Understanding each other’s models: a standard representation of 16 global water models to support intercomparison, improvement, and communication

Camelia-Eliza Telteu1, Hannes Müller Schmied1,2, Wim Thiery3, Guoyong Leng4, Peter Burek5, Xingcai Liu4, Julien Eric Stanislas Boulange6, Lauren Seaby Andersen7, Manolis Grillakis8, Simon Newland Gosling9, Yusuke Satoh6, Oldrich Rakovec10,11, Tobias Stacke12, Jinfeng Chang13,14, Niko Wanders15, Harsh Lovekumar Shah16, Tim Trautmann1,10, Naota Hanasaki6, Aristeidis Koutoulis18, Yadu Pokhrel19, Luis Samaniego10, Yoshihide Wada5, Vimal Mishra16, Junguo Liu17, Petra Döll1,2, Fang Zhao20, Anne Gädeke7, Sam Rabin21, Florian Herz1

1Institute of Physical Geography, Johann Wolfgang Goethe University Frankfurt, Frankfurt am Main, 60438, Germany
2Senckenberg Leibniz Biodiversity and Climate Research Centre (SBiK-F), Frankfurt am Main, 60325, Germany
3Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels, 1050, Belgium
4Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China
5International Institute for Applied Systems Analysis, Laxenburg, 2361, Austria
6National Institute for Environmental Studies, Tsukuba, 305–8506, Japan
7Potsdam Institute for Climate Impact Research, Potsdam, 14473, Germany
8Institute for Mediterranean Studies, Foundation for Research and Technology-Hellas, Rethymno, 74100, Greece
9School of Geography, University of Nottingham, Nottingham, NG7 2RD, United Kingdom of Great Britain and Northern Ireland
10Department Computational Hydrosystems, UFZ-Helmholtz Centre for Environmental Research, Leipzig, 04318, Germany
11Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Prague, 16500, Czech Republic
12Institute of Coastal Research, Helmholtz-Zentrum Geesthacht, Geesthacht, 21502, Germany
13College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, 310058, China
14Laboratoire des Sciences du Climat et de l’Environnement, CEA-CNRS-UVSQ/IPSL, Université Paris Saclay, Gif sur Yvette, 91191, France
15Department of Physical Geography, Utrecht University, Utrecht, 3508, The Netherlands
16Indian Institute of Technology Gandhinagar, Palaj, Gandhinagar, 382355, India
17School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, 518055, China
18School of Environmental Engineering, Technical University of Crete, Chania, 73100, Greece
19Department of Civil and Environmental Engineering, Michigan State University, East Lansing, Michigan, 48824, United States of America
20School of Geographic Sciences, East China Normal University, Shanghai, 200241, China
21Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research / Atmospheric Environmental Research, Garmisch-Partenkirchen, 82467, Germany

Correspondence to: Camelia-Eliza Telteu (telteu@em.uni-frankfurt.de, camelia.telteu@gmail.com)

Abstract. Global water models (GWMs) simulate the terrestrial water cycle, on the global scale, and are used to assess the impacts of climate change on freshwater systems. GWMs are developed within different modelling frameworks and consider different underlying hydrological processes, leading to varied model structures. Furthermore, the equations used to describe various processes take different forms and are generally accessible only from within the individual model codes. These factors
have hindered a holistic and detailed understanding of how different models operate, yet such an understanding is crucial for explaining the results of model evaluation studies, understanding inter-model differences in their simulations, and identifying areas for future model development. This study provides a comprehensive overview of how 16 state-of-the-art GWMs are designed. We analyze water storage compartments, water flows, and human water use sectors included in 16 GWMs that provide simulations for the Inter-Sectoral Impact Model Intercomparison Project phase 2b (ISIMIP2b). We develop a standard writing style for the model equations to further enhance model intercomparison, improvement, and communication. In this study, WaterGAP2 used the highest number of water storage compartments, 11, and CWatM used 10 compartments. Six models used six compartments, while four models (DBH, JULES-W1, Mac-PDM.20, and VIC) used the lowest number, three compartments. WaterGAP2 simulates five human water use sectors, while four models (CLM4.5, CLM5.0, LPJmL, and MPI-HM) simulate only water for the irrigation sector. We conclude that even though hydrological processes are often based on similar equations, in the end, these equations have been adjusted or have used different values for specific parameters or specific variables. Ultimately, we consider that similarities and differences found among the models analyzed in this study enable us to reduce the uncertainty of multi-model ensembles, to improve the existing hydrological processes, and to integrate new processes.

1 Introduction

Many multi-model intercomparison projects (MIPs) have been designed to provide insights into various Earth system processes. They provide many multi-model ensembles that consist of multiple models driven by the output of multiple other models. These multi-model ensembles offer the opportunity to inter-compare models for an increased understanding of process representation and inter-model differences as well as for model improvement. Some MIPs examples include FireMIP for the fire regime and its drivers (Rabin et al., 2017); CMIP for past, present, and future climate changes and their drivers (Eyring et al., 2016; Kageyama et al., 2018); LakeMIP for physical and biogeochemical processes of lakes (Stepanenko et al., 2010; Thiery et al., 2014); AgMIP for crop growth (Rosenzweig et al., 2013), and WaterMIP or ISIMIP for the water cycle (Haddeland et al., 2011; Frieler et al., 2017).

MIPs have encountered many challenges in how to inter-compare models and interpret various model results (von Lampe et al., 2013), realize the standardization of data and scenarios and integrate transdisciplinary knowledge in modeling (Rosenzweig et al., 2013), identify and reduce uncertainties (Sitch et al., 2008). They have been affected by scientific complexity, input data quality, technical infrastructure, and even cultural and organizational challenges (Eyring et al., 2016). Hence, they have evaluated models’ performance in the past and have focused on the models’ agreement for the future. Ultimately, MIPs and their multi-model ensembles have been blocked in interpreting inter-model differences because of models’ complexity, missing information about others models, incomplete or missing information about heterogeneity and dynamism of natural systems (Clark et al., 2011).

Nevertheless, MIPs have underlined the need to go beyond good overall model performance and to improve process representation in the models (Guseva et al., 2020), integrate missing processes (Friend et al., 2014), and reduce uncertainties (Warszawski et al. 2014). MIPs showed that robust similarities exist among models and, as a result, models are not strictly independent of each other given previous and legacy versions, and existing links among modeling communities who indirectly transfer some models’ strengths and weaknesses by sharing their ideas and codes (Masson and Knutti, 2011; Knutti et al.,
It has been concluded that there is no perfect model (Essery et al., 2013; Ullrich et al., 2017) and there is a need to understand better how models work. Thereby, the modeling communities are still testing and learning how to improve modeling and how to realize multi-model inter-comparison studies. However, few studies have undertaken model experiments on process representation and evaluated the models for specific events or characteristics specifically on the catchment scale (de Boer-Euser et al., 2017; Duethmann et al., 2020; Bouaziz et al., 2021).

In this complex scientific context, the present study represents a step forward to increase understanding of process representation and inter-model differences within one large MIP, specifically, ISMIP – the Inter-Sectoral Impact Model Intercomparison Project (Frieler et al., 2017). We assessed the equations applied by 16 state-of-the-art global water models (GWMs) to simulate the vertical and lateral water balance, and human water use sectors, on the global scale. We created a standard writing style of these equations to identify similarities and differences among models. Thereby, the global water community has through this study an overview of the model structures and the basis required to interpret various model results, to design future experiments on how model equations, model configurations, and model parameter values influence the model outputs.

In summary, our three main goals are:

- to provide a better understanding of how 16 state-of-the-art global water models are designed;
- to show similarities and differences among them, based on their equations;
- to underline future research potential in global water modeling.

Therefore, this study supports intercomparison, improvement, and communication among 16 modelling teams. It also provides the basis for (i) further water model (inter)comparison studies, including model outputs; (ii) selecting the right model(s) for a given application; and (iii) identifying data needs for a given analysis and application.

The target audience includes students, junior and senior scientists, modellers, interested stakeholders, and members of the general public interested in understanding global water models and modelling the impact of climate change on freshwater systems.

We present the modeling approaches and terminology used in global water modelling in section 2. In section 3, we present key characteristics of the models analyzed in the present study. In section 4, we describe our standard writing style of model equations. In section 5, we present similarities and differences among models. In section 6, we present the number of water flows, water storage compartments, and human water use sectors included in the 16 GWMs. In section 7, we discuss potential future research in global water modeling. In section 8, we present recommendations for future multi-model intercomparison projects and extended assessments.
2. Modeling approaches and terminology used in global water modelling

2.1 Differences in modeling approaches

On global scale, the terrestrial water cycle is simulated by three different communities that have developed three types of models: (i) the climate community has developed land surface models (LSMs); (ii) the global hydrological community has developed global hydrological models (GHMs); (iii) the vegetation community has developed dynamic global vegetation models (DGVMs). However, these three communities focus on specific hydrological and atmospheric processes, as well as anthropogenic impacts. These key aspects are important for their specific research leading to different modelling approaches, specific evaluation studies of model performance (Archfield et al., 2015), and different field-specific meanings of terminology used (Beven and Young, 2013). Thus, combining the expertise in their key aspects would create a strong synergy and improve the models of these communities, but for this goal, they have to interact with each other, identify their similarities and differences and share experiences. They need to undertake joint experiments, present and discuss their results, discuss how they influence and depend on each other, and how water modeling can be improved (Cucchi et al., 2020).

The global hydrological community focuses primarily on surface water and groundwater availability, its human interference, and their daily to century-scale changes. GHMs simulate the water cycle with its water flows, water compartments, and human water use sectors. These models simulate water abstracted for the irrigation, domestic use, livestock, industry (manufacturing and electricity), and desalination sectors. Furthermore, reservoir management and its streamflow alteration are included. One of their main foci is streamflow simulation and their ability to reproduce historical observations of this variable. They focus on lateral and vertical flows, comprehensively simulating the following surface water bodies: (i) lakes, (ii) wetlands, (iii) rivers.

The climate community focuses on simulating climate and its change over decades and centuries using global climate models (GCMs) and Earth system models (ESMs). A fundamental component of these are the LSMs, which simulate the water and energy exchanges between the land surface and the atmosphere, specifically focusing on vertical flow exchanges. Therefore, these models simulate the energy cycle, the water cycle, the carbon and nitrogen cycles, and vegetation and crop responses to temperature, precipitation, and CO₂ concentrations. Further, they represent the soil with a higher vertical resolution and evapotranspiration and snow dynamics in a more physical manner than the global hydrological models (GHMs; Döll et al., 2016; Pokhrel et al., 2016; Wada et al., 2017).

The vegetation community focuses on vegetation distribution and growth in an area and over a time interval and is primarily interested in the global carbon cycle. DGVMs simulate shifting vegetation, driven by biogeochemistry, hydrology, and anthropogenic influences. These models simulate the vegetation composition and distribution as well as compartments and flows of carbon and water, for both natural and agricultural ecosystems. Specifically, they model the active response of vegetation to changes in air temperature, precipitation, and CO₂ concentrations.

The different viewpoints of these communities are readily visible in very basic concepts such as the solar energy. This is the main driver that connects the processes of terrestrial water cycle simulated by these communities. It specifically links the water and energy budgets with vegetation processes. This link can be exemplified by the latent heat flux of evaporation that describes...
the heat or the energy required to change the liquid water into water vapor. This heat or energy is locked in the humid air as water vapor, and is released when the humid air touches cold air and water vapor condensation starts. Therefore, continental evaporation is considered to be water loss by the global hydrological and vegetation modeling communities, but a water source (for cloud formation) by the climate community (those that simulate the atmosphere), with implications for agriculture and ecosystems (Abbott et al., 2019). Additionally, transpiration represents a water source for the vegetation community, necessary for photosynthesis and plant growth, and water loss for the global hydrological community.

In the end, these three communities have developed three types of models to simulate the terrestrial water cycle on the global scale despite fundamental differences in model structure, model equations, and output variables. Hence, we decided to include the three types of models in one group and call them global water models (GWMs).

2.2 Definitions used in global water modelling

A global water model describes the dynamic behavior of a hydrological system that includes input variables, state variables, parameters, constants, and output variables (Bierkens and van Geer, 2007). State variables define how much water is in a compartment or storage at the beginning of the simulation, and can change in space and time, for example, canopy water storage. Their variation is caused by a variation of the input variables, for example, precipitation. State variables are related to the input variables and output variables through parameters, for example, infiltration capacity of the soil. Parameters and coefficients represent numbers that describe a particular characteristic of reality, of the model, of the catchment area or flow domain. Some examples are runoff coefficient, soil porosity, hydraulic conductivity of different soil horizons, maximum soil water storage, maximum canopy water storage, mean residence time in the saturated zone, surface roughness, and vegetation properties (Beven, 2012). A model also uses physical and mathematical constants meaning characteristics of the model that do not change in space and time such as catchment area. Physical constants are physical quantities that can be measured and have a constant value in time, for example, the density of water at 0°C, the density of ice. Mathematical constants cannot be measured, but can be calculated and have a fixed numerical value, for example, \( e = 2.718 \ldots \), \( \pi = 3.142 \), \( i^2 = -1 \). Ultimately, output variables vary in space and time, for example, streamflow in a river catchment.

Thus, a water global model includes many equations written with a programming language in a model code to simulate freshwater systems. During simulations, many parameters receive specific values because they cannot be measured everywhere, therefore, they are calibrated or tuned to attain the best match between simulated and observed data. The final steps of a simulation are to validate simulated and observed data, to find out how well they fit, and to evaluate the simulated results through analysis and visualization.

3 Key characteristics of 16 global water models included in the study

In this study, we analyze 16 state-of-the-art global water models included in the global water sector of the Inter-Sectoral Impact Model Intercomparison Project phase 2b (ISIMIP2b: Frieler et al., 2017). GWMs include six land surface models (LSMs),
nine global hydrologic models (GHMs), and one dynamic global vegetation model (DGVM: LPJmL, Tables 6 and 12). Land surface models are CLM4.5, CLM5.0, DBH, JULES-W1, MATSIRO, and ORCHIDEE. Global hydrologic models are CWatM, H08, Mac-PD20, mHM, MPI-HM, PCR-GLOBWB, VIC, WaterGAP2, and WAYS. Generally, these models are suitable for application over a catchment size of not smaller than 9,000 km² or at least four grid cells (Döll et al., 2003; Hunger and Döll, 2008). For smaller catchments, the results are often not reasonable (e.g., Beck et al., 2016) and require corrections due to inaccurate input data, spatial heterogeneity, and the missing representation of some hydrological processes (Döll et al., 2003; Hunger and Döll, 2008).

### 3.1 General setup

These models contribute to an experiment setup designed to assess the impact of historical and future warming under the Paris Agreement (Frieler et al., 2017). They are driven by the same climate input datasets under representative concentration pathways (RCPs) and socioeconomic scenarios (SSPs). The time span of the simulations is divided into pre-industrial (1661–1860), historical (1861–2005) and future (RCP2.6, RCP6.0, and RCP8.5, 2006–2099 (2299)). These models simulate the terrestrial water cycle, on the global land area (except Antarctica) with a spatial resolution of $0.5\degree \times 0.5\degree$ (~55 km × 55 km at the Equator), and quantify water flows, water storage compartments, and human water use under the given climatic and socioeconomic conditions. They do not simulate the ocean component of the global water cycle or water quality. Some of these models also consider reservoir operations.

### 3.2 Temporal and spatial characteristics

Twelve models have a daily temporal resolution (Table 6), while MATSIRO has an hourly temporal resolution. Four models (CLM4.5, CLM5.0, MATSIRO, and ORCHIDEE) have 30-min temporal resolution, while JULES-W1 has a 1-hour resolution. Fifteen models run with a spatial resolution of $0.5\degree$. ORCHIDEE runs with a spatial resolution of $1.0\degree$ and has its outputs converted to $0.5\degree$ spatial resolution. Some models include subgrids for some components: CLM4.5 and CLM5.0 for vegetation, surface runoff and evapotranspiration; H08 and CWatM for land cover; MPI-HM for surface runoff and evapotranspiration; PCR-GLOBWB for vegetation and land cover; WaterGAP2, CWatM and MATSIRO for snow; VIC for vegetation and elevation. Furthermore, MATSIRO divides a subgrid cell in snow-covered and snow-free portions with flows and storages resolved separately for these portions, both for land and canopy surfaces.

### 3.3 River networks used

Nine models (CLM4.5, CLM5.0, CWatM, H08, LPJmL, MATSIRO, MPI-HM, PCR-GLOBWB, WaterGAP2) use the 30-min global drainage direction map DDM30 (Döll and Lehner, 2002), a raster map with a spatial resolution of $0.5\degree \times 0.5\degree$ (~ 50 km × 50 km), to outline the drainage directions of surface water collected by creeks, rivulets, and rivers. In this map, 67,420 discrete grid cells are characterized by their specific drainage direction and are organized into drainage basins that drain from the Earth’s land surface into the ocean or inland sinks. The mHM uses a river network ($0.5\degree \times 0.5\degree$) upscaled from
HydroSHEDS (Lehner et al., 2006). ORCHIDEE uses the river network from the Simulated Topological Networks (STN-30p: Vörösmarty et al., 2000). Five models (DBH, JULES-W1, Mac-PDM.20, VIC, and WAYS) do not use any river routing scheme for the ISIMIP2b because they do not compute streamflow.

3.4 Calibration approaches

Six GHMs perform calibration of their hydrological components, using different approaches (Table 6). CWatM calibrates monthly or daily streamflow for 12 catchments using the Distributed Evolutionary Algorithms in Python (DEAP) approach (Burek et al., 2020), while WaterGAP2 uses a basin-specific approach to match long-term mean annual observed streamflow at the outlet of 1,319 river basins. It considers runoff as a nonlinear function of soil moisture and uses a runoff coefficient plus up to two additional factors for calibration (Müller Schmied et al., 2014; Müller Schmied et al., 2021). Mac-PDM.20 is calibrated using the generalized likelihood uncertainty estimation (GLUE) approach (Smith, 2016). In mHM, calibration of global model parameters is performed against the daily observed streamflow, along with gridded global fields of FLUXNET evaporation (Jung et al., 2011) and a GRACE terrestrial water storage anomaly, using the ERA5 climate forcing (Landerer and Swenson, 2012). VIC uses the source datasets and parameter sets from Nijssen et al. (2001), namely the AVHRR-derived landcover dataset (Hansen et al., 2000) and the FAO soil textures (FAO, 1995), and is sub-sampled to 0.5° × 0.5° via a nearest-neighbor approach. WAYS is calibrated against data from the International Satellite Land Surface Climatology Project (ISLSCP) Initiative II of the University of New Hampshire and GRDC composite monthly runoff data (Fekete et al., 2011), from 1986 to 1995 at a 0.5° spatial resolution. CLM5.0 performs hydrological calibration in a Bayesian framework using a sequential Monte Carlo method (Lawrence et al., 2019). Five models (CLM4.5, DBH, MATSIRO, ORCHIDEE, and PCR-GLOBWB) adjust some parameters according to vegetation or soil properties, but they have no hydrologic calibration. Neither JULES-W1 nor LPJmL calibrate hydrology, although they do calibrate biophysical processes and crop yield, respectively. MPI-HM and H08 are not hydrologic calibrated.

4. Creating the standard writing style of model equations

In this study, the rationale in finding similarities and differences among 16 GWMs is based on how models simulate terrestrial water cycle. We created a standard writing style for model equations and used the same symbols to write those equations, following seven steps to achieve our main goal.

4.1. Investigation of 16 global water models

Generally, the models have different style in describing their structure, defining their variables, and writing their equations. Furthermore, a unique equation can be implemented in various ways (e.g., discrete vs. analytical form, focusing on flows or water compartments) or parameterized differently. Therefore, we started our study with a literature review on the 16 GWMs...
analyzed in the present study. We analyzed the nomenclature of each model to identify a good way of writing the model equations and habits that exist in global water modelling. Another aim was to familiarize ourselves with model equations.

4.2 Generation of the lists with water storage compartments, flows, human water use sectors, included in 16 global water models

In the next step, we assembled a list with water storage compartments and human water use sectors included in the models to simulate terrestrial water cycle. We decided to describe 16 GWMs based on the equations implemented for eight water storage compartments, six human water use sectors. The analyzed water storage compartments are canopy, snow, soil, groundwater, lake, wetland, reservoir, and river. The human water use sectors are irrigation, domestic (households), livestock, manufacturing, electricity. Thus, the present model intercomparison study is based on the lists presented in Tables 1 to 5.

4.3 Creation of glossary with variables definitions

We decided upon clear definitions of the analyzed variables. However, we encountered many ambiguities and challenges in defining the analyzed variables and labeling processes as being similar or different among them. Some examples are presented in the following lines. We present other definitions of water storage compartments, flows, and human water use sectors, used in this paper work, in the supplementary information (Table S84).

We decided to use the expression *input data* for climate variables of the 16 GWMs to avoid confusion among readers. We define *parameterization* as changes of model parameter values (Samaniego et al., 2010).

We use *active vegetation* to highlight if models include the photosynthesis scheme in their structure and if they have the ability to simulate actively changes in vegetation, in an area, because of changes in the CO$_2$ concentration, air temperature, and precipitation. We use *dynamic vegetation* to define changes in vegetation from one geographical area to another because of competitive and biogeographical processes determined by climate change (geographical distribution of plants) or human activities. We decided to use *subsurface runoff* synonymously with *interflow* and to define it as the amount of water that leaves the soil layer laterally.

We define *baseflow* as the low part of the streamflow that is supplied by groundwater, drainage from lakes, wetlands, glaciers, and interflow during long periods when no precipitation or snowmelt occurs. Ultimately, we have excluded the variable *baseflow* from the analysis because it is not simulated by 16 GWMs in ISIMIP2b.

We discovered that *groundwater runoff* and *baseflow* are used synonymously and define the water that leaves groundwater storage. We also found that *baseflow* and *subsurface runoff* are used synonymously, and define the amount of water estimated for the third soil layer (VIC). We noticed that MPI-HM includes additional storage, called *baseflow storage*, that collects the drainage leaving through the bottom of the soil storage and applies a substantial time lag before passing it on to the river storage. In ISIMIP2b, the drainage computed by MPI-HM was submitted as subsurface runoff, but considering that this baseflow storage acts similarly to a groundwater storage, drainage could be used as groundwater recharge in ISIMIP3a/b.
Consequently, its outflow could be submitted as groundwater runoff. However, the purpose of this baseflow storage, for MPI-HM, is predominantly to cause a delay in river discharge and not to simulate groundwater in detail.

We decided to define groundwater recharge as the amount of water that reaches the groundwater storage, because of its hydrological meaning. However, we found out that the words drainage (MPI-HM), aquifer recharge (CLM4.5), and groundwater recharge (GHMs) are used synonymously among 16 GWMs. ISIMIP2b relates seepage with groundwater recharge for the models that do not include a groundwater storage, supposing that this water would reach groundwater storage if it would exist.

Another discovery was that throughfall and drip in some models were considered synonyms and they were used to describe precipitation that falls to the ground through canopy spaces (CLM4.5, CLM5.0, MATSIRO). In this case, we decided to separate these words and to define throughfall as being precipitation that falls to the ground through canopy spaces and drip as being precipitation that leaks at the edge of canopy.

In summary, in global water modelling, we need to be aware of differences in vocabulary. A widely accepted list of definitions would avoid confusion and facilitate successful interaction and collaboration. Furthermore, we need to clarify hydrological terms to peers from other disciplines, stakeholders, and a general audience (Brunner et al., 2018) to facilitate easier communication, understanding, and analysis.

### 4.4 Variable naming

We notated each variable of model equations. We used multiple subscripts and superscripts to properly identify water storage compartments, flows, and human water use sectors because of the large number of storage compartments included in the model structures. We selected “S” to describe water stock, “P” to describe everything connected to precipitation, “E” for everything related to evaporation, “R” for everything related to runoff, “Q” for everything related to streamflow and outflow, and “A” for water abstractions. We used two letters for subscripts and superscripts, ideally, the first two letters of the word, for example, “ca” for canopy; “sn” for snow; “so” for soil, and so on (see list of symbols and glossary in the Supplement), while we used the first letter of each word in case of compounds words such as groundwater (“gw”) or surface water (“sw”). We separated subscripts and superscripts from one another using comma. We did not write full words for subscripts and superscripts, because equations became too long and difficult to read and understand. Some of these decisions correspond with some habits that exist in the hydrological community (e.g., gw and sw) and we decided to keep them to make a comfortable and easy workflow for modelers and readers.

### 4.5 Collection of the equations from the modelling teams

In the next step, modelling teams created and provided the model equations, used to provide simulations for ISIMIP2b, according to the generated lists. Each modelling team, involved in this study, internally checked and reviewed its model, based on the model code and peer-review articles mentioned in Table 11 or only on the peer-review articles on model description mentioned in Table 11. In some cases, modelling teams provided the equations using our standard writing style and symbols.
presented in subsection 4.4, while in other cases using their specific writing style. Therefore, the modelling teams checked the model equations on their plausibility.

### 4.6 Homogenization of the model equations

We homogenized all variables and standardized variables’ units in Tables S1–S83. We used overleaf platform, an online LaTeX editor with its glossaries package, to homogenize all model equations of 16 GWMs, write some model equations, and rewrite other model equations using our symbols. This online LaTeX editor enabled us an online collaboration, correction of model equations many times, and saving a lot of time in all this process. Therefore, the supplementary information provides an overview of the 16 GWMS, analyzed in this study, and enables readers to understand similarities and differences among these models and identify included water compartments and human water use sectors and their flows. Ultimately, the readers get an overview of hydrological knowledge complexity behind these models (Tables S1–S84).

### 4.7 Evaluation of collected information

In the final step, we reevaluated the collected and homogenized model equations for plausibility. We found similarities and differences among 16 GWMS analysed in this study. We analyzed the model equations to find the models that simulate the same water flow (e.g., evaporation), the same water storage compartment (e.g., canopy storage), the same human water use sector (e.g., irrigation sector). For example, five models (CWatM, JULES-W1, MATSIRO, MPI-HM, and WAYS) use the same equation to compute potential evapotranspiration in Table S2. Ten models (CWatM, DBH, JULES-W1, LPJmL, mHM, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, and WAYS) compute changes in canopy water storage taking into account the same variables such as total precipitation, throughfall, and canopy evaporation. Other three models (CLM4.5, CLM5.0, and MATSIRO) compute changes in canopy water storage differently than the nine models, by taking into account the precipitation intercepted by canopy storage, liquid and solid throughfall, additional to canopy evaporation (Table S3). We also conclude that 12 models compute canopy evaporation (Tables S3, S7). Therefore, in the next section (section 5), we present our results according to two main parts of the terrestrial water cycle: hydrological part and water use part. The hydrological part includes the eight water storage compartments and their flows, while the water use part includes five human water use sectors and their flows.

### 5 Similarities and differences among 16 global water models

Several studies highlighted the need to understand better modeling approaches, model structures, model equations, and similarities and differences among models (Zhao et al., 2017; Veldkamp et al., 2018; Schewe et al., 2019). Therefore, in this section, we present some similarities and differences among 16 GWMS in simulating the terrestrial water cycle. This information enables us the interpretation of the different model results found in some model comparison and ensemble studies (Zaherpour et al., 2018; Wartenburger et al., 2018; Scanlon et al., 2019), as well as those by Gudmundsson et al., 2021;
5.1 Similarities and differences in simulating eight water storage compartments

5.1.1 Canopy water storage.

The changes in canopy water storage depend on how much water evaporates (canopy evaporation) and how much water is intercepted by canopy. Thirteen models include canopy water storage in their structure, while three other models do not include it (H08, Mac-PDM.20, and MPI-HM: Table S3, Figure 1). Ten models compute canopy water storage by subtracting the throughfall amount and canopy evaporation from the total precipitation. Other three models (CLM4.5, CLM5.0, and MATSIRO) compute change in canopy water storage by subtracting the liquid or solid throughfall and canopy evaporation from the precipitation intercepted by the canopy storage. MATSIRO is the only model that has two canopy water compartments: one for rainfall interception and one for snowfall interception. It also computes in detail how much water is intercepted by canopies in stormy areas with high wind speeds and in calm areas with low wind speeds. In these areas, precipitation depends, mainly, on leaf area index (LAI) and water deficit in the canopy storage. Three land surface models (CLM4.5, CLM5.0, and MATSIRO) divide total precipitation into precipitation intercepted by canopy, precipitation that penetrates the canopy and then reaches the ground (throughfall), and precipitation that falls directly on the ground (Tables S4–S6). Further, they also divide throughfall into liquid and solid phases.

Two models compute an interception scheme based on a leaf and stem area index, while seven models use only a leaf area index (Tables 7 and 8). Ten models compute this considering vegetation type (a plant functional type system) (Tables 7 and 8). MPI-HM uses prescribed data taken from Land Surface Parameter dataset version 2 (Hagemann, 2002). PCR-GLOBWB uses HYDE3.2 (Klein Goldewijk, 2017), MIRCA (Portmann et al., 2010), and GlobCover datasets (ESA GlobCover Project, 2005). Generally, prescribed vegetation ignores the decisive interaction between vegetation and runoff and interactions between the atmosphere and Earth’s surface, partly presented in section 3.2 (Gerten et al., 2004; McPherson, 2007; Nicholson, 2000).

Throughfall is estimated by 13 models (Table S5) depending on 1. total precipitation and relative canopy water content (JULES-W1); 2. difference between total precipitation and canopy storage deficit (mHM, WaterGAP2, WAYS); 3. ratio between rainfall or snowfall and total precipitation (CLM4.5, CLM5.0, MATSIRO); 4. total precipitation and minimum value of potential evapotranspiration (PET) or canopy storage (LPJmL); 5. canopy water content (PCR-GLOBWB); 6. a function of LAI then weighted by the canopy fraction in the grid cell (DBH and ORCHIDEE); 7. canopy water content and grid cell average precipitation (VIC); 8. total precipitation, canopy water content, and canopy evaporation (CWatM). Three models (H08, Mac-PDM.20, MPI-HM) do not estimate throughfall.

Four models (CLM4.5, CLM5.0, LPJmL, and ORCHIDEE; Tables 7 and 8) account for the CO₂ fertilization effect, in the LAI estimation, by using a photosynthesis scheme (active vegetation mentioned in section 4.3), and they have the ability to simulate...
the CO₂ effect on plant functioning. Generally, it was found that simulations depend on the number of plant functional types (PFTs) prescribed or defined in the model and on the processes used to estimate plants’ ability to adapt, acclimate, and grow in new environmental conditions (Sitch et al., 2008).

5.1.2 Snow water storage

Snow storage accumulates snow below freezing temperatures and declines by melting and surface and/or snowdrift sublimation. GHMs typically use the degree-day method to compute snow accumulation and snowmelt, while LSMs use the energy balance method (Tables 7 and 8, Figure 1). Among GHMs, H08 is the only one that applies the energy balance method to compute snow accumulation and melt. Additionally, three models (CLM4.5, CLM5.0, and CWatM) include glacier storage. CLM4.5 and CLM5.0 use a mechanistic snow module to calculate snow accumulation and melt; therefore, they include multiple snow layers where compaction, melt, refreezing, firn, and other snow related processes take place.

Four models (CLM4.5, CLM5.0, MPI-HM, and VIC) have two water storage compartments for snow: for estimation of frozen water and for liquid water content (Table S8). WaterGAP2 calculates snow accumulation and melting in 100 subgrid cells (Schulze and Döll, 2004; Müller Schmied et al., 2014), while CWatM calculated using 3 to 10 elevation zones per grid. Five models (CLM5.0, DBH, JULES-W1, MATSIRO, and VIC) estimate snow held on the canopy (Table S9). Further, seven models differentially estimate snow under the canopy (Table S10). Five models do not estimate sublimation: Mac-PDM.20, mHM, MPI-HM, PCR-GLOBWB, and WAYS (Table S11). All models simulate snowmelt (Table S12). MATSIRO is the only model that distinguishes between sublimation and evaporation on snow-covered ground and snow-free ground. The number of snow layers is fixed and it varies among 16 GWMs between 1 (most of the GHMs) and 12 (CLM5.0; Tables 7 and 8). Most of the GWMs present no upper limit for snow storage (Tables S48 – S51).

5.1.3 Soil water storage

Soil water storage keeps and loses water from flows above and below the ground’s surface. Hydrologically, this includes the unsaturated zone or vadose zone, the part of Earth between the land surface and the top of the phreatic zone (water table).

Soil hydrologic processes. Overall, 10 models consider initial infiltration as inflow of the soil storage, while three models (H08, JULES-W1, and WAYS) consider throughfall (Table S14). Mac-PDM.20 considers total precipitation as inflow of soil storage (Table S14). Thus, infiltration, throughfall, and total precipitation have different values among 16 models because the models compute infiltration and throughfall differently, while total precipitation represents the input data for some models. All models compute surface runoff (Table S20, Figure 1), soil evaporation (Table S24), and infiltration (Table S25), while six models compute interflow (Table S26). H08 computes runoff properties varying according to the climate zone (Table 7).

CLM4.5 includes an empirical soil evaporation resistance method, while CLM5.0 includes a mechanistically based method where the soil evaporation is controlled by a dry surface layer. Therefore, CLM5.0 has the ability to model the seasonality of soil evaporation and soil water storage in (semi-)arid regions. It also explicitly simulates spatial variation in soil thickness (0.4 to 8.5 m) and columnar water holding capacity, unlike CLM4.5 (Lawrence et al., 2019). These models have a large number of soil layers, each having moisture storage potential depending on the soil texture. They use the same approach to calculate surface runoff and have the ability to compute liquid runoff and solid runoff from snow capping. Both models consider
subsurface runoff as a product of an exponential function of the water table depth and a single coefficient (Niu et al., 2005). VIC uses the variable infiltration curve (Zhao et al., 1980) to account for the spatial heterogeneity of runoff generation, and assumes that surface runoff from the upper two soil layers is generated by those areas where precipitation exceeds the storage capacity of the soil. The mHM model has one more bucket between the soil storage and groundwater storage named “unsaturated storage” representing the source for interflow and groundwater recharge.

LPJmL was adjusted, and the water from the uppermost soil layers is considered to contribute to surface runoff if excess of storage is calculated according to the infiltration or percolation rates, which depend on soil type. LPJmL routes, what was previously lateral runoff, from “layer 0” (first 20 cm), as surface runoff.

In JULES-W1, water that reaches the soil surface is split between water that infiltrates into the soil and surface runoff. Infiltration takes place at a rate equal to saturated hydraulic conductivity multiplied by an infiltration enhancement factor, which is dependent on the presence and type of vegetation. If a soil layer becomes saturated, the water in excess of saturation is put into the layer below. JULES-W1 also uses a “zero-layer” scheme that does not use explicit model layers to represent snow, instead adapting the topsoil level to represent existent snow processes. In the original “zero-layer”, snow scheme has a constant thermal conductivity and density. Bulk thermal conductivity of snow on the surface layer decreases due to both the increased layer thickness and the different conductivities of snow and soil. Surface energy balance and heat flux between the surface layer are controlled by insulation factors and layer thickness (Best et al., 2011). WAYS simulates the water storage and flows in soil only for the entire root zone (Table 8). In the DBH model, runoff is generated directly when soil layer is saturated, or is generated when rainfall intensity is larger than the infiltration rate estimated with the Green–Ampt method (Tang et al., 2006).

Two models (CWatM and MPI-HM) have an additional water storage compartment to compute the runoff concentration in a grid cell that has a lag time before entering the river storage compartment (Table S44). Consequently, this storage serves to create a delay between runoff and streamflow, and accounts for the average distance that runoff, generated at a specific point within a grid cell, has to travel before reaching the river. This storage collects water from rivulets and creeks or concentrates runoff in rivulets and creeks before it enters the river storage, because the rivulets and creeks are smaller than the size of a single grid cell and have different water retention properties from the main river channel within the grid cell. Therefore, this compartment does not act as a floodplain, to delay floods, or as overland flow, to express too much water in the soil. In its original structure, MPI-HM named this compartment “overland flow”, but we decided to rename it “rivulet storage” to avoid confusion among readers.

Some GWMs compute vertical water movement in unsaturated soils by applying the Richards equation (Richards, 1931; e.g., CLM4.5, CLM5.0, CWatM, JULES-W1, MATSIRO, ORCHIDEE, VIC). However, the Richards equation may not be relevant for the models that have one soil layer because of its complexity and of missing capillary rise (Lee and Abriola, 1999; Farthing and Ogden, 2017). LPJmL uses a percolation scheme to estimate vertical water movement that applies the storage routine technique developed by Arnold et al. (1990) and simulates free water in the soil bucket. DBH uses the Green–Ampt equation to compute infiltration in unsaturated soils.
Five models compute capillary rise (CLM4.5, CLM5.0, CWatM, MATSIRO, and PCR-GLOBWB), with CWatM and PCR-GLOBWB using the same approach (Table S28).

**Soil column configuration.** Number of soil layers ranges between 1 (H08, MPI-HM, and WaterGAP2) and 25 (20 soil layers + 5 bedrock layers: CLM5.0), while total soil depth is between 1 m (H08) and 49.6 m (CLM5.0; Tables 7 and 8). ORCHIDEE uses a relatively deeper soil column to account for soil thermic. LPJmL has five hydrologically and thermal active soil layers plus one thermal active soil layer. MPI-HM defines soil storage in terms of the maximum water column, varying between 0 and 5 m; therefore this cannot be translated into soil depth directly.

### 5.1.4 Groundwater storage

Groundwater storage, beneath the soil water storage compartment, receives water from drainage (e.g., MPI-HM) or aquifer recharge (e.g., CLM4.5) or groundwater recharge (e.g., WaterGAP2) (Tables 9 and 10). It loses water through capillary rise, groundwater runoff, and groundwater abstraction for human water use. In GWMS, groundwater compartment simulates hydrologically the saturated zone or phreatic zone (WaterGAP2) or an unconfined aquifer (CLM4.5). Eleven models include groundwater storage in their structure, and most of them have only one groundwater layer (Table S29, Figure 1). In ISIMIP2b, two models (JULES-W1 and LPJmL) consider the water excess from the bottom soil layer as *seepage* and relate this variable with *groundwater recharge* because they do not have a groundwater compartment.

CLM4.5 simulates an unconfined aquifer as a groundwater component, below the saturated soil storage and with a prescribed maximum value (5000 mm), while CLM5.0 simulates an impermeable bedrock with five layers and therefore assumes no groundwater flow as bottom boundary conditions. In CLM4.5, the unconfined aquifer interacts with the saturated soil storage through the water table, whether it is within or below this storage. When the water table is below the soil storage, the aquifer recharge is estimated by applying Darcy’s law across the water table (Lawrence et al., 2019).

MATSRO has a dynamic groundwater scheme (Koirala et al., 2014; Pokhrel et al., 2015) in which the number of soil layers in the saturated zone (i.e., groundwater) varies in time between 1 and 13 depending on water table location (Table 7). The two-way interaction between the unsaturated zone (for which vertical moisture movement is resolved by solving the Richards equation) and the underlying aquifer is simulated through moisture flux exchange at the water table. This flux exchange is determined as the algebraic sum of downward gravity drainage from the unsaturated soil layer overlying the water table and the upward capillary flux (Koirala et al., 2014; Pokhrel et al., 2015). The water balance of the saturated zone is resolved by considering recharge to the groundwater aquifer and groundwater runoff that is determined by using a two-parameter, statistical-dynamical formulation considering soil hydraulic properties and basin geomorphology (Yeh and Eltahir, 2005). The variation in the water table is also determined by the aquifer specific yield.

In Mac-PDM.20, it is assumed that all water in excess of field capacity drains in one day to the deep store, which for ISIMIP2b is used to represent groundwater recharge ($R_{gwr}$). The total runoff ($q_{tot}$) is the sum of direct runoff ($q_r$) plus delayed runoff from the deep soil and groundwater ($q_{sb}$). This delayed runoff ($q_{sb}$) is assumed to be a non-linear function of the amount of water held in the groundwater and deep soil store (Table S31). Thus, like with MPI-HM, the purpose of the delayed runoff (or baseflow) is predominantly to cause a delay in river discharge and not to simulate groundwater in detail.
H08 separates groundwater into renewable and one nonrenewable layer (Hanasaki et al., 2008). WaterGAP2 is the only model that simulates the groundwater recharge from surface water bodies in semiarid and arid grid cells (Döll et al., 2014). Fifteen models compute groundwater recharge, three using the same approach (H08, WaterGAP2, and WAYS: Döll and Fiedler, 2008; Table S30), while twelve models compute groundwater runoff (Table S31).

5.1.5 Lake storage

Lake storage fills with water through flows above and below the ground and stores water for a certain residence time. It loses water through discharge to other storages, evaporation, groundwater recharge, and water abstraction for human water use. Ten models do not include lakes (Tables 9 and 10; Figure 2). Five models compute evaporation from lakes, three of them based on a PET approach (Table S33), while four models compute outflow from lakes (Table S34). CLM4.5 and CLM5.0 compute the lake storage as virtual storage where the difference between precipitation and evaporation is balanced automatically by their outflow, named lake runoff. CLM4.5 uses constant lake depth, while CLM5.0 uses spatially variable lake depth, and freezing and thawing are included in the lake body (Vanderkelen et al., 2020).

LPJmL treats natural lakes and rivers in a similar way in terms of inputs and output. Lake inputs to a river can also include upstream river inputs to the lake. LPJmL also keeps track of a lake fraction in the river input. WaterGAP2 and CWatM have two types of lake storage: “local lake storage”, gets water from runoff resulting within the cell, and “global lake storage”, gets water from runoff resulting within the cell and the upstream cell (Müller Schmied et al., 2021).

5.1.6 Reservoir storage

Reservoir storage fills with water behind dams through flows above and below the ground and stores water for a residence time. It loses water through discharge to other storages, evaporation, groundwater recharge, and water abstraction for human water use. Ten models (CLM4.5, CLM5.0, DBH, JULES-W1, Mac-PDM.20, mHM, MPI-HM, ORCHIDEE, VIC, and WAYS) do not include reservoir storage for ISIMIP2b (Tables 9, 10, S35; Figure 2). Six models compute outflow from reservoirs (Table S37), while evaporation from reservoirs is computed by four models (Table S38).

In general, most of the models use the Global Reservoir and Dam database (GRanD: Lehner et al., 2011), but with a different number of active managed reservoirs, used for reservoir operation during simulations. Three models (LPJmL, WaterGAP2 and PCR-GLOBWB) merge more than one reservoir per grid cell into one reservoir, if required.

Four models (CWatM, H08, MATSIRO, and WaterGAP2) use two water compartments, global and local reservoirs, to represent the reservoirs, following the reservoir algorithm developed by H08. However, there are some differences on how the scheme was implemented in the models, mainly, because of model structure, but the approach is essentially the same. These four models use the same approach in selecting active managed reservoirs for reservoir operation, but they use different thresholds. WaterGAP2 considers 1109 active managed reservoirs and handles reservoirs below 0.5 km³ storage capacity as local lakes. MATSIRO considers only 728 out of 6862 reservoirs for reservoir operation. In MATSIRO, global reservoirs have more than 1 km³ total storage capacity and ”local reservoirs” or ”ponds” have less than 1 km³ (around 6134 reservoirs; Hanasaki et al., 2006; Pokhrel et al., 2012a and b). H08 considers 963 active managed reservoirs (global reservoirs) and 5824 local reservoirs; therefore, global reservoirs regulate river flow, while local reservoirs do not. Global reservoirs have 4773 km³ of...
total storage capacity, while local reservoirs have 1300 km$^3$ of total storage capacity. In H08, when multiple local reservoirs are present in a grid cell, their capacity is added together. CWatM considers 3663 active managed reservoirs, while PCR-GLOBWB considers 6177. LPJmL includes 4134 reservoirs that become active after the first year of operation. In LPJmL, reservoirs are not managed according to an operation scheme, they are modeled as lakes with a maximum storage amount and the water over this amount is released as reservoir outflow; irrigation water can also be taken from the reservoir.

Five models (CWatM, H08, LPJmL, MATSIRO, and WaterGAP2) use a retrospective reservoir algorithm, while one model (PCR-GLOBWB) uses a prospective reservoir algorithm. The retrospective reservoir algorithm uses river flows and water demand, which were processed in a previous step, while the prospective reservoir algorithm uses forecasts of river flows and water demand (van Beek et al., 2011).

5.1.7 Wetland storage
Wetland storage fills and empties with water similarly to lake and reservoir compartments, except that water use is not satisfied from wetlands. Two models (MPI-HM and WaterGAP2) compute wetland compartment, evaporation, and outflow from land (Tables S39–S42; Figure 2). WaterGAP2 has two types of wetland storage: “local wetland storage”, which obtains water from runoff resulting within the cell, and “global wetland storage”, which obtains water from runoff resulting within the cell and the upstream cell (Döll et al., 2012).

5.1.8 River storage
River storage is increased by surface and sub-surface runoff. It loses water through streamflow, evaporation, channel transmission, and water abstraction for human water use. Five models (DBH, JULES-W1, Mac-PDM2.0, VIC, WAYS) do not include river storage for ISIMIP2b simulations, because of computational and resource constraints, nor do they compute streamflow (Tables 9, 10, S43, and S46; Figure 2). Four models (LPJmL, MATSIRO, MPI-HM, WaterGAP2) use a linear reservoir cascade approach to compute the water balance of the river storage (Tables 9 and 10). Furthermore, MATSIRO uses Total Runoff Integrating Pathways (TRIP) for river routing through a channel. Three models (CWatM, H08, and LPJmL) consider the minimum release for environmental flow. CWatM adopts a kinematic wave approach, approximation of the Saint-Venant equation (Chow et al., 1998), linked with dynamic reservoir and lake operation. Further, CWatM computes runoff concentrated in creeks and rivulets, with a lag time before entering the river storage, by using a triangular weighting function (Burek et al., 2020). ORCHIDEE includes a river transport module that involves the Simulated Topological Network (STN-30p). PCR-GLOBWB uses a travel time routing (characteristic distance) linked with dynamic reservoir operation. For runoff and streamflow simulation, CLM4.5 uses a river transport model (RTM), while CLM5.0 uses a new mechanistic model for streamflow routing, called the Model for Scale Adaptive River Transport (MOSART; Oleson et al., 2013, Lawrence et al., 2019). The mHM model uses a mesoscale routing model with an adaptive time step according with the spatially varying celerity (Thober et al., 2019). Only MPI-HM and ORCHIDEE include a routing model with a wetlands and floodplain scheme, in which wetlands act as floodplains. Furthermore, ORCHIDEE includes swamps.

Six models (CLM5.0, CWatM, MPI-HM, ORCHIDEE, PCR-GLOBWB, WaterGAP2) apply the Manning–Strickler equation to estimate river flow velocity and use various values for it. CLM4.5 uses a standard river flow velocity of 0.35 m s$^{-1}$, while
H08 and MATSIRO use 0.5 m s\(^{-1}\) (Tables 9 and 10). LPJmL considers a standard river flow velocity of 1 m s\(^{-1}\). MPI-HM uses the Manning–Strickler equation only for flow velocity computation in wetlands, while, for rivers, it computes a slope-dependent flow velocity following the approach by Sausen et al. (1994).

Inflow from upstream grid cell surface water bodies represents the sum of inflow water from neighboring upstream grid cells for CLM4.5, CLM5.0, CWatM, mHM, and WaterGAP2 (Table S45). Additionally, CWatM and WaterGAP2 route this water also through lakes and reservoirs before it reaches its final point. H08 computes it as being the product between a 0.5 m s\(^{-1}\) flow velocity and river storage from upstream grid cells. LPJmL considers it as being the outflow of river storage reduced by evaporation from lakes and reservoirs, while MPI-HM considers it as being the sum of outflow from rivulet storage, groundwater runoff, and streamflow from the upstream grid cells, then reduced by inflow from the wetland of an upstream grid cell. MATSIRO considers it as being the sum of inflow water from the neighboring upstream grid cell multiplied by outflow of river from an upstream grid cell. ORCHIDEE calculates it as being the sum of stream river storage of upstream grid cells divided by topographic index of the retention time and a reduction factor of stream river storage. PCR-GLOBWB takes into account the outflow from river storage, time of process duration, length of river sections, and the coefficient friction of the reservoir weir.

Evaporation from rivers is computed only by three models, CWatM, LPJmL, and PCR-GLOBWB, based on a PET approach (Table S47).

### 5.2 Similarities and differences in simulating human water use sectors

Some GWMs simulate water extracted from surface water compartments and/or a groundwater compartment that is used for human activities. Human water abstraction represents the sum of the water consumed by humans, evaporative water and other water losses (named water consumption), and water returned to the groundwater or surface water compartments (named return flow, being the part of the water not consumed). Generally, three models extract water for human activities from groundwater or surface water bodies (H08, PCR-GLOBWB, and WaterGAP2). Seven models (DBH, JULES-W1, Mac-PDM.20, mHM, ORCHIDEE, VIC, and WAYS) do not include any human water use sectors in their structures (Table 6).

#### 5.2.1 Irrigation sector

Irrigation water demand (potential irrigation water abstraction) is computed by nine models (Table S52). Groundwater abstraction for the irrigation sector is simulated by six models (CWatM, H08, MATSIRO, MPI-HM, PCR-GLOBWB, and WaterGAP2: Tables S53), while five models compute the return flow (Table S55). Irrigation surface water abstraction is calculated by nine models (Table S56, Tables S93–S94). CWatM includes a “normal irrigation scheme”, to mimic rainfall when the plants need it, and a paddy rice irrigation scheme, to mimic the flooding of the rice area (Table S56). The water source for the irrigation sector is river for nine models (CLM4.5, CLM5.0, CWatM, H08, LPJmL, MATSIRO, MPI-HM, PCR-GLOBWB, WaterGAP2: Table S93). Six models (CWatM, H08, MATSIRO, MPI-HM, PCR-GLOBWB, WaterGAP2) consider groundwater a source for the irrigation sector (Table S93). Four models take water from lakes for the irrigation sector and five models take water from reservoirs (Figure 3). Return flows from irrigation sector recharge mainly
the soil and groundwater (seven models), while the return flows from domestic and manufacturing recharge mainly rivers (four models; Figure 4).

5.2.2 Domestic, livestock, and industry sectors

Five models (CWatM, H08, MATSIRO, PCR-GLOBWB, and WaterGAP2) simulate water abstraction, water consumption, and return flow for the domestic sector (household: Tables S59–S64). Three models (MATSIRO, PCR-GLOBWB, and CWatM) combine manufacturing and electricity sectors in one sector, the industry sector. CWatM only calculates total abstraction from groundwater or surface water. MATSIRO and LPJmL used prescribed data for water demand of the domestic and industry sectors, offered by the ISIMIP2b framework. These input datasets provide water consumption, but not return flow from these sectors. Consumption water can return to the atmosphere as evapotranspiration. LPJmL used prescribed data for domestic and industrial water consumption data and assumed that only the consumed water amount is withdrawn. MATSIRO used prescribed data for domestic and industrial water demand and it computed itself the water abstraction and consumption for these sectors. ISIMIP2b does not offer prescribed data for livestock sector as the global numbers are, compared to other sectors, low (Müller Schmied et al., 2016). PCR-GLOBWB computes amount of water abstracted and consumed for livestock sector, taken from groundwater and surface water bodies (Tables S65–S68), while WaterGAP2 computes only the amount of water taken from surface water bodies for livestock (Tables S67–S68).

5.2.3 Surface water abstractions

Four models (CWatM, MATSIRO, MPI-HM, and WaterGAP2) compute total groundwater abstraction (Table S77). Five models (CWatM, LPJmL, MPI-HM, PCR-GLOBWB, and WaterGAP2) compute total lake abstraction (Table S78). Six models (CWatM, H08, LPJmL, MATSIRO, PCR-GLOBWB, and WaterGAP2) compute total reservoir abstraction (Table S79). Three models (CWatM, CLM5.0, and WaterGAP2) compute total river abstraction (Table S80). CWatM calculates the water withdrawal in total from all users and afterwards it distributes the total withdrawal to different sources: surface water, sustainable groundwater (available groundwater = long-term groundwater recharge of the last 30 years in the analyzed time interval), unsustainable groundwater human water use sectors (domestic, livestock, irrigation, industry). Each withdrawal that is depleting the groundwater storage beyond groundwater recharge is using fossil groundwater (unsustainable groundwater).

MATSIRO and WaterGAP2 take similar approaches to compute groundwater abstraction: groundwater abstraction for the irrigation sector is reduced by the sum of groundwater abstraction for the domestic and industry sectors. MPI-HM considers groundwater abstraction as being equal only to groundwater abstraction for the irrigation sector, as other sectors are not included in the model. MPI-HM considers lake abstraction equal to surface water abstraction for the irrigation sector.

H08 considers reservoir abstraction as being sum of monthly water abstraction for the irrigation, industry, and domestic sectors. LPJmL computes lake and reservoir abstraction by adding up the gross irrigation requirement and household, industry, and livestock demand at the grid cell with the gross irrigation requirement and household, industry, and livestock demand at the downstream grid cell.
MATSIRO computes reservoir abstraction by adding up water abstraction from reservoir for the domestic, industry, and irrigation sectors.

PCR-GLOBWB computes lake and reservoir abstraction by adding up water abstraction demand for the industry, irrigation, domestic (household), and livestock sectors.

CLM5.0 considers river abstraction equal to water abstraction for irrigation sector.

WaterGAP2 computes lake, reservoir, and river abstractions as the sum of water abstraction for the irrigation, livestock, domestic, manufacturing, and electricity sectors taken from surface water bodies. The net surface water abstraction is satisfied in WaterGAP2 in the following order: 1) river, 2) global lakes and reservoirs, and 3) local lakes.

6 Number of water flows, water storage compartments, and human water use sectors included in the 16 GWMs

One way of showing the model structures is to count the number of water flows, compartments, and human water use sectors included in each model participating in ISIMIP2b. For example, a model includes three water compartments if it computes canopy water storage, soil water storage, and snow water storage. In this section, we want to increase readers’ awareness of model structures and offer the readers a final overview of how the models work, and how many water storage compartments, flows, and human water use sectors are included in their structures. We consider that two issues are useful to interpret model results, first, knowing model structures, and, second, identifying the effect of model structures on model results. However, the present study is focused only on the first issue, respectively, knowing model configurations needed to interpret various model results.

Generally, GHMs have a high number of water storage compartments because their main purpose is to simulate the water cycle. LSMs and DGVMs have a relatively smaller number of processes (in this count and in this study), but each process has a mechanistic interpretation. LSMs exclude some hydrological processes because they are not relevant for their research purpose, spatial resolution, or cannot be parametrized in a general manner, adding some uncertainty.

In this study, WaterGAP2 includes the highest number of water storage compartments (11; see Figure 5), while DBH, JULES-W1, Mac-PDM.20, and VIC have the lowest, three water compartments (Figure 5). Others include CWatM, with ten compartments, then MATSIRO (seven compartments), followed by six models (CLM4.5, CLM5.0, H08, LPJmL, MPI-HM, and PCR-GLOBWB) with six compartments.

Water flows range between 13 (Mac-PDM.20) and 29 (CWatM) and water storage compartments range between 3 (VIC and Mac-PDM.20) and 11 (WaterGAP2), among nine GHMs.

Water flows range between 15 (JULES-W1) and 25 (MATSIRO) and water storage compartments range between 3 (DBH and JULES-W1) and 7 (MATSIRO), among six LSMs.

LPJmL, as a DGVM, simulated 22 water flows and 6 water storage compartments.

Seven models do not simulate water used by humans for economic purposes such as irrigation, domestic, livestock, manufacturing, electricity, and desalination (Figure 6). Three models (CWatM, MATSIRO, and PCR-GLOBWB) combine the
manufacturing and electricity sectors in one sector: the industry sector. WaterGAP2 simulates five human water use sectors: irrigation, domestic, livestock, manufacturing, and industry. Two models (PCR-GLOBWB and CWatM) simulate four human water use sectors such as irrigation, domestic, livestock, and industry. H08 simulates four human water use sectors: irrigation, domestic, industry, and desalination. MATSIRO simulates three human water use sectors: irrigation, domestic, and industry.

Four models (CLM4.5, CLM5.0, LPJmL, and MPI-HM) simulate only water used by humans for the irrigation sector. WaterGAP2 and CWatM have the highest number of water flows (23) to simulate human water use, while MPI-HM has the lowest number (3; Figure 6). Water flows range between 3 (MPI-HM) and 23 (CWatM and WaterGAP2), among five GHMs. Water flows range between 4 (CLM4.5 and CLM5.0) and 19 (MATSIRO), among three LSMs. LPJmL used four water flows to simulate irrigation sector.

Ultimately, GWMS include in their structure similar processes, but they are lacking other processes, mentioned in section 5, or include other processes resulting in different model structures or have used other parameter values determining various model results (Figure 3 and 4). Therefore, in section 7 we present future research on model development of 16 modelling groups involved in the present study.

### 7 Potential future research of 16 global water models

Each model, analyzed in this study, is continuously updated with the purpose to improve simulations. Therefore, in this section, we summarize model developments done outside ISIMIP framework and potential future research of 16 GWMS (Tables S95 and S96). Each modelling team collected and provided these model developments. Some of the 16 analyzed GWMS include, in their original structure, additional water storage compartments, water flows, and human water use sectors that have not been used for ISIMIP2b. Additional information on the 16 analyzed GWMS can be found in the peer-review articles mentioned in Table 1.

Some analyzed GWMS have the ability to operate at various spatial–temporal scales: CWatM, CLM4.5, CLM5.0 (3 h time step at around 11 km).

CLM team improved the irrigation scheme (Thiery et al., 2017; 2020), the representation of land cover and land management (Meier et al., 2018; Hirsch et al., 2017; 2018), and the implementation of reservoirs (Hauser et al., 2019). Numerous developments can be followed on the model’s GitHub page ([https://github.com/ESCOMP/CTSM](https://github.com/ESCOMP/CTSM)).

CWatM developed a groundwater scheme with linkages to MODFLOW for 5 arcmin and 30 arcsec spatial resolution. The CWatM modeling group plans to develop a reservoir storage including different operation schemes (e.g., energy, irrigation), to increase the temporal resolution (at 1 h), to apply a global calibration also for ungauged catchments, such as using the Budyko framework (Greve et al. 2020), applying both the day-degree method and energy balance method to estimate snow accumulation and melt, and applying several methods to estimate evaporation based on changing CO₂ concentration.
DBH plans to include human water uses (industrial and domestic sectors), either by developing a new module or using the simulations from other models (e.g., WFaS dataset), to calibrate the model in the new ISIMIP3 simulation round, and to improve the input/output module to read and write netcdf files.

The H08 modeling team used an approximate Bayesian computation technique to calibrate four parameters that are transferred to other regions containing no observations, mainly, based on Köppen–Geiger regions. The modeling group also increased the spatial resolution to 5 min and improved the representation of crops used for biofuel in the model.

The JULES-W1 modeling group plans to make a technical update that will enable the river routing module to estimate discharge.

The LPJmL group developed an improved energy balance module and soil hydrological scheme that can estimate permafrost dynamics (Schaphoff et al., 2013) and made the model source code freely available on GitHub (https://github.com/PIK-LPJmL/LPJmL; Schaphoff et al., 2018), hoping to engage a broader scientific community in LPJmL model development and applications.

The Mac-PDM.20 modeling group plans to develop a water use module.

The LPJmL group developed an improved energy balance module and soil hydrological scheme that can estimate permafrost dynamics (Schaphoff et al., 2013) and made the model source code freely available on GitHub (https://github.com/PIK-LPJmL/LPJmL; Schaphoff et al., 2018), hoping to engage a broader scientific community in LPJmL model development and applications.

MATSIRO modeling group has implemented a land-use change process, terrestrial biogeochemical processes, and an additional crop growth process into MATSIRO to develop a new modeling framework. As key interactions are taken into account and all processes are coupled, important boundary conditions for hydrological simulations can be dynamically simulated internally. This hydrological simulation modeling framework has been coupled with MIROC GCM and has been used as an Earth system model. In addition, the group recently proposed new schemes for lateral groundwater flow, water temperature, and sediment transportation.

Ongoing efforts to improve the realism of hydrological processes in the mHM include the development of the multiscale lake module (mLM), a comprehensible framework for reservoir regulation as well as natural processes in lakes. Near-future developments will focus on a glacial module, to better account for processes in cold regions, as well as coupling it to a groundwater model that will replace the current linear groundwater reservoir.

The MPI-HM modeling group plans to increase the spatial resolution of regional versions. The group currently implemented canopy storage in the latest model version and is developing experiments to integrate reservoir storage.

The ORCHIDEE group is focusing on calibration, soil storage, groundwater storage, river storage, reservoir storage and wetland storage (MacBean et al., 2019; Verbeke et al., 2019; Yin et al., 2020; Schrapffer et al., 2020; Mizuochi et al., 2020). The PCR-GLOBWB modeling group plans to increase the temporal and spatial resolution of the input data, to increase the temporal resolution (3 h) for energy balance calculations and the global spatial resolution (1 km), to improve the soil representation by including the Richards equation, to add more snow elevation layers, to include additional fast runoff component for improving daily discharge simulations, and to improve the reservoir operating scheme (Sutanudjaja et al., 2018). The VIC modeling group developed different irrigation practices (Shah et al., 2019a and b) and included a reservoir (Dang et al., 2019 and 2020) as well as a groundwater scheme in the model structure.
The WaterGAP2 modeling group plans to update the GRanD dataset used by the model, to include water temperature calculations, to couple the new developed groundwater model (Reinecke et al., 2019), and to update the non-irrigation water use datasets.

The WAYS modeling group plans to develop a new human water use module to consider agricultural, industrial, and domestic water use in the water cycle.

8 Recommendations for future multi-model intercomparison projects and extended assessments

We assert that this study was realized through a multi-model intercomparison project (ISIMIP) and is based on communication and collaboration. Ideally, through a unified perspective and effective collaborations towards physically realistic hydrologic models (Clark et al., 2015, Clark et al., 2017), communities will fill in existing knowledge gaps (Wagener, 2020), improve the quality of the input data and the processes in the models, and implement the missing processes in the models. In addition to these statements, we propose focusing the effective collaborations on effective *wish lists*, including specific research questions, goals to answer these questions, methods to achieve the goals, datasets to be used, tasks to be done, and, at the end of the project, a retrospective analysis on what has been done and improvement options. Certainly, collaboration among these communities results in new multi-model intercomparison projects (MIPs) and multi-model ensembles that facilitate new analyses, comparisons, understandings, and improvements.

However, many studies highlighted the need to design hydrological inter-model comparison studies by nominating models or research questions according to some specific criteria (Gupta et al., 2008; Clark et al., 2011; Gupta et al., 2012), for example, (i) specific model compartments (Nazemi & Wheater, 2015; Wada et al., 2017), (ii) specific evaluation metrics (Gupta et al., 2009; Veldkamp et al., 2018; Zaherpour et al., 2018), (iii) locations of specific hydrological indicators, regions, or rivers (Masaki et al., 2017; Veldkamp et al., 2018).

In global water modelling, there are some more methodologies that can be tested to evaluate multi-model structures and model equations, also considered as hypotheses on runoff generation, for example, Rainfall-Runoff Modelling Toolbox (Wagener et al., 2001); the rejectionist framework (Vaché and McDonnell, 2006); Framework for Understanding Structural Errors (FUSE, Clark et al., 2008); SUPERFLEX (Femia et al., 2011); Catchment Modelling Framework (CMF, Kraft, 2012); Structure for Unifying Multiple Modeling Alternatives (SUMMA, Clark et al., 2015 a and b). Other methodologies can be used to evaluate parameter values such as Model Parameter Estimation Experiment (MOPEX: Duan et al., 2006), multiple-try DREAM(ZS) algorithm (Laloy and Vrugt, 2012), Generalized Likelihood Uncertainty Estimation methodology (GLUE: Beven and Binley, 2014), perturbed parameter ensembles (Gosling, 2013), the Uncertainty Quantification Python Laboratory platform (UQ-PyL: Wang et al., 2016), Multiscale Parameter Regionalization (MPR, Samaniego et al., 2010 and 2017). Thus, some existing methods might offer some solutions for reducing the high number of parameters and their values still found in global water models, and to apply more reasonable regionalization schemes in global water research (Bierkens, 2015).
Other methods can be found in frameworks proposed by Döll and Romero-Lankao, 2017 and Kundzewicz et al., 2018. Furthermore, some studies have tested how model equations combined in different configurations and using different parameter values influence the simulations: Essery et al., 2013 (testing 1701 snow models); Niu et al., 2011 (Noah-MP model); Pomeroy et al., 2007 (Cold Regions Hydrologic Model, CRHM); Kuppel et al., 2018 (Ecohydrologic model, EcH2O). In summary, they found that some model configurations provide consistently good results, others provide consistently poor results, and many configurations provide good results in some cases and poor results in others (Essery et al., 2013).

We recommend, for the benefit of the MIPs, to 1. maintain very good documentation of the model code; 2. always start research with a list, for example, with water storage compartments, flows, and human water use sectors included in the model structures; 3. have clear definitions of the variables, water storage compartments, flows, and human use sectors, describing exactly their role in the model; 4. have synonyms for variables, helping to show similarities and differences among models; 5. collect all ideas, recommendations, and improvements received from everyone (in our case, they were required to complete our study); 6. collaborate and communicate with peers, which was very useful in our study for identifying synonyms among communities; 7. describe your model or a model through your eyes and other’s people eyes; 8. invest much time and patience and be meticulous about extracting equations of water storage, flow, and human water use sectors from the model code.

We encourage communities to write and convey a clear, simple, and understandable text for large audiences. We consider that simplicity improves communication, and communication starts with a common language, the same words having the same meaning for the sender and the receiver. While trivial in theory, in practice there are some discrepancies among scientists, as well as between scientists and stakeholders by using vocabulary differently in climate impact science (Sultan et al., 2020).

Our future research will include describing the GWMs analyzed in this study, through a standard visualization of the water cycle that will show the water storage compartments, water flows, and human water use sectors included in the ISIMIP2b model structures. These diagrams would be connected with the tables presented in the supplement of the present paper (Tables S1–S83). Another future study might focus on the numerical implementation of each model code.

We note that this review and description study had a positive impact on the modeling groups, motivating them to re-think and re-analyze model structures, equations, and descriptions. We affirm that ISMIP global water sector needs to organize workshops on some parameterization experiments, by changing model parameter values. Other evaluation studies could focus on the equations applied to compute water compartments, water flows, and human water use sectors, as well as considering model outputs, to identify the effect of different water compartments on model results. ISIMIP community could increase the number of regional and pilot studies, that could validate global studies, and the number of cross-sectoral climate impact assessments.

Certainly, simulating the terrestrial water cycle on the global scale involves many challenges, as we presented in this study. Other challenges have also been synthesized by reviewing articles published by the climate, global hydrological, and vegetation communities and have been classified according to the 23 unsolved problems in hydrology (UPH) identified by Blöschl et al., 2019 (Table S97). In summary, these challenges can generally be overcome through innovative and creative collaboration among communities and investment in technical infrastructure. In the end, Arheimer et al., 2020 showed that the catchment
models can be applied at a global scale because of the new global datasets, increased computational capacity, new methods to estimate parameters, and collaboration. Ultimately, GWMS may even become a part of the Earth System Models used to simulate the water cycle at a high resolution, including human water demand and use (Wood et al., 2011; Bierkens et al., 2015).

9. Conclusions

Global water models are used to simulate the climate–water–human system. However, recent evaluation studies show that there is a need to better simulate this system by including other hydrological processes, data on physical infrastructure, societal behavior, cultural behavior, water diversions, and virtual water, as well as by identifying its teleconnections on the global scale (Zaherpour et al., 2018; Veldkamp et al., 2018; Wada et al., 2017). Some studies also underline the need to better explain various model results and better understand how models work (Reinecke et al., 2021; Pokhrel et al., 2021).

We undertook the present study mainly to find similarities and differences among global water models that will facilitate interpretation of various results, as well as those of further intercomparison studies. We developed a standard equation writing style to achieve this goal. We found that there are some similarities among the models when applying similar equations for the same hydrological processes; however, model structures are different and various values have been used for parameters or variables.

In summary, we mention that our approach was affected by models’ complexity and is limited to eight water storage compartments and their flows, and six human water use sectors mainly, because of models’ complexity. We conclude that the standard writing style of the equations is useful and necessary for finding similarities and differences among models for each water storage, human water use sector, and water flow. In addition, it can be leveraged for explaining the different model outputs, for classification of the models based on cluster analysis, and for selecting the right model for the right application. It can also be used for drawing a standard schematic visualization of the water cycle, for describing models on ISIMIP and ISIpedia platforms (the open climate-impacts encyclopedia, a part of the ISIMIP, https://www.isipedia.org/), and for understanding how models work. Other modelling teams can apply, in their studies, our lists with water storage compartments, flows, and human water use sectors and the symbols presented in the supplementary information. They can follow our steps in creating a standardized writing style of model equations and they may be aware of some challenges that could encounter. This study represents a roadmap in finding similarities and differences among models. However, it should be noted that these equations are available only for model versions used for ISIMIP2b.

We consider this study as a blueprint for other studies because it offers a practical approach to identify similarities and differences among models that are necessary for a better interpretation of their various results. We highlight the need to undertake experiments on individual water compartments in order to analyze the equations and parameters used, as well as the results obtained. We also underline the need to make multi-model intercomparison projects: firstly, because they enhance collaboration and communication between modeling groups, communities, countries and
cultures; secondly, through communication and collaboration, these projects enhance creativity and open opportunities to finding new ways to improve the models.

**Supplement**

Tables with equations of each water storage, water flow, human water use sectors, datasets used by global water models, models’ structures, future research perspectives.

**Acknowledgement**

The present study was possible through the international ISIMIP framework. CET and HMS are supported by the German Federal Ministry of Education and Research (grant no. 01LS1711F). We thank to Jacob Schewe, Agnès Ducharne, Lukas Gudmundsson, and Lamprini Papadimitriou for their useful recommendations on the manuscript. We also thank to the ISIMIP coordination team for providing funds for the ISIMIP2b Global Water Paper-writing Workshop (30 July – 01 August 2018), Frankfurt am Main, Germany and the Cross-sectoral ISIMIP Workshop 2019 (04 June 2019), where this manuscript could be discussed as well as to the ISIpedia Stakeholder Engagement Team (Inga Menke, Quentin Lejeune, Gina Maskell, Kaylin Lee) and ISIpedia stakeholders (Anatolii Shmurak, Atiq Kainan Ahmed, Aimé Koudou) by providing useful recommendations to create the standard visualization of the water cycle simulated by the models analyzed in this article. GL is supported by the National Key Research and Development Program of China (No. 2020YFA060032) and the National Natural Science Foundation of China (No. 42077420).

**Code availability**

Information on the availability of source code for the models featured in this article can be found in the Table 1.

**Author contributions**

CET proposed, designed, and led the conceptualization of ISIMIP2b models’ intercomparison based on their water cycle simulation and of a standard writing style of their equations. Initial idea was proposed by HMS, respectively, to compare the ISIMIP2b models based on their equations and a standard diagram. CET, WT, GL, PB, OR, TS, TT, FH, SR, HMS provided LaTeX support to write equations on the overleaf platform. HMS, WT, GL, PB, XL, JESB, LPS, MG, YS, OR, TS, JC, NW, HLS, TT, and GM checked the code and provided models’ equations for each water storage and water flow. CET, HMS, WT, GL, PB, XL, JESB, LPS, MG, YS, OR, TS, JC, NW, HLS, TT, GM, AK, YP, LS, YW, VM, JL, and PD checked the consistency and correctness of the equations. HMS, WT, GL, PB, XL, JESB, LPS, MG, YS, OR, TS, JC, NW, HLS, GM run simulations for ISIMIP2b project. SNG and HMS coordinated the ISIMIP2b model simulations of the global water sector.
CET conducted the analysis, visualization, and wrote the original draft. All authors reviewed, commented, and contributed to the final draft.

**Competing interests**

The authors declare that they have no conflict of interest.

**References**


28


Table 1: Canopy compartment and its water flows included in ISIMIP2b Global Water Models

**Canopy water storage** ($S_{ca}$) *(Table S3)*: CLM4.5, CLM5.0, CWatM, DBH, JULES-W1, LPJmL, mHM, MATSIRO, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, WAYS.

**Inflows:**
- total precipitation ($P_{tot}$) (sum of rainfall and snowfall, as input data): CWatM, DBH, JULES-W, LPJmL, mHM, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, WAYS.
- precipitation intercepted by canopy storage: CLM4.5, CLM5.0, MATSIRO.

**Outflows:**
- evaporation of the water intercepted by canopy storage or interception loss or canopy evaporation ($E_{ca}$): CLM4.5, CLM5.0, CWatM, DBH, JULES-W1, LPJmL, mHM, MATSIRO, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, WAYS.
- throughfall ($P_{th}$): CLM4.5, CLM5.0, CWatM, DBH, JULES-W1, LPJmL, MATSIRO, mHM, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, WAYS.

Table 2: Snow and soil compartments and their water flows included in ISIMIP2b Global Water Models

**Snow storage** ($S_{sn}$) *(Table S8)*: CLM4.5, CLM5.0, CWatM, DBH, H08, JULES-W1, LPJmL, Mac-PDM.20, mHM, MATSIRO, MPI-HM, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, WAYS.

- snow held on the canopy ($S_{soc}$): CLM5.0, DBH, JULES-W1, MATSIRO, VIC.
- snow under the canopy ($S_{suc}$): CLM4.5, CLM5.0, DBH, JULES-W1, MATSIRO, ORCHIDEE, VIC.

**Inflows:**
- total precipitation ($P_{tot}$): CWatM, DBH, JULES-W1, WaterGAP2.
- snowfall ($P_{sn}$): CLM4.5, CLM5.0, H08, Mac-PDM.20, MATSIRO, mHM, MPI-HM, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, WAYS.
- throughfall ($P_{th}$): LPJmL.

**Outflows:**
- sublimation ($E_{sn}$): CLM4.5, CLM5.0, CWatM, DBH, H08, JULES-W1, LPJmL, MATSIRO, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, WAYS.
- snowmelt ($M$): CLM4.5, CLM5.0, CWatM, H08, JULES-W1, Mac-PDM.20, mHM, MATSIRO, MPI-HM, ORCHIDEE, PCR-GLOBWB, WaterGAP2, WAYS.

**Soil storage** ($S_{so}$) *(Table S14)*: CLM4.5, CLM5.0, CWatM, DBH, H08, JULES-W1, LPJmL, Mac-PDM.20, mHM, MATSIRO, MPI-HM, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, WAYS.

**Inflows:**
- total precipitation ($P_{tot}$): Mac-PDM.20
- infiltration ($R_{in}$): CWatM, DBH, H08, LPJmL, mHM, MPI-HM, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2.
- throughfall ($P_{th}$): H08, JULES-W1, WAYS.
- snowmelt ($M$): H08, JULES-W1, Mac-PDM.20, WAYS.
- capillary rise ($R_{cr}$): CWatM.

**Outflows:**
- transpiration ($T$): CLM4.5, CLM5.0, CWatM, DBH, LPJmL, MPI-HM, PCR-GLOBWB, VIC.
- evaporation from soil ($E_{so}$): CWatM, DBH, H08, JULES-W1, LPJmL, Mac-PDM.20, mHM, MPI-HM, ORCHIDEE, PCR-GLOBWB, VIC, WaterGAP2, WAYS.
- surface runoff ($R_{su}$): LPJmL, JULES-W1, Mac-PDM.20, ORCHIDEE.
- total runoff ($R_{tot}$): WaterGAP2, WAYS.
- interflow ($R_{if}$): CWatM, JULES-W1, LPJmL, PCR-GLOBWB.
- percolation ($R_{pe}$): MPI-HM.
- groundwater recharge ($R_{gw}$): CWatM, DBH, LPJmL, PCR-GLOBWB.
- groundwater runoff ($R_{gw}$): VIC.
### Table 3: Groundwater, lake, reservoir, and wetland compartments and their water flows included in ISIMIP2b Global Water Models

**Groundwater storage (Sgw) (Table S26):** CLM4.5, CLM5.0, CWatM, H08, Mac-PDM.20, mHM, MATSIRO, MPI-HM, PCR-GLOBWB, WaterGAP2, WAYS.

**Inflows:**
- groundwater recharge ($R_{gw}$): CLM4.5, CLM5.0, H08, Mac-PDM.20, mHM, MPI-HM, WaterGAP2, WAYS.
- percolation ($R_{pe}$): CWatM, PCR-GLOBWB.
- preferential flow ($Q_{pf}$): CWatM.

**Outflows:**
- capillary rise ($R_{cr}$): CWatM, PCR-GLOBWB.
- groundwater runoff ($R_{gw}$): CLM4.5, CLM5.0, CWatM, H08, Mac-PDM.20, mHM, MPI-HM, PCR-GLOBWB, WaterGAP2, WAYS.
- total human water abstraction ($A_{tot}$): H08.

**Lake (Sla) (Table S29):** CLM4.5, CLM5.0, CWatM, LPJmL, PCR-GLOBWB, WaterGAP2.

**Inflows:**
- precipitation ($P_{tot}$): LPJmL, WaterGAP2
- inflow from upstream surface water bodies ($Q_{in}$): LPJmL, WaterGAP2
- groundwater runoff ($R_{gw}$): WaterGAP2
- return flow from human water use ($A_{rf}$): WaterGAP2
- water abstraction for human purposes: LPJmL

**Outflows:**
- evaporation from lake ($E_{la}$): LPJmL, PCR-GLOBWB, WaterGAP2
- outflow from lake ($Q_{la}$): CWatM, LPJmL, PCR-GLOBWB, WaterGAP2
- groundwater recharge below surface water bodies ($R_{gw}$swb): WaterGAP2
- return flow from human water use ($A_{rf}$): LPJmL, MATSIRO, PCR-GLOBWB, WaterGAP2
- water abstraction for human water use from lake ($A_{la}$): WaterGAP2, LPJmL

**Reservoir storage (Sre) (Table S32):** DBH, H08, LPJmL, MATSIRO, PCR-GLOBWB, WaterGAP2.

**Inflows:**
- precipitation ($P_{tot}$): WaterGAP2, LPJmL
- inflow from upstream surface water bodies ($Q_{in}$): CWatM, PCR-GLOBWB, WaterGAP2.
- total runoff ($R_{tot}$): H08, MATSIRO
- groundwater recharge below surface water bodies ($R_{gw}$swb): WaterGAP2
- return flow from human water use ($A_{rf}$): LPJmL, MATSIRO, PCR-GLOBWB, WaterGAP2

**Outflows:**
- evaporation from reservoir ($E_{re}$): WaterGAP2, CLM4.5, CLM5.0, LPJmL, VIC.
- outflow from reservoir ($Q_{re}$): DBH, H08, LPJmL, MATSIRO, PCR-GLOBWB, WaterGAP2.
- groundwater recharge ($R_{gw}$): WaterGAP2.
- water abstraction for human water use from reservoir ($A_{re}$): LPJmL, H08, MATSIRO, PCR-GLOBWB, WaterGAP2.

**Wetland storage (Swe) (Table S36):** MPI-HM, WaterGAP2.

**Inflows:**
- precipitation ($P$): MPI-HM, WaterGAP2
- inflow from upstream surface water bodies ($Q_{in}$): MPI-HM, WaterGAP2

**Outflows:**
- groundwater recharge ($R_{gw}$): WaterGAP2
- evaporation from wetland ($E_{we}$): MPI-HM, WaterGAP2
- outflow from wetland ($Q_{we}$): MPI-HM, WaterGAP2
Table 4: River compartment and its water flows included in the ISIMIP2b Global water Models

<table>
<thead>
<tr>
<th>River storage ($S_r$) (Table S40):</th>
<th>CLM4.5, CLM5.0, CWatM, DBH, H08, LPJmL, mHM, MATSIRO, MPI-HM, ORCHIDEE, PCR-GLOBWB, WaterGAP2, WAYS.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inflows:</strong></td>
<td></td>
</tr>
<tr>
<td>- inflow from upstream surface water bodies ($Q_{iu}$): CLM4.5, CLM5.0, CWatM, H08, LPJmL, mHM, MATSIRO, PCR-GLOBWB, WaterGAP2.</td>
<td></td>
</tr>
<tr>
<td>- total runoff ($R_{n}$): mHM</td>
<td></td>
</tr>
<tr>
<td>- surface runoff or overland flow or fast runoff ($R_{su}$): CLM4.5, CLM5.0, CWatM, ORCHIDEE, PCR-GLOBWB, WaterGAP2.</td>
<td></td>
</tr>
<tr>
<td>- interflow ($R_{if}$): CWatM, PCR-GLOBWB</td>
<td></td>
</tr>
<tr>
<td>- groundwater runoff ($R_{gw}$): CLM4.5, CLM5.0, CWatM, H08, mHM, MATSIRO, MPI-HM, ORCHIDEE, PCR-GLOBWB, WaterGAP2.</td>
<td></td>
</tr>
<tr>
<td>- return flow from human water use ($A_{ri}$): WaterGAP2.</td>
<td></td>
</tr>
<tr>
<td>- streamflow ($Q_{ri}$): H08, MPI-HM</td>
<td></td>
</tr>
<tr>
<td><strong>Outflows:</strong></td>
<td></td>
</tr>
<tr>
<td>- streamflow or outflow or river discharge ($Q_{o}$): CLM4.5, CLM5.0, LPJmL, mHM, MPI-HM, WaterGAP2</td>
<td></td>
</tr>
<tr>
<td>- inflow upstream of a grid cell ($Q_{ui}$): H08</td>
<td></td>
</tr>
<tr>
<td>- mean total annual inflow in a lake ($Q_{mi}$): LPJmL</td>
<td></td>
</tr>
<tr>
<td>- outflow downstream of a grid cell ($Q_{oi}$): CWatM, PCR-GLOBWB</td>
<td></td>
</tr>
<tr>
<td>- water abstraction for irrigation ($A_{irr}$): LPJmL,</td>
<td></td>
</tr>
<tr>
<td>- water abstraction for irrigation from surface water bodies ($A_{irr}^{sw}$): CWatM, PCR-GLOBWB</td>
<td></td>
</tr>
<tr>
<td>- water abstraction for domestic sector from surface water bodies ($A_{dom}^{sw}$): CWatM, PCR-GLOBWB</td>
<td></td>
</tr>
<tr>
<td>- water abstraction for livestock from surface water bodies ($A_{liv}^{sw}$): CWatM, PCR-GLOBWB</td>
<td></td>
</tr>
<tr>
<td>- water abstraction for manufacturing from surface water bodies ($A_{man}^{sw}$): CWatM, PCR-GLOBWB</td>
<td></td>
</tr>
<tr>
<td>- water abstraction for human water use from river ($A_{ri}$): WaterGAP2, H08, MATSIRO</td>
<td></td>
</tr>
<tr>
<td>- water abstraction for irrigation sector ($A_{irr}$): LPJmL</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Human water use sectors estimated by ISIMIP2b GWMs

<table>
<thead>
<tr>
<th>Human water use sectors (A) (Tables S40-S80):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrigation ($A_{irr}$): CLM4.5, CLM5.0, CWatM, H08, LPJmL, MATSIRO, MPI-HM, PCR-GLOBWB, WaterGAP2.</td>
</tr>
<tr>
<td>Domestic ($A_{dom}$): MATSIRO, PCR-GLOBWB, WaterGAP2, CWatM</td>
</tr>
<tr>
<td>Manufacturing ($A_{man}$): MATSIRO, PCR-GLOBWB, WaterGAP2, CWatM</td>
</tr>
<tr>
<td>Electricity ($A_{ele}$): PCR-GLOBWB, WaterGAP2, CWatM</td>
</tr>
<tr>
<td>Livestock ($A_{liv}$): CWatM, PCR-GLOBWB, WaterGAP2.</td>
</tr>
</tbody>
</table>
Table 6: Key characteristics of the Global Water Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Model Type</th>
<th>Temporal resolution</th>
<th>Discretization Type</th>
<th>Calibration / Ability to calibrate / Details</th>
<th>Human water use sectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLM4.5</td>
<td>LSM</td>
<td>6 hours</td>
<td>grid, subgrid for vegetation, surface runoff, and evapotranspiration</td>
<td>no / no, adjustment of some parameters according to vegetation or soil properties / not available</td>
<td>sim $A_{irr}$</td>
</tr>
<tr>
<td>CLM5.0</td>
<td>LSM</td>
<td>6 hours</td>
<td>grid, subgrid for vegetation, surface runoff, and evapotranspiration</td>
<td>no / yes / calibration performed in a Bayesian framework based on sequential Monte Carlo</td>
<td>sim $A_{irr}$</td>
</tr>
<tr>
<td>CWatM</td>
<td>GHM</td>
<td>1 day</td>
<td>grid, subgrid for land cover, snow</td>
<td>yes, calibrated for 12 catchment / monthly or daily discharge / hydrological calibration uses DEAP (Burek et al., 2020)</td>
<td>sim $A_{irr}$, $A_{dom}$, $A_{ind}$, $A_{liv}$</td>
</tr>
<tr>
<td>DBH</td>
<td>LSM</td>
<td>1 day</td>
<td>grid</td>
<td>no / no hydrological calibration, adjustment of some parameters according to vegetation or soil properties / most parameters derived from satellite data.</td>
<td>not included</td>
</tr>
<tr>
<td>H08</td>
<td>GHM</td>
<td>1 day</td>
<td>grid</td>
<td>no / can be calibrated but generally done at the regional scale / the model can be applied at the global or regional scale</td>
<td>sim $A_{irr}$ and $A_{ocean}$, $A_{dom}$, $A_{ind}$, $A_{liv}$</td>
</tr>
<tr>
<td>JULES-W1</td>
<td>LSM</td>
<td>1 day</td>
<td>grid</td>
<td>biophysical processes are calibrated / no hydrologic calibration / ✗</td>
<td>not included</td>
</tr>
<tr>
<td>LPJmL</td>
<td>DGVM</td>
<td>1 day</td>
<td>grid</td>
<td>yield calibration to match FAO stats / no hydrological calibration</td>
<td>sim $A_{irr}$, ISIMIP2b prescribed $A_{dom}$ and $A_{ind}$ not included</td>
</tr>
<tr>
<td>Mac-PDM.20</td>
<td>GHM</td>
<td>1 day</td>
<td>grid</td>
<td>no / yes / calibration uses a 100,000 GLUE ensemble with WATCH Forcing Data (Smith, 2016)</td>
<td>sim $A_{irr}$, ISIMIP2b prescribed $A_{dom}$ and $A_{ind}$ not included</td>
</tr>
<tr>
<td>MATSIRO</td>
<td>LSM</td>
<td>1 hour</td>
<td>grid</td>
<td>no / yes / adjustment of some parameters according to vegetation or soil properties, no calibration capability in TRIP model for routing discharge</td>
<td>not included</td>
</tr>
<tr>
<td>mHM</td>
<td>GHM</td>
<td>1 day</td>
<td>grid</td>
<td>yes / yes / calibration is performed against observed daily discharge GRDC stations, gridded fields of TWS and gridded ET from FLUXNET with the ERA5 climate forcing</td>
<td>not included</td>
</tr>
<tr>
<td>MPI-HM</td>
<td>GHM</td>
<td>1 day</td>
<td>grid, subgrid for surface runoff and evapotranspiration</td>
<td>no / ✗ / ✗</td>
<td>sim $A_{irr}$</td>
</tr>
<tr>
<td>ORCHIDEE</td>
<td>LSM</td>
<td>30 min</td>
<td>grid</td>
<td>no / yes / adjustment of some parameters</td>
<td>not included</td>
</tr>
<tr>
<td>PCR-GLOBWB</td>
<td>GHM</td>
<td>1 day</td>
<td>grid, subgrid for vegetation, land cover</td>
<td>no / yes / adjustment of some parameters</td>
<td>sim $A_{irr}$, $A_{dom}$, $A_{ind}$, $A_{liv}$, $A_{ocean}$</td>
</tr>
<tr>
<td>VIC</td>
<td>GHM</td>
<td>1 day</td>
<td>grid</td>
<td>no calibration for ISIMIP2b</td>
<td>not included</td>
</tr>
<tr>
<td>WaterGAP2</td>
<td>GHM</td>
<td>1 day</td>
<td>grid</td>
<td>yes / mean annual discharge / Beta function, 1319 GRDC stations</td>
<td>sim $A_{irr}$, $A_{dom}$, $A_{ind}$, $A_{ele}$, $A_{liv}$ not included</td>
</tr>
<tr>
<td>WAYS</td>
<td>GHM</td>
<td>1 day</td>
<td>grid</td>
<td>yes / yes / calibrated against the ISLSCP, Initiative II UNH or GRDC composite monthly runoff data (Fekete et al., 2011) from 1986 to 1995 at a 0.5° resolution</td>
<td>not included</td>
</tr>
</tbody>
</table>

Legend: ✗ = no details; DEAP = Distributed Evolutionary Algorithms in Python; DGVM = dynamic global vegetation model; EB = energy balance; GHM = global hydrological model; GRDC = Global Runoff Data Centre; ISLSCP = International Satellite Land Surface Climatology Project; LSM = land surface model; sim = simulated by the model; UNH = University of New Hampshire; $A_{irr}$ = water abstractions for irrigation; $A_{dom}$ = water abstractions for domestic; $A_{liv}$ = water abstractions for livestock; TRIP = Total Runoff Integrating Pathways; **Bold** = LSMs; *Italic* = GHMs; Underline = DGVMs.
Table 7: Representation of the water storages and water flows included in the Global Water Models – PART I

<table>
<thead>
<tr>
<th>Model</th>
<th>Interception scheme</th>
<th>Vegetation scheme Partition / Photosynthesis scheme</th>
<th>(Potential) evapotranspiration scheme</th>
<th>Number of soil layers</th>
<th>Soil scheme Soil layer depth SLD [m]</th>
<th>Snow scheme Snow accumulation and snowmelt</th>
<th>Snow acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLM4.5</td>
<td>f(LAI, SAI)</td>
<td>tile approach with 24 PFTs (including 10 crop types) / *; CO₂</td>
<td>Mon-in-Obukhov Similarity Theory computes only AET</td>
<td>15</td>
<td>depth at layer interface: 0.0175, 0.0451, 0.0906, 0.1655, 0.2891, 0.4929, 0.8289, 1.3828, 2.2961, 3.8019, 6.2845, 10.3775, 17.1259, 28.2520, 42.1032, 49.556</td>
<td>42.1</td>
<td>physically based / mechanistic snow module</td>
</tr>
<tr>
<td>CLM5.0</td>
<td>f(LAI, SAI)</td>
<td>naturally vegetated surfaces are comprised of up to 14 possible plant functional types (PFTs) / dynamic global vegetation model (DGVM); CO₂</td>
<td>Mon-in-Obukhov Similarity Theory computes only AET</td>
<td>25, f(depth to bedrock)</td>
<td>depth at layer interface: 0.020; 0.060; 0.120; 0.200; 0.320; 0.480; 0.680; 0.920; 1.200; 1.520; 1.880; 2.280; 2.720; 3.260; 3.900; 4.640; 5.480; 6.420; 7.460; 8.600; 10.990; 15.666; 23.301; 34.441; 49.556</td>
<td>49.6</td>
<td>physically based snow module</td>
</tr>
<tr>
<td>CWatM</td>
<td>f(veg)</td>
<td>subgrid</td>
<td>Penman-Monteith</td>
<td>3</td>
<td>0.05; 0.05; 0.3; 0.3; 1.7 depends on HWSD data from 1.5 to 3.5m; top layer = 0.020m; root layer = 1.0 to 1.5m.</td>
<td>2.0</td>
<td>Degree-day Method Energy Balance Method</td>
</tr>
<tr>
<td>DBH</td>
<td>f(LAI)</td>
<td>prescribed, 10 vegetation types (PFTs) with fixed vegetation characteristics / *</td>
<td>Energy balance model with Monin-Obukhov similarity theory computes only AET</td>
<td>3</td>
<td></td>
<td>3.5</td>
<td>Degree-day Method Energy Balance Method</td>
</tr>
<tr>
<td>H08</td>
<td></td>
<td>tile approach / *</td>
<td>Bulk, Bulk transfer coefficient set to 0.003</td>
<td>1 / RCZ</td>
<td>1; 1; 1; 1; 1; 1; 1; 1; 1; 1</td>
<td>1</td>
<td>Energy Balance Method Energy Balance Method</td>
</tr>
<tr>
<td>JULES-W1</td>
<td>f(LAI)</td>
<td>5 static vegetation types (PFTs) with fixed plant characteristic / *</td>
<td>Penman-Monteith</td>
<td>4</td>
<td>0.10; 0.25; 0.65; 2.00</td>
<td>3.0</td>
<td>Energy Balance Method Energy Balance Method</td>
</tr>
<tr>
<td>LPJmL</td>
<td>f(LAI)</td>
<td>9 PFTs /L, W, S / DVPNV; CO₂</td>
<td>Priestley-Taylor modified for transpiration</td>
<td>5+1</td>
<td>0.20; 0.30; 0.50; 1; 1m. 1 thermally active soil of 10m</td>
<td>13</td>
<td>Degree-day Method with precipitation factor Energy Balance Method</td>
</tr>
<tr>
<td>Mac-PDM.20</td>
<td>f(veg)</td>
<td>prescribed, 16 PFTs with fixed vegetation characteristics / *</td>
<td>Penman-Monteith</td>
<td>1</td>
<td>none</td>
<td>none</td>
<td>Energy Balance Method</td>
</tr>
<tr>
<td>MATSIRO</td>
<td>f(LAI)</td>
<td>11 static vegetation types with fixed characteristics (PFTs) / *</td>
<td>Mon-in-Obukhov Similarity Theory, to compute only actual evapotranspiration</td>
<td>13</td>
<td>0.05; 0.2; 0.75; 0.75; 1; 1; 1; 90m.</td>
<td>100</td>
<td>Energy Balance Method</td>
</tr>
</tbody>
</table>

Legend: AET = actual evapotranspiration; CO₂ = CO₂ fertilization effect; DGVM = dynamic global vegetation model; DVPNV = dynamic vegetation composition on potential natural vegetation areas; f(LAI) = function of leaf area index; f(LAI, SAI) = function of leaf area index (LAI) and stem area index (SAI); f(veg) = function of vegetation type; HWSD = Harmonized World Soil Database (FAO et al., 2012: http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/); L= light; PFTs = Plant functional types; RCZ = runoff properties varies with climate zones; SLD = soil layers depth from top to bottom; TSD = total soil layer depth; Snow acc = snow accumulation; S = space; W = water; * = not included in the model; Bold = LSMs; Italic = GHMs; Underline = DGVMs.
Table 8: Representation of the water storages and water flows included in the Global Water Models – PART II

<table>
<thead>
<tr>
<th>Model</th>
<th>Interception scheme</th>
<th>Vegetation scheme (Potential) evapotranspiration scheme</th>
<th>Number of soil layers</th>
<th>Soil scheme Soil layer depth</th>
<th>TSD [m]</th>
<th>Snow scheme Snow accumulation and snowmelt</th>
<th>Snow acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>mHM</td>
<td>f(veg)</td>
<td>3 major vegetation classes: (forest, impervious, pervious) for parameter regionalization + long-term dynamics based on LAI-based on GIMMS</td>
<td>6</td>
<td>soil layers correspond to SoilGrids250 vertical discretization, i.e.: 0-5cm, 5-15cm, 15-30cm, 30-50cm, 50-100cm, 100-200cm</td>
<td>2.0</td>
<td>Degree-day Method 1 layer</td>
<td></td>
</tr>
<tr>
<td>MPI-HM</td>
<td>×</td>
<td>prescribed, Land Surface Parameter dataset 2 / × Penman-Monteith reference Evapotranspiration</td>
<td>1, f(FC) none</td>
<td>none</td>
<td>none</td>
<td>Degree-day Method 1 layer</td>
<td></td>
</tr>
<tr>
<td>ORCHIDEE</td>
<td>f(LAI)</td>
<td>tile approach with 17 vegetation types (PFTs); CO₂ Penman &amp; Monteith (Monteith, 1965) based on the correction term developed by Chris Milly (1992) Hamon</td>
<td>11</td>
<td>0.001; 0.003; 0.006; 0.012; 0.023; 0.047; 0.094; 0.188; 0.375; 0.750; 0.500.</td>
<td>2</td>
<td>Energy Balance Method 3 layers</td>
<td></td>
</tr>
<tr>
<td>PCR-GLOBWB</td>
<td>f(veg)</td>
<td>natural vegetation (short and tall vegetation) and agriculture (rainfed, rice irrigated and non-rice irrigated) prescribed annually by HYDE dataset and MIRCA, GLOBCOV/ ×</td>
<td>2</td>
<td>variable up from 0 to 0.3 (first layer) and variable from 0.3 to 1.5 (second layer)</td>
<td>1.5</td>
<td>Degree-day Method 1 layer</td>
<td></td>
</tr>
<tr>
<td>VIC</td>
<td>f(veg)</td>
<td>any number of vegetation types with fixed characteristics can be represented (PFTs) / × Penman-Monteith</td>
<td>3</td>
<td>variable, first layer is fixed to 0.1-0.3m, second and third layers are calibrated</td>
<td>6.15</td>
<td>Energy Balance Method variable</td>
<td></td>
</tr>
<tr>
<td>WaterGAP2</td>
<td>f(LAI)</td>
<td>LAI development model based on T and P / × Priestley-Taylor with varying alpha-values for arid and humid areas</td>
<td>1</td>
<td>from 0.1 to 4m</td>
<td>4</td>
<td>Degree-day Method SG</td>
<td></td>
</tr>
<tr>
<td>WAYS</td>
<td>f(LAI)</td>
<td>14 static vegetation types (PFTs) with fixed characteristics / × Penman-Monteith</td>
<td>1/the complete root zone</td>
<td>variable, derived separately from remote sensing data</td>
<td>variable</td>
<td>Degree-day Method 1 layer</td>
<td></td>
</tr>
</tbody>
</table>

Legend: AET = actual evapotranspiration; CO₂ = CO₂ fertilization effect; f(LAI) = function of leaf area index; f(veg) = function of vegetation type; f(FC) = function of field capacity; P = precipitation; PFTs = Plant functional types; SG = subgrid; SLD = soil layers depth from top to bottom; TSD = total soil layer depth; Snow acc = snow accumulation; T = subgrid temperature (daily average) (0°C); × = not included in the model; Bold = LSMs; Italic = GHMs.
Table 9: Representation of the water storages and water fluxes included in the Global Water Models – PART III

<table>
<thead>
<tr>
<th>Model</th>
<th>Groundwater scheme / groundwater layer</th>
<th>Runoff generation scheme / surface runoff / subsurface runoff</th>
<th>River scheme / River routing(^1) / flow velocity(^2) / floodplain scheme / Details</th>
<th>Reservoir scheme / reservoir operation / Number / Details</th>
<th>Lakes scheme / Details</th>
<th>Wetlands scheme / Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLM4.5</td>
<td>✓ / 1</td>
<td>TOPMODEL(^1) Beven and Kirkby, 1979 / R(<em>{\text{lin}}) R(</em>{\text{sat}}) / f(gw)</td>
<td>✓ / River Transport Model (RTM) / 0.35 m(^{-1}) / diagnostic tool, conserves water globally</td>
<td>× / MOSART based on Manning's equation / × / MOSART based on kinematic wave method</td>
<td>× / / 3663, HydroLakes(^1) / retrospective: following H08: Hansaki et al. (2018) and Wisser et al. (2010), × / / ×</td>
<td>× / Modifed Puls</td>
</tr>
<tr>
<td>CLM5.0</td>
<td>✓ / 1</td>
<td>TOPMODEL(^1) Beven and Kirkby, 1979 / R(<em>{\text{lin}}) R(</em>{\text{sat}}) / f(gw)</td>
<td>✓ / Kinematic wave, approximation of the Saint-Venant equation(^1) Chow et al., 1998 / variable Manning-Strickler equation / × / linear storage</td>
<td>× / / 963 global reservoirs and 5824 local reservoirs / retrospective: Hansaki et al., 2006</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>CWatM</td>
<td>✓ / 1</td>
<td>ARNO(^1) Dümenil and Todini, 1992 / R(_{\text{sat}}) / f(soil and gw)</td>
<td>✓ / / 4134, GRanD / retrospective: Biemans et al., 2011</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>DBH</td>
<td>× / ×</td>
<td>Green-Ampt method / R(<em>{\text{sat}}) / × leaky Bucket(^1) Manabe, 1969 / R(</em>{\text{sat}}) / f(soil) / RCZ</td>
<td>× / / / / / / /</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>H08</td>
<td>✓ / 1</td>
<td>TOPMODEL(^1) Beven and Kirkby, 1979 / R(<em>{\text{lin}}) R(</em>{\text{sat}}) / f(gw)</td>
<td>✓ / based on 30° flow drainage direction map (DDM30) / 0.5 m(^{-1}) / × / Rfd</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>JULES-W1</td>
<td>seepage as gw recharge / ×</td>
<td>TOPMODEL(^1) Beven and Kirkby, 1979 / R(<em>{\text{lin}}) R(</em>{\text{sat}}) / f(gw)</td>
<td>× / / / / / / /</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>LPJmL</td>
<td>seepage as gw recharge / ×</td>
<td>TOPMODEL(^1) Beven and Kirkby, 1979 / R(<em>{\text{lin}}) R(</em>{\text{sat}}) / f(gw)</td>
<td>✓ / continuity equation derived from linear reservoir model / 1 m(^{-1}) / × / linear storage buffer; Rfd</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Mac-PDM.20</td>
<td>✓ / 1</td>
<td>Probability Distributed Moisture (PDM)(^1) Moore and Clarke, 1981 / R(<em>{\text{lin}}) R(</em>{\text{sat}}) / f(gw)</td>
<td>× / / / / / / /</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>MATSIRO</td>
<td>✓ Dgws / ~13</td>
<td>TOPMODEL(^1) Beven and Kirkby, 1979 / R(<em>{\text{lin}}) R(</em>{\text{sat}}) / f(soil)</td>
<td>✓ / linear reservoir, TRIP / 0.5 m(^{-1}) / × / × TRIP</td>
<td>× / / 728 global reservoirs and 6134 'local reservoirs' / following H08, retrospective: Pokhrel et al., 2012</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

Legend: ✓ = included in the model; × = not included in the model for ISIMIP2b simulations; Dgws = dynamic groundwater scheme; GRanD = Global Reservoir and Dam database according to Lehner et al., 2011; gw = groundwater; R\(_{\text{lin}}\) = surface runoff; R\(_{\text{sat}}\) = R\(_{\text{sat}}\) modelled as saturation excess overland flow; R\(_{\text{soil}}\) = R\(_{\text{soil}}\) modelled as infiltration excess or Hortonian overland flow; f(gw) = subsurface flow or interflow modelled as a function of groundwater; f(soil) = subsurface flow or interflow modelled as a function of soil moisture (soil); Rfd = the model routes runoff along flow direction; RTM = routing model; TRIP = Total Runoff Integrating Pathways; Bold = LSMs; Italic = GHMs; Underline = DGVMs.

Table 10: Representation of the water storages and water fluxes included in the Global Water Models – PART IV

<table>
<thead>
<tr>
<th>Model</th>
<th>Groundwater scheme / groundwater layer</th>
<th>Runoff generation scheme</th>
<th>River scheme / River routing(^1) / flow velocity(^2) / floodplain scheme / Details</th>
<th>Reservoir scheme / reservoir operation / Number / Details</th>
<th>Lakes scheme / Details</th>
<th>Wetlands scheme / Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>(mHM)</td>
<td>× / 1</td>
<td>HBV(^\text{Bergström,1976}) + VIC (^\text{3 Layers / R}_{\text{sat}}/ f(\text{soil}))</td>
<td>✓ / mesoscale Routing Model with adaptive timestep, spatially varying celerity(^4) / ✓ / ✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>(MPI-HM)</td>
<td>× / 1</td>
<td>ARNO(^\text{Dümenil and Todini,1992)} / (\text{R}_{\text{sat}}/ f(\text{soil}))</td>
<td>✓ / linear reservoir cascade / variable, Manning-Strickler Equation / ✓ / RtMwefp</td>
<td>✓</td>
<td>lake storage is part of the wetland storage</td>
<td>×</td>
</tr>
<tr>
<td>ORCHIDEE</td>
<td>× / ×</td>
<td>SECHIBA(^\text{Ducoudré et al, 1993)} / (\text{R}_{\text{dso}}/ f(\text{soil}))</td>
<td>✓ / STN, 30p river network / variable, Manning-Strickler Equation / ✓ / wetlands act as floodplains</td>
<td>×</td>
<td>×</td>
<td>✓ / wetlands act as floodplains</td>
</tr>
<tr>
<td>(PCR-GLOBWB)</td>
<td>✓ / 1</td>
<td>ARNO(^\text{Dümenil and Todini,1992)} / (\text{R}_{\text{sat}}/ f(\text{soil and gw}))</td>
<td>✓ / travel time routing (characteristic distance) linked with dynamic reservoir operation / variable based on channel dimension and Manning-Strickler Equation / ✓ / ✓</td>
<td>✓ / ✓ / 6862: GRanD / prospective; Wada et al., 2014</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>VIC</td>
<td>× / × / ✓ / seepage as gw recharge and gw runoff / ✓</td>
<td>XIANJJIANG(^\text{Zhao, 1980)} / (\text{R}_{\text{sat}}/ f(\text{soil}))</td>
<td>✓ / ✓ / ✓ / ✓ / ✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>WaterGAP(^2)</td>
<td>✓ / 1</td>
<td>HBV(^\text{Bergström,1976)} / (\text{R}_{\text{sat}}/ \text{Beta function} / ✓)</td>
<td>✓ / linear reservoir cascade / variable, Manning-Strickler Equation / ✓ / ✓</td>
<td>✓ / ✓ / 11097: GRanD / retrospective, following H08: Döll et al, 2009</td>
<td>✓ / local and global lakes(^4)</td>
<td>✓ / local and global wetlands(^5)</td>
</tr>
<tr>
<td>WAYS</td>
<td>✓ / 1</td>
<td>Bucket(^\text{Manabe,1969)} / Beta function(^1) / (f(\text{soil}))</td>
<td>× / ✓ / ✓ / ✓ / ✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

**Legend:** ✓ = included in the model; × = not included in the model for ISIMIP2b simulations; GRanD = Global Reservoir and Dam database according to Lehner et al., 2011; gw = groundwater; \(\text{R}_{\text{sat}}\) = surface runoff; \(\text{R}_{\text{dso}}\) = \(\text{R}_{\text{sat}}\) modelled as saturation excess overland flow; \(\text{R}_{\text{dso}}\) = \(\text{R}_{\text{sat}}\) modelled as infiltration excess or Hortonian overland flow; \(f(gw)\) = subsurface flow or interflow modelled as a function of soil moisture (soil); RtMwefp = routing model with wetlands (we) and floodplain (fp) scheme; **Bold** = LSMs; **Italic** = GHMs.

**Notes:** 1: Data source: [www.isimip.org](http://www.isimip.org). 2: Zhao et al., 2017. 3: Thober et al, 2019; 4 and 5: [WaterGAP\(^2\), Döll et al., 2012.](http://www.isimip.org)
Figure 1: Scheme with number of models that compute water storage compartments and water flows in ISIMIP2b.

Legend: DDM = degree-day method; EBM = energy balance method; $E_{sa}$ = sublimation; $E_a$ = soil evaporation; $M$ = snowmelt; $P_I$ = total precipitation falls directly to the ground; $P_{in}$ = precipitation intercepted by canopy; $P_{in}$ = snowfall; $P_{th}$ = throughfall; $P_{tot}$ = total precipitation; $P_{can}$ = snowfall that is affected by the canopy interception and dripping; $R_{c}$ = capillary rise; $R_{gr}$ = groundwater runoff; $R_{gw}$ = groundwater recharge; $R_{in}$ = infiltration; $R_{s}$ = surface runoff; $T$ = transpiration; $\theta$ = air temperature; $\theta_{mf}$ = snow freeze temperature; $\theta_{mt}$ = melting temperature. **Bold:** number of models that compute water flows and water storage compartments in ISIMIP2b, **Blue:** water flow, **Orange:** evaporation.

Figure 1: Number of global water models that compute lateral water balance in ISIMIP2b.

Legend: $E_{la}$ = evaporation from lake; $E_{re}$ = evaporation from reservoir; $E_{ri}$ = evaporation from river; $E_{we}$ = evaporation from wetland; $Q_{in,up}$ = inflow from upstream cell for reservoir storage; $Q_{in,we,up}$ = inflow from upstream cell for wetland storage; $Q_{la}$ = outflow from lake; $Q_{re}$ = outflow from reservoir; $Q_{ri}$ = streamflow, $Q_{we}$ = outflow from wetland. **Bold:** number of models that compute lateral water balance in ISIMIP2b, **Blue:** water flow, **Orange:** evaporation.
Figure 3: Number of global water models that consider water source for human water use sectors in ISIMIP2b

Figure 4: Number of global water models that consider return flow destination in ISIMIP2b
Figure 5: Number of water storage compartments and water flows included in the ISIMIP2b global water models

Figure 6: Number of human water use sectors and related water flows included in the ISIMIP2b global water models.
Table 11: Code availability of the ISIMIP2b Global water models

<table>
<thead>
<tr>
<th>Model Abbreviation</th>
<th>Code availability</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLM4.5</td>
<td>CLM4.5 is under active development by the University Corporation for Atmospheric Research (UCAR) - National Center for Atmospheric Research (NCAR; <a href="https://ncar.ucar.edu">https://ncar.ucar.edu</a>). The model version is licensed under CC BY 4.0. The exact version of the model, used to produce the results of this paper, is archived on Zenodo (Thiery, 2020).</td>
<td>Oleson and Lawrence, 2013</td>
</tr>
<tr>
<td>CLM5.0</td>
<td>CLM5.0 is under active development by the University Corporation for Atmospheric Research - National Center for Atmospheric Research and hosted at the National Center for Atmospheric Research (NCAR; <a href="https://ncar.ucar.edu">https://ncar.ucar.edu</a>). The version of model is licensed under CC BY 4.0. The exact version of the model, used to produce the results of this paper, is archived on Zenodo (CTSM Development Team, 2020).</td>
<td>Lawrence et al., 2019</td>
</tr>
<tr>
<td>CWaM</td>
<td>CWaM is under active development funded by the International Institute for Applied Systems Analysis (IIASA, Austria; <a href="http://www.iiasa.ac.at/cwam">http://www.iiasa.ac.at/cwam</a>). CWaM is open source and available online via GNU General Public License v3. The code can be used on different platforms (Unix, Linux, Window, Mac) and is provided through a GitHub repository <a href="https://github.com/cwam/cwam">https://github.com/cwam/cwam</a>. The version of the model used to produce the results in this paper are stored as version 1.04 in the GitHub repository (<a href="https://github.com/CWaM/CWaM">https://github.com/CWaM/CWaM</a>) and at Zenodo (Burek et al., 2019).</td>
<td>Burek et al., 2020 Burek et al., 2019</td>
</tr>
<tr>
<td>DBH</td>
<td>DBH is under active development by the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China (<a href="http://english.igsnrr.cas.cn/">http://english.igsnrr.cas.cn/</a>). The exact version of the model (global version 1), used to produce the results of this paper, is not open source. It is only available by request to the editors / reviewers in charge of this paper.</td>
<td>Tang et al., 2006; Liu et al., 2016</td>
</tr>
<tr>
<td>H08</td>
<td>H08 is under active development by the National Institute for Environmental Studies, Japan (Japan; <a href="http://www.nies.go.jp/index-e.html">http://www.nies.go.jp/index-e.html</a>; <a href="http://h08.nies.go.jp">http://h08.nies.go.jp</a>). H08 is open source and available online via <a href="http://h08.nies.go.jp">http://h08.nies.go.jp</a>. The version of model is licensed under the terms and conditions: <a href="https://h08.nies.go.jp/h08/files/licence_en.pdf">https://h08.nies.go.jp/h08/files/licence_en.pdf</a>. The version of model is licensed under CC BY 4.0. The exact version of the model (version 20190101), used to produce the results of this paper, is archived on Zenodo (Hanasaki, 2020).</td>
<td>Hanasaki et al., 2006; Hanasaki et al., 2008a,b; Hanasaki et al., 2018.</td>
</tr>
<tr>
<td>JULES-W1</td>
<td>JULES (the Joint UK Land Environment Simulator) is a community land surface model under continuous development by a wide community of UK researchers, coordinated by UKMO and CEH. The exact version of the model (version 4.7) used in these simulations is available from the Met Office Science Repository Service (registration required) at <a href="https://code.metoffice.gov.uk/trac/jules">https://code.metoffice.gov.uk/trac/jules</a>. To access the code a freely available non-commercial research license is required (<a href="https://jules-lsm.github.io">https://jules-lsm.github.io</a>).</td>
<td>Best et al., 2011; Clark et al., 2011</td>
</tr>
<tr>
<td>LPJmL</td>
<td>LPJmL is under active development funded by the Potsdam Institute for Climate Research (Germany; <a href="https://www.pik-potsdam.de/en/home">https://www.pik-potsdam.de/en/home</a>). The exact version of the model (model version 3.5), used to produce the results of this paper, is not open source. It is only available by request to the editors / reviewers in charge of this paper.</td>
<td>Gerten, 2004; Bondeau et al., 2007; Rost et al., 2008; Biemans et al., 2011</td>
</tr>
<tr>
<td>Mac-PDM.20</td>
<td>Mac-PDM.20 is under active development by the University of Nottingham (UK; <a href="https://www.nottingham.ac.uk">https://www.nottingham.ac.uk</a>) and the University of Reading (UK; <a href="https://www.reading.ac.uk/">https://www.reading.ac.uk/</a>). The version of the model (version 20), used in ISIMIP2b and in this paper, is not open source as it under active development. It is only available by request to the editors / reviewers in charge of this paper.</td>
<td>Gosling and Arnell, 2011; Smith, 2016</td>
</tr>
<tr>
<td>MATSIRO</td>
<td>MATSIRO is under active development by the University of Tokyo (Japan; <a href="https://www.u-tokyo.ac.jp/en/index.html">https://www.u-tokyo.ac.jp/en/index.html</a>) and National Institute for Environmental Studies (Japan; <a href="http://www.nies.go.jp/index-e.html">http://www.nies.go.jp/index-e.html</a>). The exact version of the model (model version MIROC-INTEGI), used to produce the results of this paper, is not open source. It is only available by request to the editors / reviewers in charge of this paper.</td>
<td>Takata et al., 2003; Pokhrel et al., 2012; 2015</td>
</tr>
<tr>
<td>mHM</td>
<td>mHM is under active development funded by the Helmholtz Centre for Environmental Research – UFZ (Germany; <a href="https://www.ufz.de/index.php?en=33573">https://www.ufz.de/index.php?en=33573</a>; <a href="https://git.ufz.de/mhm">https://git.ufz.de/mhm</a>). The version of model is licensed under GNU General Public License v3: <a href="https://git.ufz.de/mhm/-/blob/develop/LICENSE">https://git.ufz.de/mhm/-/blob/develop/LICENSE</a>. The exact version of the model (model version 5.10), used to produce the results of this paper, is archived on Zenodo (Samaniego et al., 2017).</td>
<td>Samaniego, 2017; Samaniego et al., 2010; Kumar et al., 2013; Thober et al., 2019</td>
</tr>
<tr>
<td>MPI-HM</td>
<td>MPI-HM was developed at the Max Planck Institute for Meteorology (Germany; <a href="https://mpimet.mpg.de/en/homepage">https://mpimet.mpg.de/en/homepage</a>). The exact version of the model (model version 1.2), used to produce the results of this paper, is not open source. It is only available by request to the editors / reviewers in charge of this paper.</td>
<td>Stacke and Hagemann, 2012</td>
</tr>
</tbody>
</table>
The source code for ORCHIDEE- MICT version 8.4.1 is available online, but its access is restricted. Consequently, one is required to communicate with the corresponding author for a username and password. The source code can be found at the following address: https://forge.ipsl.jussieu.fr/orchidee/browser/branches/ ORCHIDEE-MICT/tags/ORCHIDEE_MICT_8.4.1 The exact version of the model (model version v8.4.1), used to produce the results of this paper, is not open source. It is only available by request to the editors / reviewers in charge of this paper.

PCR-GLOBWB  
PCR-GLOBWB is under active development funded by the Utrecht University (The Netherlands; https://www.uu.nl/en/research/department-of-physical-geography). PCR-GLOBWB is open source and available online via: https://github.com/UU-Hydro/PCR-GLOBWB_model. The version of model is licensed under GNU General Public License v3. The exact version of the model (model version 2.0), used to produce the results of this paper, is archived on Zenodo: https://doi.org/10.5281/zenodo.1045338 (Sutanudjaja et al., 2017).

VIC  
VIC is under active development funded by the University of Washington, (USA; https://vic.readthedocs.io/en/master/). It was applied by the Indian Institute of Technology Gandhinagar, Gandhinagar (India; http://www.iitgn.ac.in/). VIC is open source and available online via https://github.com/UW-Hydro/VIC. The version of model is licensed under GNU General Public License v2.0. The exact version of the model (model version 4.1.2.g), used to produce the results of this paper, is archived on Zenodo (Shah and Vimal, 2020).

WaterGAP2  
WaterGAP2 is under active development funded by the Goethe University Frankfurt (https://www.goethe-university-frankfurt.de/en?legacy_request=1; https://www.uni-frankfurt.de/45218063/WaterGAP) and Kassel University (https://www.uni-kassel.de/uni/) (Germany). The exact version of the model (model version 2.2c), used to produce the results of this paper, is not open source due to licensing issues. It is only available by request to the editors / reviewers.

WAYS  
WAYS is under active development funded by the Southern University of Science and Technology – SUSTech (China; https://www.sustech.edu.cn). The version of model is licensed under Creative Commons Attribution 4.0 International. The exact version of the model used to produce the results used in this paper is archived on Zenodo (Mao and Liu, 2019).

Van Beek et al., 2011; Wada et al., 2011; Wada et al., 2014; Sutanudjaja et al., 2018

Gao et al., 2009

Döll et al., 2012, 2014; Portmann et al., 2010; Müller Schmied et al., 2014, 2016; Verzano et al., 2012; Flörke et al., 2013

Mao and Liu, 2019