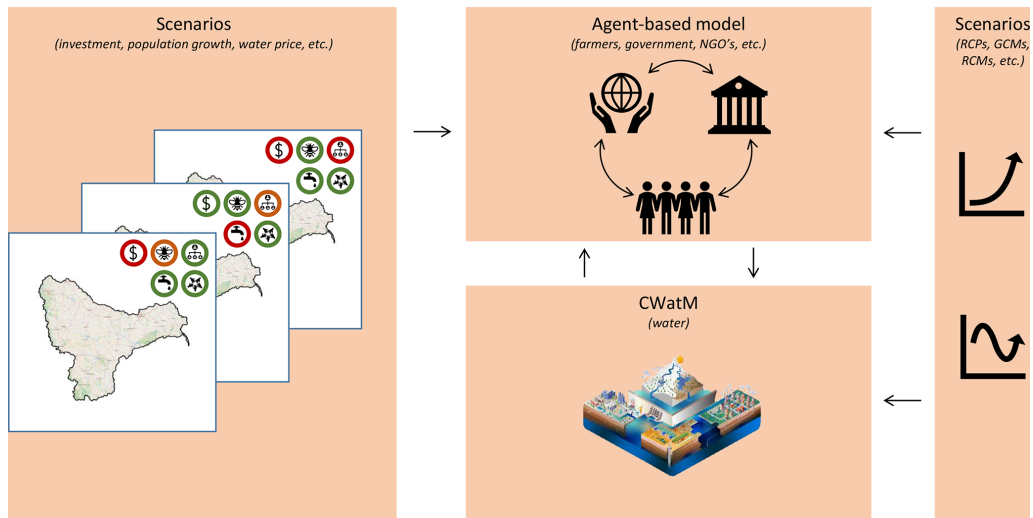
Figure 1. GEB: high-level interaction between CWatM and the agent-based model. © OpenStreetMap contributors 2022. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.
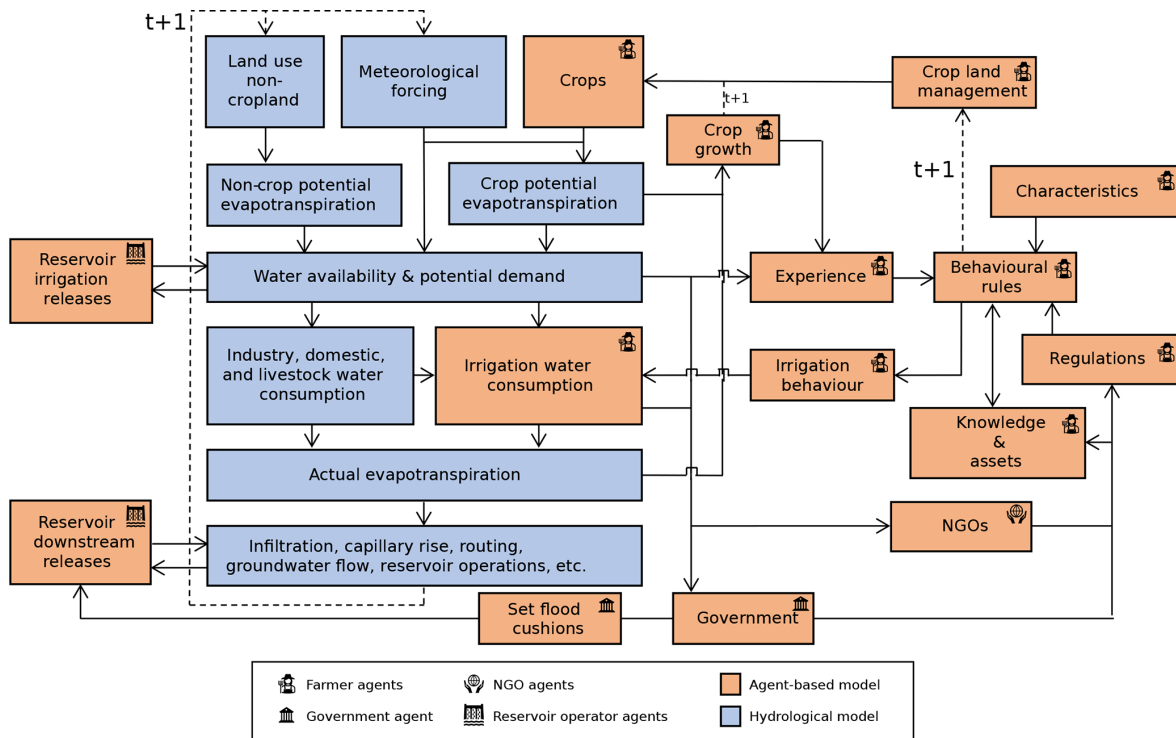

Figure 2. A schematic overview of GEB.

water from their reservoir based on considerations such as current demand and water levels. Then, after *water consumption* by industrial, domestic, and livestock sectors, farmers can abstract irrigation water.

Here, the calculation of water consumption is removed from CWatM and is instead calculated in the agent-based model, addressing irrigation behaviour, selection of crop types, and assets and characteristics of individual agents.

These factors are not necessarily static over time, as agents can invest in management options such as irrigation wells, drip irrigation, or change crop types. Moreover, external bodies such as governments and non-governmental organizations (NGOs) can influence the behaviour of farmers and reservoir operators by imposing regulations, providing knowledge to the farmer population or investing in the wider availability

of assets (e.g. create an irrigation reservoir). Knowledge can also be obtained from other (neighbouring) agents.

After the application of irrigation water, CWatM simulates infiltration, capillary rise within soils, groundwater recharge, surface routing, and groundwater flow (using MODFLOW; Langevin et al., 2017). Here, CWatM again communicates with the reservoir operator agents to determine the amount of water released downstream. Then, as a new time step is initiated, each farmer can decide to plant or harvest crops based on *experience*, *assets*, *characteristics*, *knowledge*, and *regulations*. The land use classes in CWatM are then updated accordingly. Finally, the next time step is initiated, starting with meteorological forcing, as described above.

The model is implemented in Python 3, with all computationally intensive parts written in compiled Python libraries such as NumPy (Harris et al., 2020) and Numba (Lam et al., 2015), and it includes optional GPU vectorization of soil components through CuPy. The model can be run on all major platforms (i.e. Linux, Windows, and Mac). An optional model interface is extended from Mesa (Kazil et al., 2020; Fig. 3). A high-level description of the technical model integration can be found in Appendix A, while further details can be found in the model documentation.

Simulating hydrological processes at the field scale

Most hydrological models implement several different land use types (e.g. Burek et al., 2020; Sutanudjaja et al., 2018; Müller Schmied et al., 2021). In these models, soil processes in all land use types are simulated individually. Runoff and several other hydrological fluxes are then computed by aggregating to the grid cell level while considering the relative size of each land use type in a grid cell. In other words, each land use type within a grid cell is simulated as a HRU (Flügel, 1997; Chaney et al., 2016). Farmers usually occupy cropland land use types such as non-irrigated land, paddy-irrigated land, and non-paddy irrigated land (Burek et al., 2020; Sutanudjaja et al., 2018; Hanasaki et al., 2018; Alcamo et al., 2003). When a single land use type within a grid cell is occupied by multiple farmers, these farmers share a HRU (i.e. hydrological environment) and are thus simulated as a single unit (of multiple farmers).

This introduces an issue for agent-based models that focus on the implementation of heterogeneous decision-making at the field scale. For example, when two farmers share a HRU and farmer no. 1 decides to irrigate while farmer no. 2 does not, the soil moisture in the field of farmer no. 1 should increase relative to the soil moisture in the field of farmer no. 2. However, when both farmers share a HRU, the soil moisture in their field cannot be separately simulated in the model.

The most straightforward solution is to run the model at a higher resolution, such that the smallest field is simulated as a single grid cell while larger fields are simulated as multiple grid cells. However, as small farms of less than 1 ha make up 72 % of global farms (Lowder et al., 2016), this solution

requires the use of grid cells that are less than 100×100 m, which would use an enormous amount of computational resources, making the approach unfeasible in larger basins.

As a solution, we simulate the field of each farmer as a single HRU and adapt CWatM to be able to work with these HRUs (Fig. 4). In this concept, cropland land use types are further subdivided into dynamically sized HRUs based on the land ownership (or rent) of the agent (e.g. a farmer). These HRUs can be independently operated by agents in the ABM, such as farmers. In this manner, the land management decisions (e.g. crop planting date and irrigation) and soil processes (e.g. percolation, capillary rise, and evaporation) are independently simulated in a HRU for each farmer, thus allowing for simulation of multiple independently operated farms within a single grid cell. These HRUs can also be split, allowing, for example, for farmland expansion into other land use types and the sale of (part of) a farmer's land.

Each crop farm that is owned by a farmer is thus an individual HRU. An exception is when a crop farm is spread across multiple grid cells, in which case it is represented by multiple HRUs across those grid cells. However, as these split HRUs are owned by a single farmer, management decisions still affect all HRUs and thus the entire farm. In addition, each land use type that is not operated by crop farmers in a grid cell is a separate HRU, thus operating independently from other land use types, such as water areas, grasslands, and forests.

While most primarily vertical hydrological fluxes (e.g. infiltration, percolation) occur within HRUs, river discharge and groundwater flow are simulated at the grid cell level. To this extent, conversion of fluxes between HRUs and grid cells is required. Figure 5 shows how this works in practice, similar to hydrological models that simulate multiple land use types in a single grid cell, such as CWatM and PCR-GLOBWB (Sutanudjaja et al., 2018). Runoff is first determined per HRU and then aggregated to the grid cell level while considering the relative size of each HRU. Aggregated runoff is then added to discharge, followed by solving the kinematic wave equation at the grid cell level.

For the implementation in the Krishna basin, we selected the region basin for the study using the globally available MERIT Hydro elevation map (Yamazaki et al., 2019), which was upscaled to $30''$ by using the iterative hydrography upscaling method (Eilander et al., 2021) and by subsequently selecting all upstream cells of the Krishna River outlet. Other routing maps, such as river slope and width, were obtained similarly (Eilander et al., 2020). Reservoir and lake footprints were obtained from the HydroLAKES dataset (Messager et al., 2016). If available, flood cushions and reservoir volumes were obtained from the Andhra Pradesh WRIMS (https://apwrims.ap.gov.in/, last access: 7 September 2021) database. If not available, flood cushions were assumed to be zero, while reservoir volumes were taken from the original HydroLAKES data. Reservoir command areas were obtained from the India Water Resources Information Sys-
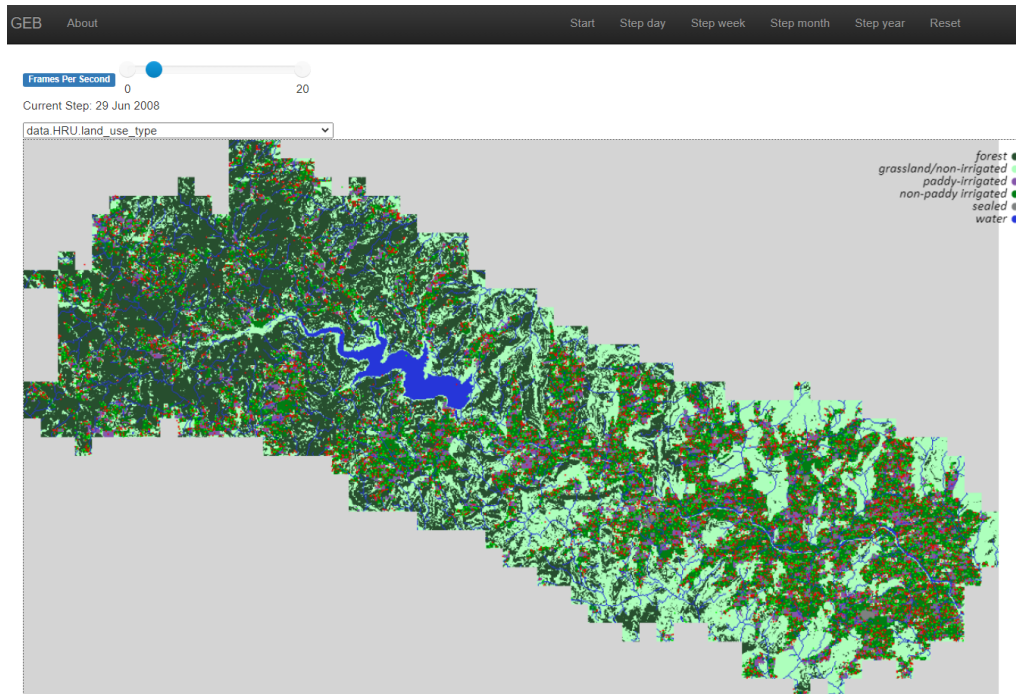
Figure 3. Optional model interface. The model can be run for one more time step, and one can show all model variables on a map. Here, the land use type is shown. The red dots represent farmer agents. Optional line charts can be added to show variables like discharge and mean groundwater level over time. For visualization purposes only, a small subbasin northwest of Pune (Maharashtra) is shown here. Land use map is derived from Jun et al. (2014).
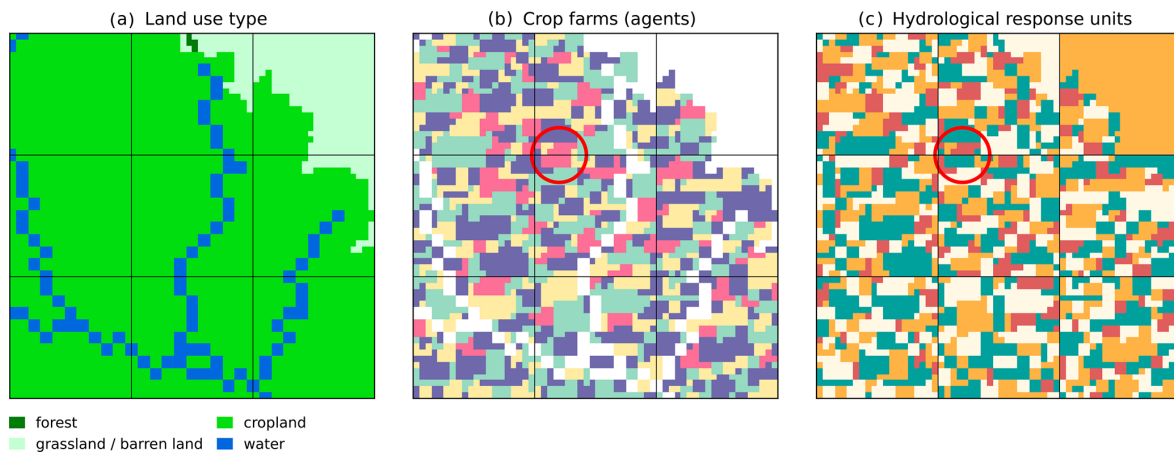

Figure 4. In this figure, 3×3 grid cells are shown, delineated by horizontal and vertical black lines ($30''$ resolution). Panel **(a)** displays various land use types at 20 times higher subgrid resolution, panel **(b)** shows the crop farms owned by agents, and panel **(c)** shows the resulting HRUs. Each contiguous area of one colour in panel **(b)** represents a farm, while each contiguous area of one colour in panel **(c)** represents a HRU. One exception is non-crop HRUs of the same land use type within a grid cell, which belong to one HRU (e.g. all rivers within a grid cell are 1 HRU). Crop farms owned by one farmer that cross grid cell boundaries are represented by multiple HRUs; see, for example, the crop farm in the centre of the red circle. Land use map is derived from Jun et al. (2014).

tem (India-WRIS, https://indiawris.gov.in/wris/, last access: 7 September 2021) and manually linked to the previously obtained reservoir using satellite imagery. Reservoir operator agents are assumed to release a maximum fraction of the current reservoir volume for irrigation, limited by the irriga-

tion demand in the command area. Land use was obtained at 30 m resolution from GlobeLand30, downscaled to $1.5''$ and mapped to CWatM land use types. Pixels that were classified as "waterbody" in GlobeLand30 and all cells with at least $100\,\text{km}^2$ upstream area were classified as "water cov-
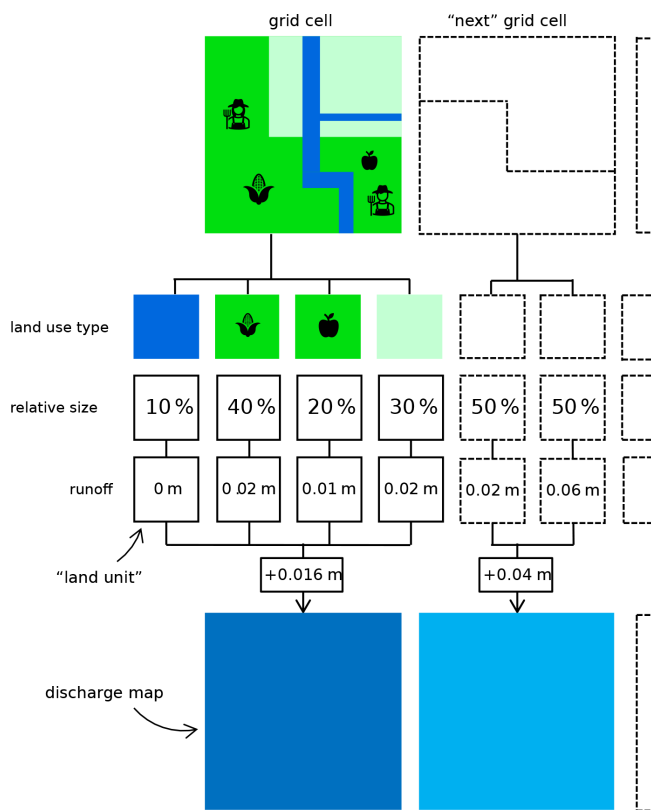
Figure 5. Schematic overview of the implementation of farm-level HRUs. Here, a grid cell consists of four HRUs: one water-covered area, two crop farms, and one grass-covered area (see coloured land use types). Runoff is determined per HRU and then aggregated by considering the relative size of the HRUs to compute runoff for the entire grid cell.

ered area" in CWatM. All other input data were obtained from CWatM input maps at 5′ resolution and downscaled to 30″ for CWatM input. The groundwater MODFLOW model is defined by an orthogonal grid at a 1000 m resolution. Only one homogeneous unconfined aquifer layer is considered. One pumping well is set up in each MODFLOW cell to satisfy the water demand from farmers and other sectors.

Water demand and consumption for industrial, domestic, and livestock sectors are estimated using the approach developed by Wada et al. (2011) and are then downscaled to the size of the HRUs by distributing the demands over cells with relevant land uses: grassland for livestock demands and sealed area for industrial and domestic demands. The model was forced with GSWP3 (Dirmeyer et al., 2006), provided as part of ISIMIP3a (Warszawski et al., 2014).

3 Model application

Here, we show the application of the model in the heavily managed Krishna basin in India, simulating the behaviour of more than 12 million farmers and the water system. With a

size of roughly 8 % of India's land area, the Krishna basin in India (Fig. 6) is a complex socio-ecological system experiencing several sustainability and equity challenges particularly related to water resources management. For example, some farmers have access to large amounts of irrigation water, leading to dropping groundwater tables in the entire region, affecting small farmers without a (deep) well disproportionately. The basin is important for agricultural production while being exposed to floods, droughts, and dropping groundwater tables (Surinaidu et al., 2013). A large number of reservoirs with a total volume of approximately 42×10^9 m^3 ($\sim$ 20 % of annual rainfall) were built primarily for irrigation purposes. Farmers in a reservoir command area can access the reservoir water that is distributed through a system of canals. In addition, following the Indian Agriculture Census (http://agcensus.dacnet.nic.in/, last access: 7 October 2022), approximately 8 % of farmers have access to groundwater through a well, depending on the farm size and location. The mean farm size is $\sim$ 1.5 ha. Soils largely consist of clay, silt, and sand. The predominant rock type found in the Western Ghats is basalt.

3.1 Agents

The ABM has farmer (Sect. 3.1.1) and reservoir operator agents (Sect. 3.1.2), which can make autonomous decisions affecting the hydrological system as well as each other. Farmers and reservoir operators directly interact with CWatM.

3.1.1 Farmers

Farmer initialization

First, an initial agent population needs to be generated with heterogeneous characteristics similar to the real population living in the basin. As with most agent-based models, and in particular large-scale models, we do not have specific information for every person. Therefore, we generate a synthetic population of farmers which has statistically similar properties to the real population (income, irrigation type, household size, etc). These statistics are based on available survey data combined with regional marginal statistics using the iterative proportional fitting algorithm (IPF; Fig. 7). For the implementation in the Krishna basin, the IPF algorithm reweights survey data from the India Human Development Survey (IHDS; Desai et al., 2005) such that the overall distribution of the adjusted survey data fits the marginal distributions of farm sizes and crop types at subdistrict-level based on the Indian Agriculture Census (https://agcensus.dacnet.nic.in/, last access: 7 October 2022). Here, we consider all crops that are grown by at least 2 % of the farmer population. Because farmers with multiple crop types throughout the year are counted multiple times in the census, an adjusted version of IPF is used (Appendix B). In this manner, a heterogeneous
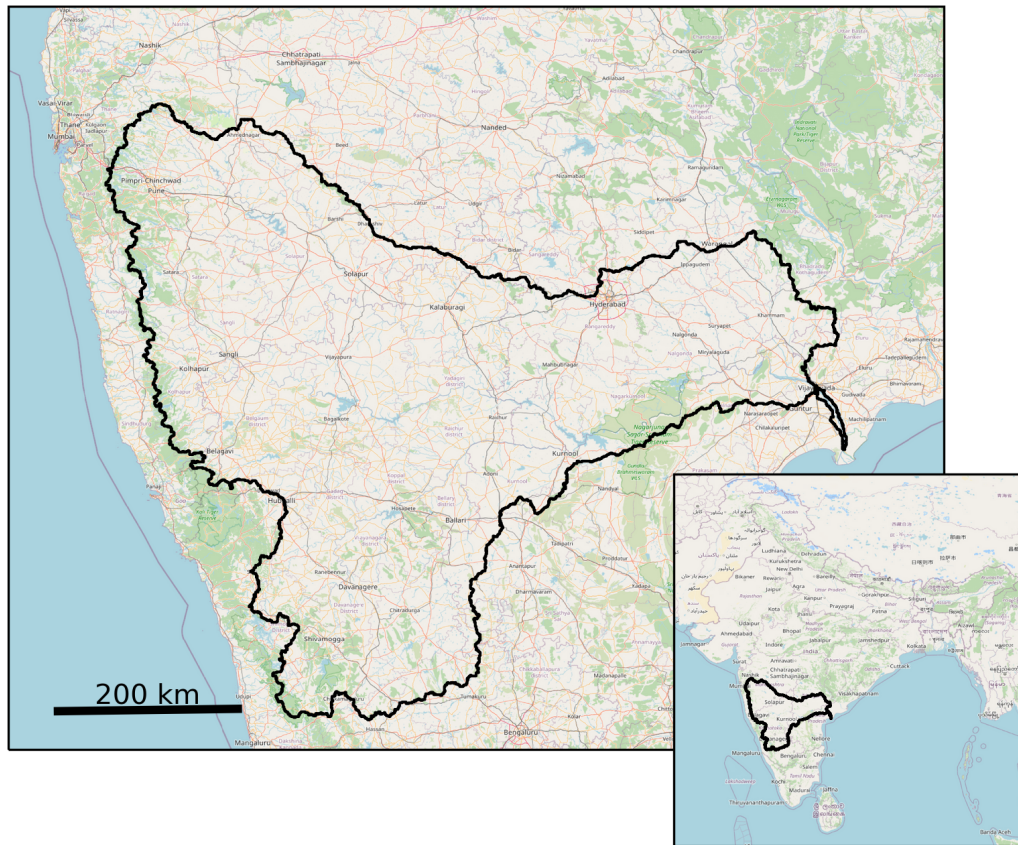
Figure 6. Outline of the Krishna basin in India. © OpenStreetMap contributors 2022. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.
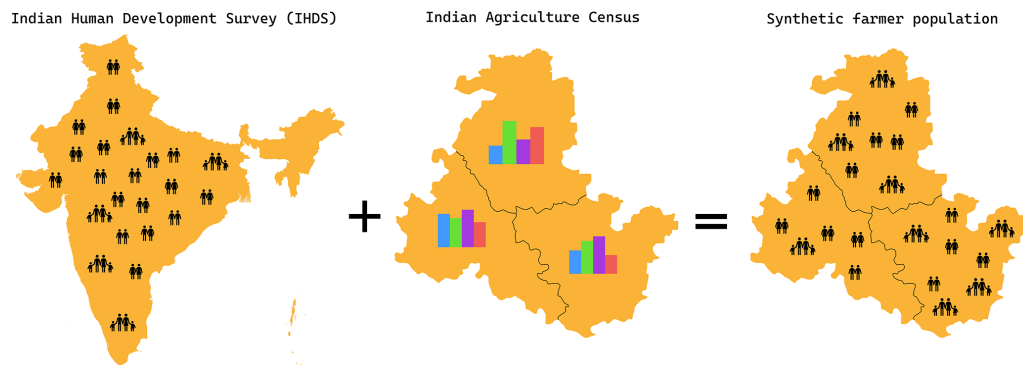

Figure 7. Schematic overview of the iterative proportional fitting (IPF) algorithm.

population of 12.2 million farmers is generated with the following characteristics: household size; crop type in kharif, rabi, and summer seasons; irrigation type; daily non-farm income; and daily consumption per capita.

Next, the farmer population is randomly distributed within their subdistrict on farmland as specified in GlobeLand30 (Jun et al., 2014), at a 1.5″ resolution (i.e. 20 times higher resolution than the CWatM grid; < 50 m at the Equator). The smallest field size is thus approximately 0.25 ha.

Cropping

The generated farmer agents grow the following crops: pearl millet, groundnut, sorghum, paddy rice, sugar cane, wheat, cotton, chickpea, maize, green gram, finger millet, sunflower and red gram. Each crop has four growth stages ($d_1, \ldots, d_4$). The crop factor (Kc) is then calculated based on the following equation (Fischer et al., 2021):

$$Kc_t = \begin{cases} Kc1, & t < d_1 \\ Kc1 + (t - d1) \times \frac{Kc2 - Kc1}{d2}, & d_1 \le t < d_2 \\ Kc2, & d_2 \le t < d_3 \\ Kc2 + (t - (d1 + d2 + d3)) \times \frac{Kc3 - Kc2}{d4}, & \text{otherwise};\end{cases} \quad (1)$$

where t is the number of days after the crop has been planted, and d_1 to d_4 represent the length of each crop stage. At the harvest stage, actual yield (Y_a) is calculated using a reference yield (Y_r; Siebert and Döll, 2010), the water-stress reduction factor (KyT), and the ratio of actual evapotranspiration (AET) to potential evapotranspiration (PET) over the entire growing period (Fischer et al., 2021).

$$Y_a = Y_r \times \left(1 - KyT \times \left(1 - \frac{\sum_{t=0}^{t=h} AET_t}{\sum_{t=0}^{t=h} PET_t}\right)\right) \quad (2)$$

All crop-specific factors used in Eqs. (1) and (2) can be found in Table C1.

Farmer income and expenses

After harvesting, it is assumed that farmers sell their crops for the state-wise market price for that month. Historic monthly market prices are obtained from Agmarknet (https://agmarknet.gov.in/, last access: 27 July 2022) for all crops except sugar cane. For sugar cane, which is brought directly to sugar cane mills, it is assumed that farmers receive the yearly indexed fair and remunerative price (FRP). These prices are fixed by the government. Yearly cultivation costs (e.g. purchasing seeds, manure, labour cost, annual depreciation) per hectare per crop type are obtained from the Ministry of Agriculture and Farmers Welfare (https://eands.dacnet.nic.in/Cost_of_Cultivation.htm, last access: 15 July 2022). Additional farmer income is obtained (e.g. from non-farming work) from the IHDS survey data. Similarly, living expenses are calculated from the daily consumption per capita in each household and household size, both available from the IHDS survey. Finally, disposable income is calculated by subtracting income and expenses.

Irrigation

During the model run, when a farmer F has irrigation equipment and is cultivating a crop, they irrigate the HRUs they own (Fig. 8).

Agents are allowed to abstract water first on a first-come, first-served principle, starting with upstream agents as determined by their elevation. As agents have no incentive to consider environmental flow conditions, these are not enforced. Farmer irrigation demand (dem) is determined by the difference between field capacity (FC) and soil moisture (SM) and is limited by infiltration capacity (IC). If farmers have access to the right equipment for surface (F_{sw}), reservoir (F_{res}), and groundwater irrigation (F_{gw}), irrigation demand (dem)

is then satisfied (Eq. 3), first from surface water (Eq. 4), then from reservoirs (Eq. 5), and finally from groundwater (Eq. 6). All sources are limited to current water availability from the streamflow ($avail_{sw}$) in grid cell G, reservoirs ($avail_{res}$) that supply the command area of grid cell G, and groundwater in grid cell G. In addition, farmers only have access to water resources if they have the relevant irrigation equipment.

$$dem_{HRU} = \min(FC_{HRU} - SM_{HRU}, IC_{HRU}); \quad (3)$$

$$irr_{HRU,sw} = \begin{cases} 0, & F_{sw} = \text{False} \\ \min\left(dem_{HRU}, avail_{sw,G}\right), & F_{sw} = \text{True};\end{cases} \quad (4)$$

$$irr_{HRU,res} = \begin{cases} 0, & F_{res} = \text{False} \\ \min\left(dem_{HRU} - irr_{HRU,sw}, avail_{res}\right), & F_{res} = \text{True};\end{cases} \quad (5)$$

$$irr_{HRU,gw} = \begin{cases} 0, & F_{gw} = \text{False} \\ \min\left(dem_{HRU} - irr_{HRU,sw} - irr_{HRU,res}, avail_{gw,G}\right), & F_{gw} = \text{True};\end{cases} \quad (6)$$

When a farmer decides to irrigate, the water is subtracted from the relevant sources in CWatM and then applied to the land in the relevant HRU.

The planting and harvesting dates are dependent on crop type and growing pattern. Once the farmer decides to harvest their crop, the respective HRU is set to "barren land" in CWatM. Then, as the farmer decides to plant a new crop, the land use type is changed accordingly in CWatM (e.g. to "irrigated").

Investing in irrigation wells

On the first day of each year, farmers can choose to invest in an irrigation well to improve their ability to irrigate their land. Here, we use the expected utility theory (e.g. Schrieks et al., 2021) to assess whether farmers make such an investment. Due to the strong social network effects (e.g. Tripathi and Mishra, 2017), we consider that farmers assess the potential benefit of installing an irrigation well based on the profit of neighbouring farmers with an identical cropping pattern but with an existing irrigation well. More specifically, we first calculate the farmer "profit ratio" (r_{own}) [0–1], defined by the ratio of actual profit to potential profit given abundant water (i.e. actual evapotranspiration is equal to potential evapotranspiration). Each farmer without an irrigation well compares their profit ratio to the profit ratio of 10 nearby farmers with an identical cropping pattern but no irrigation well ($r_{neighbours}$). The benefit of installing an irrigation well is then calculated by multiplying the difference by the current crop price (P) and reference yield (Y_r). Similarly, costs for well installation are calculated by considering the costs of a loan for that amount (Y_r), considering loan duration in years (n), the current interest rate (i), and yearly upkeep cost (U). Finally, incremental profit ($\Delta\pi$) is determined by subtracting costs and benefits. If $\Delta\pi$ is positive, the farmer invests in the
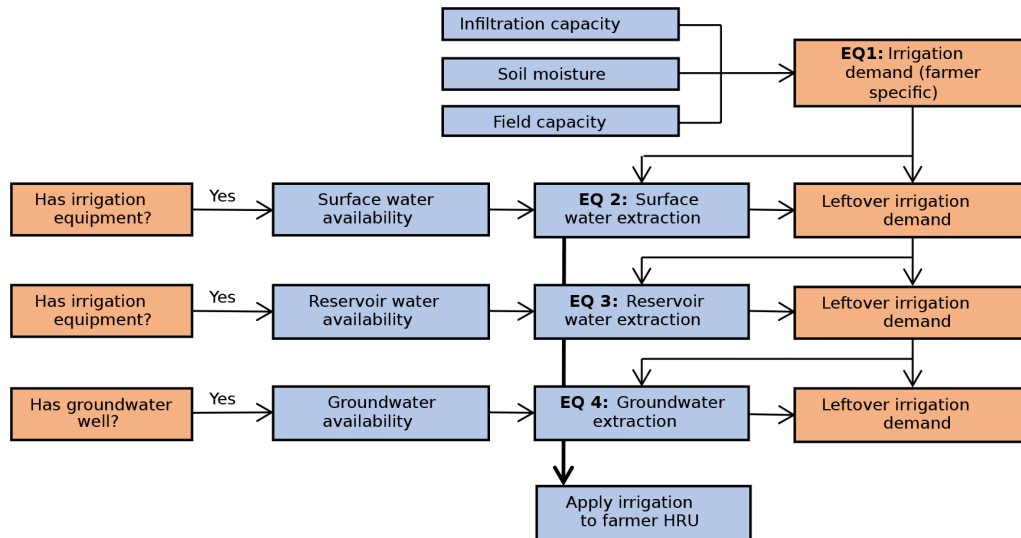
Figure 8. Farmer irrigation procedure; see also Eqs. (3)–(6).

irrigation well.

$$\Delta\pi = P \times Y_{\mathrm{r}} \times \left(r_{\mathrm{neighbours}} - r_{\mathrm{own}}\right) - \left(\frac{C}{n} + C \times i + U\right)$$

The interest rate is set to the lending interest rate in India for that year (https://data.worldbank.org/indicator/FR.INR.LEND?locations=IN, last access: 3 January 2023); tube well installation and maintenance costs for the year 2008 are set at INR 146 000 and 3000 per hectare, respectively (Sharma et al., 2008), and corrected for inflation for other years. Loan duration is set to 10 years, following offerings of several major agricultural banks in India. Finally, it is determined whether current disposable income is sufficient to pay for the loan. All installed wells are assumed to be 30 m deep.

3.1.2 Reservoir operators

The reservoir operator agents communicate with farmer agents in their reservoir command area on a daily basis and release the requested water for irrigation purposed, maximized by a maximum percentage of the current reservoir water volume. As upstream farmers get to abstract water first, this can lead to limited access to reservoir water for farmers at the tail end of command areas. The amount of water released for other purposes (e.g. maintaining outflow, reducing water level) depends on the rating curve of the reservoir and relevant flood cushions (Burek et al., 2020).

3.2 Calibration and validation

The model is calibrated based on daily river discharge from the India Water Resources Information System (WRIS) for the Wadenepally station in the Krishna River, nearby the river outlet, roughly 60 km upstream of Vijayawada. Calibration is performed based on several hydrological param-

eters (Burek et al., 2020), as well as the maximum amount of water released from a reservoir for irrigation purposes on a given day, the normal reservoir outflow, and the irrigation return fraction. Using the NSGA-II genetic algorithm (Deb et al., 2002), as implemented in DEAP (Fortin et al., 2012), the calibration procedure optimizes the modified version of the Kling–Gupta efficiency score (KGE; Eq. 5; Kling et al., 2012):

$$\mathrm{KGE'} = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}, \qquad (7)$$

where r is the correlation coefficient between monthly simulated and observed discharge, $\beta = \frac{\mu_{\mathrm{s}}}{\mu_{\mathrm{o}}}$ is the bias ratio, and $\gamma = \frac{\mathrm{CV_s}}{\mathrm{CV_o}} = \frac{\sigma_{\mathrm{s}}/\mu_{\mathrm{s}}}{\sigma_{\mathrm{o}}/\mu_{\mathrm{o}}}$ is the variability ratio; r, β, and γ are all optimal at 1.

The period 2004–2006 is used as a spin-up period, the period 2006–2012 is used for calibration, and the period 2012–2019 is used for validation. The genetic calibration algorithm first generates 60 parameter sets within a predefined range of plausible options (i.e. "the population"), and the model is subsequently run for each parameter set (i.e. "individual"). The 10 most optimal parameter sets are then combined (i.e. "mated") with a probability of 0.7 or altered (i.e. "mutated") with a probability of 0.3 to create 12 new parameter sets for which the model is also run. This procedure is repeated for 10 iterations (i.e. "generations"), and the most optimal parameter set is selected. This set is then re-run until the year 2019, and the KGE score is calculated for 2013–2019 for validation.

4 Results and discussion

Figure 9 shows observed versus simulated discharge (in $\mathrm{m^3\,s^{-1}}$) for the calibration model. The KGE during the cali-
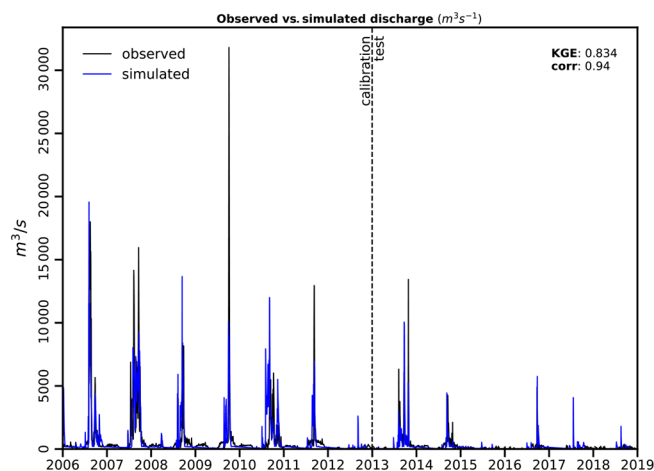
Figure 9. Observed versus simulated discharge for the calibrated model.

bration period is 0.810, while the KGE during the test period is 0.834 (1 is optimal), showing a good calibration performance for the model during both periods.

Figure 10 shows irrigation from channels, reservoirs, and groundwater for all agents for the Krishna basin. The insets show the detailed heterogeneous quantities at the field scale for a small portion of the basin. Yet, on the larger scale, it is clearly visible that farmers along rivers and within reservoir command areas have better access to water for irrigation. Differences in irrigation quantities from various sources are explained due to the location of farmers (and hence access to water from various irrigation sources), crop types, irrigation equipment, etc. For example, farmers without an irrigation well cannot access the groundwater and thus can irrigate less, while farmers with sugar cane are expected to irrigate more than other farmers.

In Fig. 11, we show how several variables change over time for specific scenarios, with the aim of showing how the model behaves. In the baseline scenario, farmers that have access to a reliable irrigation well have a 20 % probability of switching to sugar cane, a crop that generally ensures a higher income but also uses a lot of water. Here, we define a reliable irrigation well as a well that has not fallen dry in the last 3 years. The resulting mean annual discharge is shown in panel (a), while the number of farmers with sugar cane in each state gradually increases over time, as shown in panel (b). Similarly, the number of wells increases over time (panel c), due to investment in irrigation wells (see Sect. "Irrigation"). Panels (d) and (e) show the amount of irrigation from the groundwater and total irrigation from all sources. Because of the increasing number of irrigation wells, as well as other factors, groundwater abstraction generally outpaces recharge, and thus the groundwater table decreases in all four states. The mean reservoir storage also decreases over time. However, profit, when adjusted for inflation, decreases over time (panel h) as less water is available for farmers and prices

increase for crops at the market (e.g. 83 % increase in average price at the crop market for the analysed crops for Maharashtra, while general prices are 122 % higher). Finally, panels (i) and (j) show how the number of irrigation wells and profit change for farmers practicing small- and larger-scale farming. Here, smallholder farmers are the 50 % of farmers with the smallest fields, while farmers practicing large-scale farming are those with the 50 % largest fields. It is clear that smallholder farmers have less irrigation wells (panel i) at the beginning of the run but also can invest less in irrigation wells due to their limited income compared with farmers practicing large-scale farming (Sect. "Irrigation"). This is also reflected in the profits. The relative increase in profits is approximately 2 % higher for large farms over the entire simulation timeline.

In another hypothetical scenario called drip (see Fig. 11), the state of Maharashtra, one of the upstream states, provides subsidies to farmers, making 20 % of farmers switch to drip irrigation each year, corresponding with a 90 % irrigation efficiency (Brouwer and Heibloem, 1986). Here, yearly discharge (at the basin outlet) increases slightly in some years (here: 2016 only), likely because most additional water is used downstream due to the large number of reservoirs and reservoir command areas downstream of Maharashtra. The mean combined reservoir storage is slightly higher in this scenario (panel g). Interestingly, the number of installed irrigation wells (panel c) also goes up in Maharashtra, Karnataka, and Telangana, likely because fewer wells fall dry. This means that the benefit of having an irrigation well increases (mostly Maharashtra), and the means to invest in an irrigation well also increase due to higher water availability and thus higher yield (mostly Maharashtra and Karnataka). In Karnataka and Telangana, this even leads to a slight decrease in the water table.

While some adaptation options are considered here, additional adaptation options could be considered at a later stage such as crop switching and rainwater harvesting (Tamburino et al., 2020). Moreover, other factors can be included, such as threat appraisal (e.g. perception of drought risk), the coping appraisal of individual farmers (e.g. knowledge, information, and financial resources; Wens et al., 2020; Schrieks et al., 2021; Streefkerk et al., 2023), an extension of farmer networks (Wens et al., 2020), and collective adaptation behaviour by farmer cooperatives. In addition, model parameterization could benefit from the automatic delineation of smallholder fields using machine learning (Waldner and Diakogiannis, 2020) and the recognition of crop types at the field scale (Gumma et al., 2020) when such datasets become freely available in the future.

5 Conclusions

Here, we present a coupled agent-based hydrological model which for the first time allows for simulation of millions of individual households and their bidirectional interactions
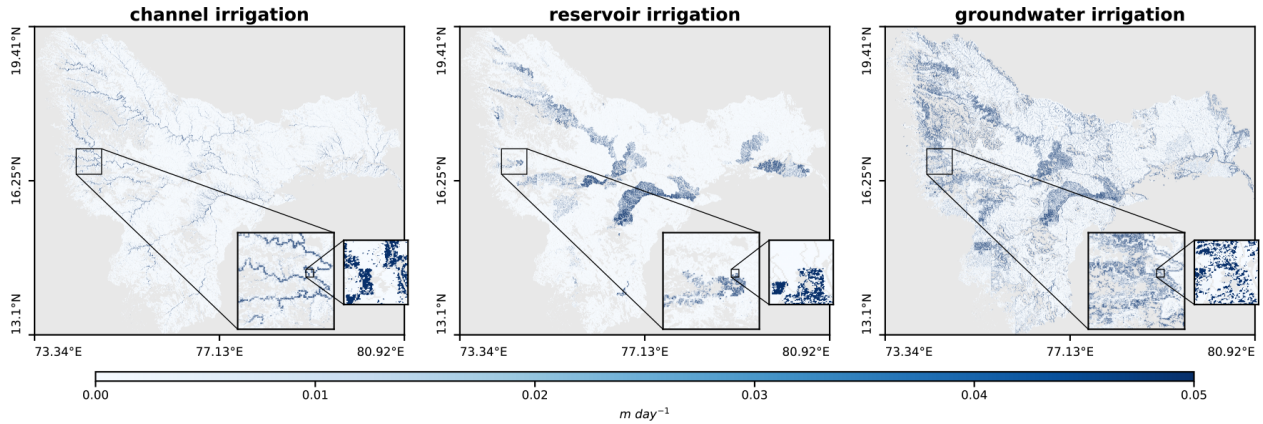
Figure 10. Average daily irrigation from channels, reservoirs, and groundwater.
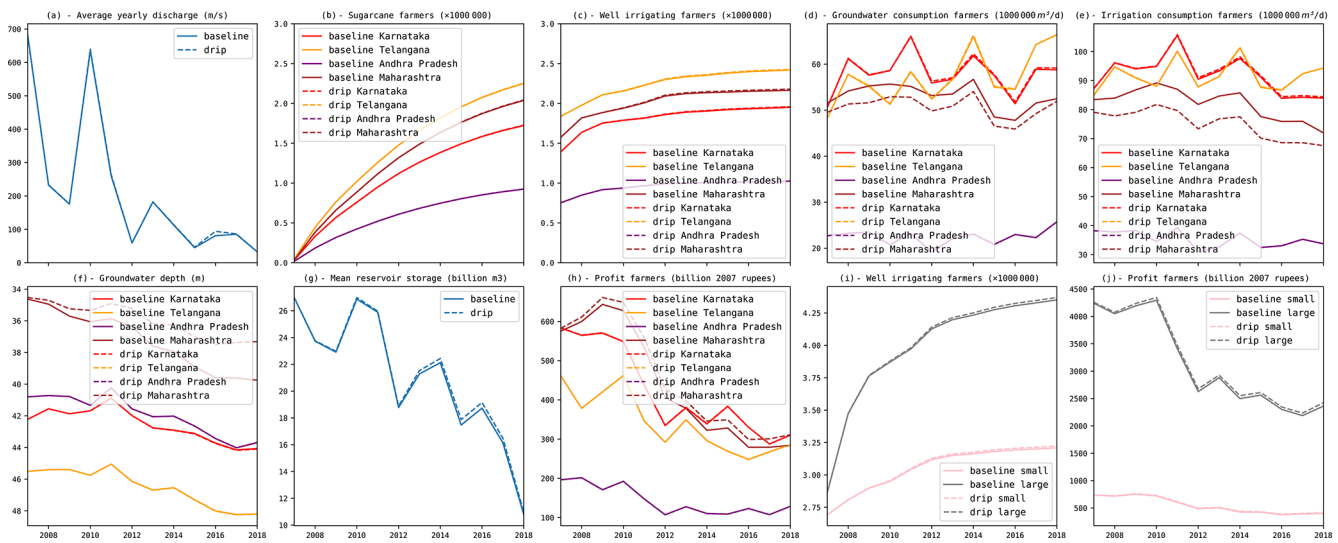

Figure 11. Yearly simulation results for baseline scenario and scenario with investment in drip irrigation.

with the hydrological system while assessing large-scale hydrological processes. By combining survey data and aggregate statistics at the subdistrict level, the model uses a statistically representative population, considering the heterogeneity of the population and the spatial variability. The open-source model is developed in Python, simulating more than 12 million farmers on an computer with 16 GB of RAM. By adopting a fully distributed hydrological model with dynamically sized field-scale hydrological response units (HRUs), the model can effectively represent subsistence farmers, ensuring that they can individually interact with the hydrological model while keeping computational demands reasonable.

Using the model, we quantitatively show how farmer behaviour and the hydrological system are intricately woven across both small and large scales. Changes in behaviour or investment in irrigation measures affect hydrology and other farmers locally, but they also affect river discharge and other farmers further downstream. Effects are visible both in hy-

drological variables as well as farmer behaviour and profit. Using a scenario where drip irrigation is promoted in an upstream state, we show how the effect of policies can be assessed on local and large-scale processes across both the hydrological and human domains. This provides opportunities to study large- and small-scale socio-hydrological processes simultaneously in large river basins worldwide. Moreover, capturing the full heterogeneity of the farmer populations allows for investigation of equity issues, particularly related to distributional justice.

The agent-based model can be separately coupled to other hydrological models, assuming the hydrological model facilitates simulation of field-scale hydrological processes. Alternatively, at the price of losing the ability to fully capture heterogeneity of human processes, the ABM can be adapted to simulate aggregated agents within a grid cell, facilitating coupling to other hydrological models that do not support field-scale hydrological simulation. Similarly, the adapted

version of CWatM, which now simulates field-scale hydrological processes, can be coupled to other ABMs.

Future studies could include additional adaptation measures to make the model more realistic and further investigate how policies and infrastructural projects, such as through reservoir construction and management, water/electricity pricing (Parween et al., 2021), water rights (Nouri et al., 2019a), and enforcing specific crop types (Wallach, 1984) can affect the human–natural system. As humans play a key role in the environment, the human component of GEB can also be central in coupling further models, e.g. to economic models or land use change models (Dou et al., 2020), allowing us to investigate the land–water–food use nexus. In addition, coupling GEB to a hydrodynamic model such as DIM (Farrag et al., 2021) would allow us to investigate the interactions between human behaviour and flood and drought risk (Ward et al., 2020). Finally, the integration of future scenarios such as climate change, population growth, and exogenous land use can be used to project how the coupled human–natural system changes in the future.

Computational requirements

The model for the entire Krishna basin can be run on a computer with 16 GB of RAM. Model run time was ∼ 10 s per daily time step (i.e. ∼ 1 h for 1 year) using a single core on an AMD EPYC 7302 while requiring no more than 7 GB of RAM and an 8 GB RTX1070 GPU. Without GPU, the run time was ∼ 30 s per time step while requiring 12 GB of RAM. Model run time and requirements scale near-linearly with basin size, assuming identical farm sizes. Larger farm sizes reduce the requirements, while smaller farms increase the requirements. In the implementation in the Krishna basin as shown here, the average farm size is 1.6 ha.

Appendix A

This section discusses the coupling between the ABM and hydrological model, serving as both an explanation to how the coupling has been performed in GEB and, if required, allowing the reader to couple their hydrological model to the ABM or the ABM to their hydrological model. The ABM can be found in the GEB repository (https://github.com/jensdebruijn/GEB, last access: 14 February 2023), while the adapted version of CWatM (see Sect. 2.1) can be found in the ABCWatM repository (https://github.com/jensdebruijn/ABCWatM, last access: 14 February 2023). Because both models are written in Python, the coupling is performed by creating a Python class, which *subclasses* both models while synchronizing their time steps. By subclassing both models, both models can access variables from the other model, and functions can be adapted, allowing the models to communicate during each time step.

The model class first loads the configuration file (default: "*GEB.yml*") which contains configuration parameters for both models, such as the start and end dates of the simulation. Then, the shared data class ("self.data") is loaded, which contains all data that are shared between the models. The data class loads the study area and automatically creates the grid and HRUs using the given land use and farm map (see Sect. 2). This class also contains convenience functions to convert data between the grid and HRUs (e.g. adding up runoff for individual HRUs to the grid level). An example is given in Fig. A1.

Then, the agent-based model is initialized with agent attributes (see Sect. 2). For farmers, this consists of a raster map that indicates the area that is managed by a specific farmer, the locations of the farms (e.g. the centre point of the each farmers field), and other attributes, such as crop type, cropping schedule, and irrigation status. Subsequently, CWatM is initialized as per Burek et al. (2020) while loading initial land use and crop parameters from the ABM.

After initialization, the (daily) time steps of the ABM and the hydrological model are synchronized. First, farmers make decisions based on the current conditions (e.g. planting, harvesting, investing), followed by the hydrological model which simulates hydrological processes, including water demand. For all sectors excluding crop farming, water demand is simulated as exogenous input. However, for crop water, demand and consumption are determined by the ABM. To this extent, CWatM calls a function (i.e. "abstract_water()") in the ABM that determines these demand and consumption rates using current soil moisture content; available water from surface, reservoir, and groundwater; and farmer characteristics such as irrigation equipment, crop type, and location.

Appendix B

The iterative proportional fitting (IPF) algorithm reweights the sample weights from survey data in such a way that the aggregate statistics of the reweighted survey data match those of the marginal distributions. To do so, it is required that before initialization of the algorithm the sum of initial weights matches the marginal distribution (e.g. when there are 500 women and 500 men in the marginal distribution, the sum of the initial weights must be 1000). However, here one of the statistics we use is crop type. In the synthetic population, we require crop type per agricultural season (i.e. kharif, rabi, and summer). In the IHDS, this is specified accordingly but not in the Indian Agricultural Census. Here, only the total area per crop is specified. If a farmer uses different crops per season, the area is counted twice. In the IPF procedure, it is thus required to combine the crops from the survey data and then fit the combined crops to the census data, which means that the total number of crops is higher than the total number of farmers (assuming at least one farmer is multi-cropping

```
# subgrid hydrological response units (colors indicate data belonging to the same grid cell.
land_use_type = [ 0, 1 ,  3, 3, 1  ,  3]  # 0: forest, 1: grassland, 3: irrigated cropland
land_owner = [ -1 ,  -1 ,  1, 2 ,  - 1,  2]  # -1: no owner, 1: owned by farmer #1, 2: owned by farmer #2.
land_use_ratio = [   .3 , .7  ,  1, .2, .7  ,  1.]  # HRUs of one grid cell always sum to 1
cell_boundaries = [0, 2, 5, 6]          # 1st  grid cell contains the 0    th  up to the 2 nd  HRU, the 2 nd  grid cell the 2    nd up to the 5  th, etc.

# set runoff from natural sources
runoff_m_HRU = [ 1, 2 ,  4, 1, 1  ,  0]

# farmer #2 has a groundwater pump running, creating additional runoff from irrigation return flow.
for i, land_owner in enumerate(land_owners):
    if land_owner == 2:   # increase runoff by 1 in fields owned by farmer #2
        runoff_m_HRU[i] += 1

runoff_m_HRU
>> [ 1, 2 ,  4, 2, 1  ,  1]

# translate subgrid to grid       – unit in meters in the example.
runoff_m_grid = []
for left, right in zip(cell_boundaries[:            - 1], cell_boundaries[1:]):
    runoff_m_grid.append(sum([
            runoff * land_size
            for runoff, land_size in zip(runoff_m_HRU[left:right], land_use_ratio[left:right])
    ]))

runoff_m_grid
>> [ 1.7 ,  1.5 ,  1]

# translate grid to subgrid        – unit in meters in the example.
precipitation_m_grid = [    1 ,  3 ,  2]
precipitation_m_HRU = []
for left, right, precipitation_HRU_m in zip(cell_boundaries[:             - 1], cell_boundaries[1:], precipitation_m_grid):
    number_of_HRUs_in_grid_cell = right     – left
    for i in range(number_of_HRUs_in_grid_cell):
        precipitation_m_HRU.append(precipitation_HRU_m)

precipitation_m_HRU
>> [ 1, 1 ,  3, 3, 3  ,  2]
```

Figure A1. A simplified code example of the conversion between hydrological response units (HRUs) and the grid.

in the region. To this extent, we adapt the IPF procedure in three ways to ensure correct fitting:

1. Entries in the survey data are reweighted based on the mean target for all crops within that entry. For example, to obtain the correct values at the tehsil level, sugar cane needs to be multiplied by 2.0, while rice needs to be multiplied by 0.8. In that case, the weight attributed to a farmer who grows both sugar cane and rice is multiplied by 1.6 (i.e. 2.0×0.8), while the weight attributed to another farmer who only grows rice is multiplied by 0.8.

2. A learning rate is introduced, meaning that the reweighting is not done at once, but changes are gradual following the learning rate. Here we set the learning rate to 0.1.

3. After reweighting all farmers, the entire survey is reweighted to ensure that the total number of farmers is still identical to the initial number of farmers.

Appendix C

Table C1 displays the crop growth stages, crop factors, and water-stress reduction factors for all crops included in the implementation in the Krishna basin.

Table C1. Crop growth stages, crop factors, and water-stress reduction factors.

Crop	d1	d2	d3	d4	Kc1	Kc2	Kc3	KyT	Y_r
Pearl millet	10	25	40	25	0.3	1	0.3	0.9	800
Groundnut	20	30	30	20	0.4	1.15	0.6	0.7	850
Sorghum	10	25	40	25	0.3	1.05	0.55	0.9	1500
Paddy	15	20	45	20	1.05	1.2	0.9	2	1200
Sugarcane	9	15	47	29	0.4	1.25	0.75	1.2	15 000
Wheat	15	25	40	20	0.3	1.1	0.3	1.15	1000
Cotton	15	25	50	10	0.35	1.15	0.7	0.85	550
Chickpea	20	30	30	20	0.4	1.05	0.35	1.15	600
Maize	15	30	40	15	0.3	1.2	0.35	1.25	2400
Green gram	20	30	30	20	0.4	1.05	0.35	1.15	600
Finger millet	10	25	40	25	0.3	1	0.3	0.9	800
Sunflower	15	30	40	15	0.35	1.05	0.35	0.95	400
Red gram	20	30	30	20	0.4	1.05	0.35	1.15	600

References

Ablo, A. D. and Yekple, E. E.: Urban water stress and poor sanitation in Ghana: perception and experiences of residents in the Ashaiman Municipality, GeoJournal, 83, 583–594, https://doi.org/10.1007/s10708-017-9787-6, 2018.

Aerts, J. C. J. H., Botzen, W. J., Clarke, K. C., Cutter, S. L., Hall, J. W., Merz, B., Michel-Kerjan, E., Mysiak, J., Surminski, S., and Kunreuther, H.: Integrating human behaviour dynamics into flood disaster risk assessment, Nat. Clim. Change, 8, 193–199, https://doi.org/10.1038/s41558-018-0085-1, 2018.

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

Arnold, R. T., Troost, C., and Berger, T.: Quantifying the economic importance of irrigation water reuse in a Chilean watershed using an integrated agent-based model, Water Resour. Res., 51, 648–668, https://doi.org/10.1002/2014WR015382, 2015.

Batchelor, C. H., Rama Mohan Rao, M. S., and Manohar Rao, S.: Watershed development: A solution to water shortages in semi-arid India or part of the problem?, Land Use and Water Resources Research, 3, 1–10, https://doi.org/10.22004/ag.econ.47866, 2003.

Becu, N., Perez, P., Walker, A., Barreteau, O., and Le Page, C.: Agent based simulation of a small catchment water management in northern Thailand Description of the CATCHSCAPE model, Ecol. Model., 170, 319–331, 2003.

Bert, F. E., Podestá, G. P., Rovere, S. L., Menéndez, Á. N., North, M., Tatara, E., Laciana, C. E., Weber, E., and Toranzo, F. R.: An agent based model to simulate structural and land use changes in agricultural systems of the argentine pampas, Ecol. Model., 222, 3486–3499, https://doi.org/10.1016/j.ecolmodel.2011.08.007, 2011.

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

Biggs, T., Gaur, A., Scott, C., Thenkabail, P., Gangadhara Rao, P., Gumma, M. K., Acharya, S., and Turral, H.: Closing of the Krishna Basin: Irrigation, streamflow depletion and macroscale hydrology, Volume 111 of IWMI Research Report, ISBN 9290906634, 9789290906636, 2007.

Bouma, J. A., Biggs, T. W., and Bouwer, L. M.: The downstream externalities of harvesting rainwater in semi-arid watersheds: An Indian case study, Agr. Water Manage., 98, 1162–1170, https://doi.org/10.1016/j.agwat.2011.02.010, 2011.

Brouwer, C. and Heibloem, M.: Irrigation water management: irrigation water needs, Training manual 3, 1–5, 1986.

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, 2020.

Chaney, N. W., Metcalfe, P., and Wood, E. F.: HydroBlocks: a field-scale resolving land surface model for application over continental extents, Hydrol. Process., 30, 3543–3559, https://doi.org/10.1002/HYP.10891, 2016.

Chaney, N. W., Torres-Rojas, L., Vergopolan, N., and Fisher, C. K.: HydroBlocks v0.2: enabling a field-scale two-way coupling between the land surface and river networks in Earth system models, Geosci. Model Dev., 14, 6813–6832, https://doi.org/10.5194/gmd-14-6813-2021, 2021.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE T. Evol. Comput., 6, 182–197, https://doi.org/10.1109/4235.996017, 2002.

de Bruijn, J.: jensdebruijn/GEB: v0.1.1 (v0.1.1), Zenodo [code], https://doi.org/10.5281/zenodo.7820962, 2023a.

de Bruijn, J.: jensdebruijn/ABCWatM: v0.1.1 (v0.1.1), Zenodo [code], https://doi.org/10.5281/zenodo.7820968, 2023b.

de Bruijn, J.: VU-IVM/honeybees: v0.2 (0.2), Zenodo [code], https://doi.org/10.5281/zenodo.7820973, 2023c.

Desai, S., Vanneman, R., and National Council of Applied Economic Research New Delhi: India Human Development Survey (IHDS), Inter-university Consortium for Political and Social Research, https://doi.org/10.3886/ICPSR22626.v12, 2005.

Di Baldassarre, G., Cloke, H., Lindersson, S., Mazzoleni, M., Mondino, E., Mård, J., Odongo, V., Raffetti, E., Ridolfi, E., Rusca, M., Savelli, E. and Tootoonchi, F.: Integrating Multiple Research Methods to Unravel the Complexity

of Human-Water Systems, AGU Adv., 2, e2021AV000473, https://doi.org/10.1029/2021AV000473, 2021.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., and Hanasaki, N.: GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface, B. Am. Meteorol. Soc., 87, 1381–1398, https://doi.org/10.1175/BAMS-87-10-1381, 2006.

Dou, Y., Yao, G., Herzberger, A., da Silva, R. F. B., Song, Q., Hovis, C., Batistella, M., Moran, E., Wu, W., and Liu, J.: Land-Use Changes in Distant Places: Implementation of a Telecoupled Agent-Based Model, JASSS, 23, 11, https://doi.org/10.18564/jasss.4211, 2020.

Eilander, D., Winsemius, H. C., Van Verseveld, W., Yamazaki, D., Weerts, A., and Ward, P. J.: MERIT Hydro IHU, Zenodo [data set], https://doi.org/10.5281/zenodo.5166932, 2020.

Eilander, D., van Verseveld, W., Yamazaki, D., Weerts, A., Winsemius, H. C., and Ward, P. J.: A hydrography upscaling method for scale-invariant parametrization of distributed hydrological models, Hydrol. Earth Syst. Sci., 25, 5287–5313, https://doi.org/10.5194/hess-25-5287-2021, 2021.

Farrag, M., Vorogushyn, S., Nguyen, D. V., de Bruijn, K., and Merz, B.: River-dike-floodplain system interactions and temporal dynamics for large-scale flood risk assessment, FLOODrisk 2020-4th European Conference on Flood Risk Management, https://doi.org/10.3311/FloodRisk2020.9.11, 2021.

Fischer, G., Nachtergaele, F. O., Van Velthuizen, H. T., Chiozza, F., Franceschini, G., Henry, M., Muchoney, D., and Tramberend, S.: Global Agro-Ecological Zones v4–Model documentation, Food & Agriculture Org., ISBN 978-92-5-134426-2, 2021.

Flügel, W. A.: Combining GIS with regional hydrological modelling using hydrological response units (HRUs): An application from Germany, Math. Comput. Simulat., 43, 297–304, https://doi.org/10.1016/S0378-4754(97)00013-X, 1997.

Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A., Parizeau, M., and Gagné, C.: DEAP: Evolutionary Algorithms Made Easy, J. Mach. Learn. Res., 13, 2171–2175, 2012.

Gumma, M. K., Tummala, K., Dixit, S., Collivignarelli, F., Holecz, F., Kolli, R. N., and Whitbread, A. M.: Crop type identification and spatial mapping using Sentinel-2 satellite data with focus on field-level information, Geocarto Int., 37, 1833–1849, https://doi.org/10.1080/10106049.2020.1805029, 2020.

Gupta, J. and van der Zaag, P.: Interbasin water transfers and integrated water resources management: Where engineering, science and politics interlock, Phys. Chem. Earth, 33, 28–40, https://doi.org/10.1016/j.pce.2007.04.003, 2008.

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., 22, 789–817, https://doi.org/10.5194/hess-22-789-2018, 2018.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357–362, https://doi.org/10.1038/s41586-020-2649-2, 2020.

Huber, L., Bahro, N., Leitinger, G., Tappeiner, U., and Strasser, U.: Agent-based modelling of a coupled water demand and supply system at the catchment scale, Sustain., 11, 6178, https://doi.org/10.3390/su11216178, 2019.

Ibisch, R. B., Bogardi, J. J., and Borchardt, D.: Integrated Water Resources Management: Concept, Research and Implementation, in: Integrated Water Resources Management: Concept, Research and Implementation, edited by: Borchardt, D., Bogardi, J. J., and Ibisch, R. B., Springer International Publishing, Cham, 3–32, ISBN 978-3-319-25071-7, 2016.

Jun, C., Ban, Y., and Li, S.: Open access to Earth land-cover map. Nature 514, https://doi.org/10.1038/514434c, 2014.

Kazil, J., Masad, D., and Crooks, A.: Utilizing Python for Agent-Based Modeling: The Mesa Framework, in: Social, Cultural, and Behavioral Modeling, edited by: Thomson, R., Bigsin, H., Dancy, C., Hyder, A., and Hussain, M., Springer International Publishing, 308–317, 2020.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, J. Hydrol., 424–425, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

Kuil, L., Evans, T., McCord, P. F., Salinas, J. L., and Blöschl, G.: Exploring the Influence of Small-holders "Smallholders" Perceptions Regarding Water Availability on Crop Choice and Water Allocation Through Socio-Hydrological Modeling, Water Resour. Res., 54, 2580–2604, https://doi.org/10.1002/2017WR021420, 2018.

Kummu, M., Guillaume, J. H. A., de Moel, H., Eisner, S., Flörke, M., Porkka, M., Siebert, S., Veldkamp, T. I. E., and Ward, P. J.: The world's road to water scarcity: shortage and stress in the 20th century and pathways towards sustainability, Sci. Rep., 6, 38495, https://doi.org/10.1038/srep38495, 2016.

Lam, S. K., Pitrou, A., and Seibert, S.: Numba: A llvm-based python jit compiler, in Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, 1–6, https://doi.org/10.1145/2833157.2833162, 2015.

Langevin, C. D., et al.: Documentation for the MODFLOW 6 groundwater flow model, No. 6-A55, US Geological Survey, 2017.

Li, F., Cook, S., Geballe, G. T., and Burch Jr, W. R.: Rainwater Harvesting Agriculture: An Integrated System for Water Management on Rainfed Land in China's Semiarid Areas, Ambio, 29, 477–483, https://doi.org/10.1579/0044-7447-29.8.477, 2000.

Llamas, M. R. and Martínez-Santos, P.: Intensive groundwater use: silent revolution and potential source of social conflicts, J. Water Res. Pl., 131, 337–341, https://doi.org/0.1061/(ASCE)0733-9496(2005)131:5(337), 2005.

Lowder, S. K., Skoet, J., and Raney, T.: The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide, World Development, 87, 16–29, https://doi.org/10.1016/j.worlddev.2015.10.041, 2016.

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, 13603, https://doi.org/10.1038/ncomms13603, 2016.

Mollinga, P. P.: On the waterfront: Water distribution, technology and agrarian change in a South Indian canal irrigation system, Orient Blackswan, ISBN 905485927X, 2003.

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., Peiris, T. A., Popat, E., Portmann, F. T., Reinecke, R., Schumacher, M., Shadkam, S., Telteu, C.-E., Trautmann, T., and Döll, P.: The global water resources and use model WaterGAP v2.2d: model description and evaluation, Geosci. Model Dev., 14, 1037–1079, https://doi.org/10.5194/gmd-14-1037-2021, 2021.

Nouri, A., Saghafian, B., Delavar, M., and Bazargan-Lari, M. R.: Agent-Based Modeling for Evaluation of Crop Pattern and Water Management Policies, Water Resour. Manag., 33, 3707–3720, https://doi.org/10.1007/s11269-019-02327-3, 2019a.

Nouri, H., Stokvis, B., Galindo, A., Blatchford, M., and Hoekstra, A. Y.: Water scarcity alleviation through water footprint reduction in agriculture: The effect of soil mulching and drip irrigation, Sci. Total Environ., 653, 241–252, https://doi.org/10.1016/j.scitotenv.2018.10.311, 2019b.

Parween, F., Kumari, P., and Singh, A.: Irrigation water pricing policies and water resources management, Water Policy, 23, 130–141, 2021.

Porporato, A., Laio, F., Ridolfi, L., and Rodriguez-Iturbe, I.: Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress: III. Vegetation water stress, Adv. Water Resour., 24, 725–744, https://doi.org/10.1016/S0309-1708(01)00006-9, 2001.

Schulla, J. and Jasper, K.: Model description wasim-eth, Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology, Zürich, 2007.

Schreinemachers, P. and Berger, T.: An agent-based simulation model of human-environment interactions in agricultural systems, Environ. Modell. Softw., 26, 845–859, https://doi.org/10.1016/j.envsoft.2011.02.004, 2011.

Schrieks, T., Botzen, W. J. W., Wens, M., Haer, T., and Aerts, J. C. J. H.: Integrating Behavioral Theories in Agent-Based Models for Agricultural Drought Risk Assessments, Front. Water, 3, 104, https://doi.org/10.3389/frwa.2021.686329, 2021.

Shah, T. and Bhattacharya, S.: Farmer Organizations for Lift Irrigation: Irrigation Companies and Tubewell Cooperatives of Gujarat, http://hdl.handle.net/10535/4600 (last access: 2 May 2023), 1993.

Sharma, B. R., Rao, K., and Massuel, S.: Groundwater externalities of surface irrigation transfers under National River Linking Project: Polavaram–Vijayawada link. Strategic Analyses of the National River Linking Project (NRLP) of India Series 2, 271, 2008.

Siebert, S. and Döll, P.: Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation, J. Hydrol., 384, 198–217, https://doi.org/10.1016/j.jhydrol.2009.07.031, 2010.

Streefkerk, I. N., de Bruijn, J., Haer, T., Van Loon, A. F., Quichimbo, E. A., Wens, M., Hassaballah, K., and Aerts, J. C. J. H.: A coupled agent-based model to analyse human-drought feedbacks for agropastoralists in dryland regions, Front. Water, 4, https://doi.org/10.3389/frwa.2022.1037971, 2023.

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., 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 5 arcmin global hydrological and water resources model, Geosci. Model Dev., 11, 2429–2453, https://doi.org/10.5194/gmd-11-2429-2018, 2018.

Tamburino, L., Di Baldassarre, G., and Vico, G.: Water management for irrigation, crop yield and social attitudes: a socio-agricultural agent-based model to explore a collective action problem, Hydrolog. Sci. J., 65, 1815–1829, https://doi.org/10.1080/02626667.2020.1769103, 2020.

Tripathi, A. and Mishra, A. K.: Knowledge and passive adaptation to climate change: An example from Indian farmers, Climate Risk Manage., 16, 195–207, https://doi.org/10.1016/j.crm.2016.11.002, 2017.

van Leeuwen, C. J., Dan, N. P. and Dieperink, C.: The challenges of water governance in Ho Chi Minh City, Integr. Environ. Assess., 12, 345–352, https://doi.org/10.1002/ieam.1664, 2016.

van Oel, P. R., Krol, M. S., Hoekstra, A. Y., and Taddei, R. R.: Feedback mechanisms between water availability and water use in a semi-arid river basin: A spatially explicit multi-agent simulation approach, Environ. Modell. Softw., 25, 433–443, https://doi.org/10.1016/j.envsoft.2009.10.018, 2010.

Veldkamp, T. I. E., Wada, Y., de Moel, H., Kummu, M., Eisner, S., Aerts, J. C. J. H., and Ward, P. J.: Changing mechanism of global water scarcity events: Impacts of socioeconomic changes and inter-annual hydro-climatic variability, Global Environ. Chang., 32, 18–29, https://doi.org/10.1016/j.gloenvcha.2015.02.011, 2015.

Veldkamp, T. I. E., Wada, Y., Aerts, J. C. J. H., Döll, P., Gosling, S. N., Liu, J., Masaki, Y., Oki, T., Ostberg, S., Pokhrel, Y., Satoh, Y., Kim, H., and Ward, P. J.: Water scarcity hotspots travel downstream due to human interventions in the 20th and 21st century, Nat. Commun., 8, 15697, https://doi.org/10.1038/ncomms15697, 2017.

Wada, Y., van Beek, L. P. H., and Bierkens, M. F. P.: Modelling global water stress of the recent past: on the relative importance of trends in water demand and climate variability, Hydrol. Earth Syst. Sci., 15, 3785–3808, https://doi.org/10.5194/hess-15-3785-2011, 2011.

Waldner, F. and Diakogiannis, F. I.: Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network, Remote Sens. Environ., 245, 111741, https://doi.org/10.1016/j.rse.2020.111741, 2020.

Wallach, B.: Irrigation Developments in the Krishna Basin since 1947, Geogr. Rev., 74, 127–144, https://doi.org/10.2307/214095, 1984.

Ward, P. J., de Ruiter, M. C., Mård, J., Schröter, K., Van Loon, A., Veldkamp, T., von Uexkull, N., Wanders, N., AghaKouchak, A., Arnbjerg-Nielsen, K., Capewell, L., Carmen Llasat, M., Day, R., Dewals, B., Di Baldassarre, G., Huning, L. S., Kreibich, H., Mazzoleni, M., Savelli, E., Teutschbein, C., van den Berg, H., van der Heijden, A., Vincken, J. M. R., Waterloo, M. J., and Wens, M.: The need to integrate flood and drought disaster risk reduction strategies, Water Secur., 11, 100070, https://doi.org/10.1016/j.wasec.2020.100070, 2020.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., and Schewe, J.: The inter-sectoral impact model intercomparison project (ISI–MIP): project framework, P. Natl. Acad. Sci. USA, 111, 3228–3232, 2014.

Wens, M., Veldkamp, T. I. E., Mwangi, M., Johnson, J. M., Lasage, R., Haer, T., and Aerts, J. C. J. H.: Simulating Small-Scale Agricultural Adaptation Decisions in Response to Drought Risk: An Empirical Agent-Based Model for Semi-Arid Kenya, Front. Water, 2, 15, https://doi.org/10.3389/frwa.2020.00015, 2020.

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, 2019.