Advances in Water Resources 143 (2020) 103682

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Modeling the surface water and groundwater budgets of the US using MODFLOW-OWHM

Mustafa H. Alattar[a], Tara J. Troy[a,b,*], Tess A. Russo[c], Scott E. Boyce[d]

[a] Department of Civil and Environmental Engineering, Lehigh University, Bethlehem, PA 18015, USA
[b] Department of Civil Engineering, University of Victoria, Victoria, BC V8W 2Y2, Canada
[c] Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA
[d] U.S. Geological Survey, California Water Science Center, 4165 Spruance Rd., Suite 200, San Diego, CA 92101-0812, USA

ARTICLE INFO

Keywords:
MODFLOW-OWHM
Surface-groundwater interactions
Water budgets
Climate variability
Evapotranspiration
Streamflow

ABSTRACT

Assessments of groundwater and surface water budgets at a large scale, such as the contiguous United States, often separately analyze the complex dynamics linking the surface and subsurface categories of water resources. These dynamics include recharge and groundwater contributions to streamflow. The time-varying simulation of these complex hydrologic dynamics, across large spatial and temporal scales, remains a scientific challenge due to the complexity of the processes and data availability. In this study, groundwater fluxes and surface hydrologic processes are simulated across the contiguous US for 1950-2010. The simulation estimates the monthly water budget components, such as groundwater recharge, surface runoff, and evapotranspiration; streamflow in major rivers is routed while accounting for groundwater exchange. Human impacts are included through groundwater pumping, and climate variability is included, including variability in precipitation, temperature and potential evapotranspiration. The simulated groundwater level and river discharge have strong correlation with USGS observation wells and streamflow gages, with R^2 values of 0.992 and 0.946, respectively. The simulated evapotranspiration is compared with three other published estimation methods, showing that it is able to capture the magnitude and seasonality of evapotranspiration over the Mississippi River basin. As such, the model is able to reasonably simulate the surface and groundwater budgets over the US, allowing for questions of the relative importance of climate and human impacts to be explored in the future.

1. Introduction

Surface water and groundwater interact with one another, with surface hydrologic processes impacting groundwater recharge and groundwater levels determining baseflow in streams. The water budgets are dynamically affected by climate conditions, such as precipitation and temperature; human activities, such as water withdrawals from rivers, groundwater pumping, and land use change; and terrestrial processes such as plant water use. Estimating the surface and groundwater budgets is critical for quantifying water resources across large spatial and temporal scales, yet it remains a challenging task due to a lack of in situ observations of critical hydrologic processes and poor characterization of subsurface hydrogeologic properties. This study presents a contiguous US (CONUS) set-up of the MODFLOW-One-Water Hydrologic Model Version 2 (MF-OWHM2) (Hanson et al., 2014a; Boyce et al., 2020), which is a surface and groundwater model capable of simulating the hydrologic fluxes and storages for water budget assessment.

Climate, subsurface conditions, and human activities all impact groundwater levels. For example, during periods of low rainfall, groundwater levels decline due to a reduction in groundwater storage from low recharge rates (Wada et al., 2014) if the lateral groundwater flow is limited (de Graaf et al 2017, Condon and Maxwell, 2017). Furthermore, excessive groundwater pumping can threaten groundwater sources with high groundwater depletion (Döll et al., 2014b; Scanlon et al., 2012). In some regions, such as the Great Plains, groundwater levels increase slowly due to low recharge rate and relatively low hydraulic conductivity (Peterson et al., 2016). The sensitivity analysis of groundwater level is analyzed on global scale with respect to hydraulic conductivity, groundwater recharge, and surface water body elevation by Reinecke et al., (2019). To maintain the groundwater availability and baseflow to rivers, some global studies (Döll et al., 2014a; Pokhrel et al., 2012; Wada et al., 2010) analyzed the groundwater recharge rate and groundwater storage from water budget components; and other studies (Peterson et al., 2016; Faunt et al., 2009) analyzed the streamflow in regional aquifer systems.

* Corresponding author.
E-mail address: tjtroy@uvic.ca (T.J. Troy).

https://doi.org/10.1016/j.advwatres.2020.103682
Received 5 December 2019; Received in revised form 7 July 2020; Accepted 7 July 2020
Available online 8 July 2020
0309-1708/© 2020 Published by Elsevier Ltd.

Although it is widely accepted that understanding surface and groundwater fluxes and storages is critical for water resources planning, collecting extensive in situ observations, such as groundwater levels and subsurface geological properties, is prohibitively expensive over large regions. Remote sensing data can provide estimates of hydrologic fluxes over large spatial scales, but remote sensing has relatively short time series and can typically only observe surface processes. The exception is GRACE (Tapley et al., 2004), which can estimate groundwater storage change through gravity field changes, but only at coarse spatial and temporal resolutions. To overcome the remote sensing limitations numerical models can be utilized to estimate and reconstruct the hydrologic fluxes and stores, thereby overcoming the observational data limitations. These models can be useful tools to address critical questions: for instance, what is the impact of climate variability and groundwater pumping on groundwater availability.

Regional, hydrological and groundwater, models have been used to simulate numerically the surface water and groundwater hydrologic processes. For example, three regional models have been set up in detail over three major aquifers in the United States in Central Valley, High Plains, and Rio Grande aquifer systems. The first groundwater flow model simulates the Northern High Plains aquifer (Peterson et al., 2016) and uses MODFLOW with the Newton-Rhapson solver (Niswonger et al., 2011). The second is the Central Valley Hydrologic Model CVHM (Faunt et al., 2009) for California's Central Valley aquifer using MODFLOW-FMP2 (Schmid and Hanson, 2009). The forthcoming new version of CVHM (CVHM2) is being updated to use MF-OWHM2. A third model is the Rio Grande transboundary integrated hydrologic model (RGTIHM) that simulates portions of New Mexico, Texas, and Mexico (Hanson et al., 2020) using MF-OWHM2. These regional models provide a better understanding of the surface and subsurface regional water budgets and their dynamics. However, regional models typically cover a single aquifer, neglecting processes at the model boundaries. This means the streamflow, surface runoff, and lateral groundwater flow may be neglected at the boundaries of these regional models. These assumptions can misrepresent the modeled water budgets (Schaller and Fan, 2009; Krakauer et al., 2014), depending on the regional conditions such as topography, geology, and climate.

Continental and global scale groundwater models were developed (de Graaf et al., 2015; Fan et al., 2007; Maxwell et al., 2015) to not only solve the problem of lateral flow at model boundaries, but because they can provide a better understanding of the hydrology of an entire large-scale system. For example, large scale models can analyze the spatial variation of climate, geology, and topography and its effect on surface water and groundwater availability, and they can analyze the surface-groundwater interaction in different aquifers.

One of the first studies to simulate the groundwater depth across the contiguous United States was conducted by Fan et al. (2007), where the water table was estimated as the equilibrium of long-term climatic forcing on groundwater level under a steady state model, followed by a study of global observations of water table depth (Fan et al., 2013). These studies led to the ability to connect groundwater with other hydrologic fluxes, which has been investigated in several later studies over the United States. For example, Maxwell et al. (2015) used an integrated groundwater model (ParFlow, Maxwell and Miller, 2005) to analyze the surface and subsurface flow system over the majority of CONUS. To do a high spatial resolution run, the model was run under steady state conditions but can be run transiently as in Kollett (2009) and Maxwell et al. (2016). Other studies simulate the groundwater on the global scale using a model run at steady state (de Graaf et al., 2015) and transient (de Graaf et al., 2017), then analyzing the groundwater depletion (de Graaf et al., 2019). These models simulate recharge and river discharge from the global hydrological model PCR-GLOBWB (Wada et al., 2011), then simulates the groundwater lateral flow from MODFLOW. The MODFLOW One-Water Hydrologic Model (MF-OWHM2) used in this paper simulates both surface hydrologic processes and groundwater flow simultaneously in one model, overcoming the limitation of coupling two models.

Some of these large scale models assume the magnitude and direction of the fluxes are constant over time (Fan et al., 2007; Maxwell et al., 2015; de Graaf et al., 2015), thereby neglecting the temporal dynamics of seasonal and interannual climate variability. In addition, human activity, such as groundwater pumping, is often neglected. However, in recent studies (de Graaf et al., 2017; Condon et al., 2019; de Graaf et al., 2019) the groundwater pumping is simulated. In addition, incorporating streamflow routing is still a challenging component to model with a coupled groundwater model because of modeling limitations or the scarcity of needed data such as river hydraulic conductance over such a vast area such as the United States. These challenges have led some past studies to simplify the modeling of such a complex system by neglecting human activity and simplifying or neglecting the streamflow routing and its interaction with groundwater.

This study's objectives are 1) to evaluate the feasibility of using MF-OWHM2 at continental scales for water budget estimation and 2) to develop and validate a CONUS-wide MF-OWHM2 model set-up that explicitly simulates hydrologic processes that link the surface and groundwater hydrologic processes and human impacts. The focus is on capturing large-scale hydrologic patterns and human impacts to provide a baseline model for the hydrologic community, which can then expand the model to include more detailed processes relevant for specific research questions. This model is complementary to other US-focused modeling efforts, such as that of ParFlow, as it is more focused on including human impacts in a parsimonious framework that would not require high performance computing resources. In this paper, we first present the model set-up, including simplifications, the observation-based input data, and the assumptions made in some input variables. Then, the model validation and water budget results are presented, including both groundwater and surface water. The conclusions are presented in the final section.

2. Model and data

To simulate the surface and subsurface hydrology of the contiguous US (CONUS), this study uses MODFLOW-One-Water Hydrologic Model Version 2 (MF-OWHM2) (Boyce et al., 2020; Boyce, 2020), which is a modular modeling software developed by the U.S. Geological Survey and U.S. Bureau of Reclamation. MF-OWHM2 builds on the widely-used MODFLOW model (Harbaugh et al., 2000, Harbaugh, 2005), which simulates groundwater fluxes at a range of spatial and temporal scales. In addition to simulating groundwater fluxes, the standard version of MODFLOW includes pumping and surface water routing (Prudic et al., 2004). MF-OWHM2 incorporates a surface water model, allowing for simulation of the partitioning of precipitation into infiltration and surface runoff, evapotranspiration, and irrigation. MF-OWHM2 is able to holistically simulate the hydrologic system, including some of the ways human's impact hydrologic processes. It also allows for better quantifying surface water and groundwater availability by directly coupling the surface and subsurface hydrologic fluxes. Table 1 compares four large-scale models simulating the lateral groundwater flow.

The model domain covers CONUS, extending from 132.437°W to 72.502°W and 22.177°N to 49.822°N (Fig. 1) and to a depth of ~60 m below the water table as a single vertical layer. The MF-OWHM2 study area is divided into a finite difference equal-area grid of approximately 13km, using the North America Albers Equal Area Conic projection. The domain covers a land area of approximately 11,790,960 km^2, which extends into Canada and Mexico. The ocean boundary conditions are represented as a Cauchy boundary condition using the MODFLOW General Head Boundary (GHB) module that specifies a single average sea level elevation. The simulation time frame is from 1950 to 2010 at a monthly time step, called a stress period in MODFLOW, with boundary conditions specified for each month. For model convergence reasons, each stress period contains two transient sub-steps, where the model breaks up the monthly data into sub-periods. The model is simulated us-

Table 1
Comparison summary of models with lateral groundwater flow simulation.

	This study	Faunt et al. (2009); Peterson et al. (2016); Hanson et al. (2020)	Maxwell et al. (2016)	de Graaf et al. (2017)
Model type	Integrated groundwater and surface water model	Integrated groundwater and surface water model	Groundwater and surface water model	Groundwater and surface water model
Software	MODFLOW-OWHM	MODFLOW-OWHM	ParFlow with optional coupling with CLM	MODFLOW
Model domain	CONUS	Regional aquifers	Majority of CONUS	Global
Spatial resolution	13 km	1.6 km, 1km, 0.27km	1 km	10 km
Time step	monthly with two transient sub-steps	monthly with two transient sub-steps	hourly (can be run transiently)	monthly
Surface runoff	simulated for each grid cell	simulated for each water balance subregion	calculated	Groundwater recharge and streamflow discharge forced with outputs from PCR-GLOBWB
Streamflow routing	routed for major rivers	routed	routed	
Groundwater recharge	simulated with respect to groundwater level	simulated with respect to groundwater level	Coupled with CLM (P-E)	
Evapotranspiration from groundwater	simulated based on the groundwater level	simulated based on the groundwater level	simulated with CLM	
Confined aquifer	not included	included	not included	included
Groundwater pumping	included	included	included in Condon and Maxwell (2019)	included

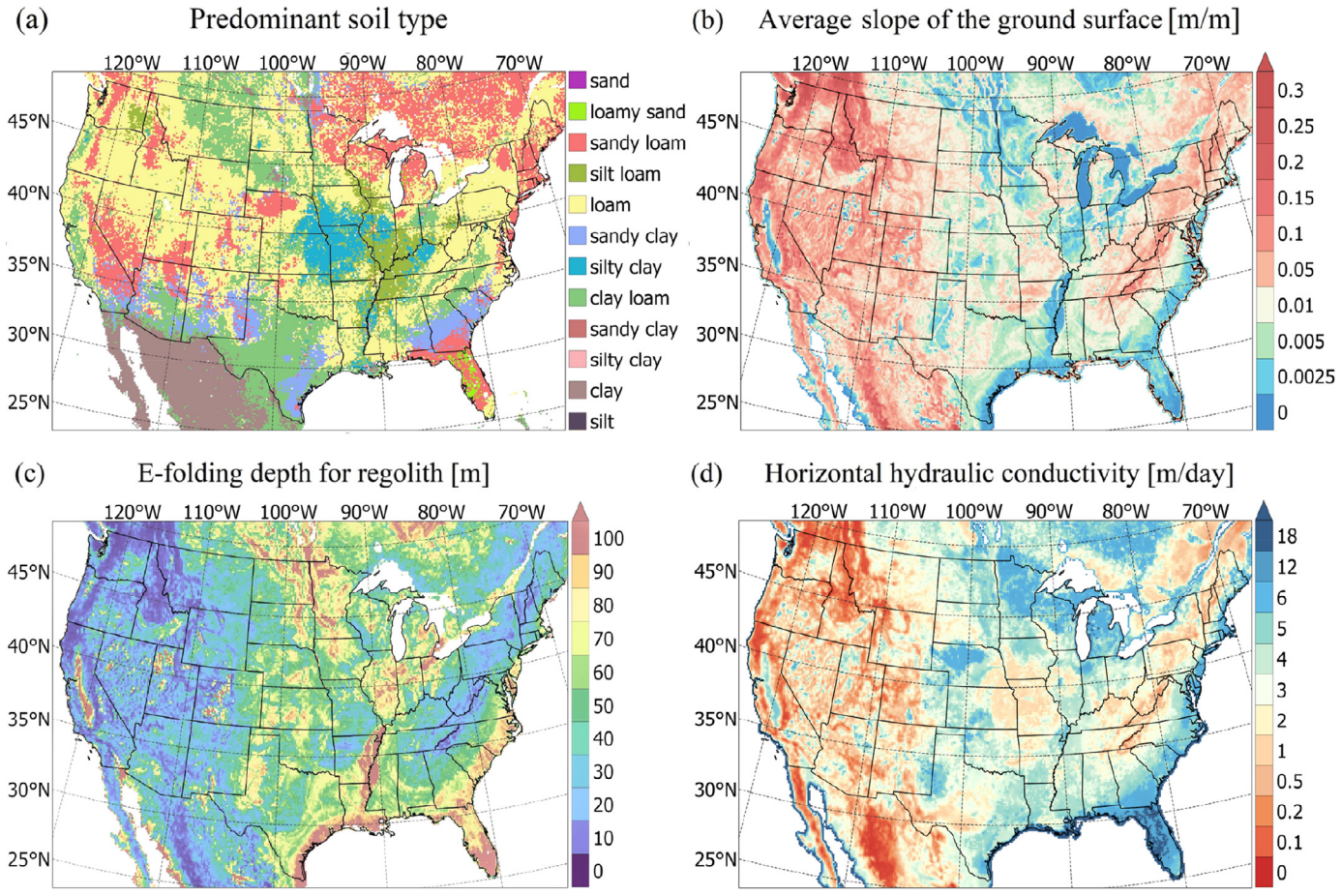

Fig. 1. (a) Predominant soil type at 1.5 m depth using the POLARIS soil map (Chaney et al., 2016), classified into 12 soil types using the USDA classification. (b) Average slope of the ground surface within each MODFLOW grid [m/m]. (c) e-folding depth for the regolith [m]. (d) Depth-averaged horizontal hydraulic conductivity for the 60 m thick unconfined model layer [m/day].

ing MODFLOW's "convertible" option, which uses the unconfined equation when the water table is above the layer bottom and converts to the confined equation when the water table is below the layer bottom of a grid cell. Confined layers are omitted due to insufficient well observations in some confined aquifers to accurately set-up and validate the model across the CONUS domain.

The model empirically estimates the evapotranspiration, based on the monthly precipitation, reference evaporation, crop coefficient, and the groundwater evapotranspiration extinction depth. The crop coefficient is the ratio of the actual evapotranspiration to the reference evapotranspiration for each crop type, and the extinction depth is the depth at which there is no evaporation from groundwater sources. The sur-

face runoff to rivers is estimated based on the surface runoff coefficient, which is the ratio of the surface runoff from the rainfall over the infiltration. Groundwater recharge is estimated as the residual of the water budget. MF-OWHM2 calculates the vertical and the horizontal groundwater flows using a finite difference scheme of the three-dimensional groundwater equations and estimates the interaction between the groundwater and surface water. These fluxes are calculated based on the subsurface properties (Harbaugh, 2005; Prudic et al., 2004), such as the hydraulic conductivity and layer thickness, and the climatic conditions. The unsaturated zone is not simulated in this study, which increases the simulation error when water that infiltrates takes longer than the model timestep to reach the water table. This is common in very dry areas such as the southwestern deserts (Scanlon et al., 2006; Flint et al., 2007) and may result in an over- or underestimate of the groundwater fluxes. Both surface runoff and baseflow from the groundwater to rivers are simulated, and streamflow is routed in the major rivers (Section 2.4). In addition to the effect of climate variability, the model includes the human impact on groundwater levels by simulating withdrawals using the MODFLOW well module and explained (Section 2.3). Details of the model set-up are described below.

2.1. Subsurface hydrogeologic parameters

The ground surface elevation was specified using the HydroSHEDS 30-second resolution digital elevation model (Lehner et al., 2006). Using HydroSHEDS, the average ground surface elevation of each 13km model grid cell was calculated. MODFLOW calculates the layer thickness as the difference between the ground surface elevation and the layer bottom (explained in Section 2.2); this layer thickness is used in the groundwater equations.

The model hydraulic conductivity was estimated using the method derived by Fan et al. (2007) that is based on soil type. The soil type was first characterized into fractions of sand, silt, and clay using the POLARIS soil database (Chaney et al., 2016), and the average for the top 1.5m depth is calculated. Then, the fractions were used to specify in which of the twelve soil types in the United States Department of Agriculture (USDA) soil type classification the grid cell belongs (Fig. 1a). This was then used to estimate the hydraulic conductivity at 1.5 m depth. The vertical hydraulic conductivity is calculated for each soil type from the Land Data Assimilation System (LDAS, http://ldas.gsfc.nasa.gov/); and the horizontal hydraulic conductivity is calculated using the anisotropy ratio estimated by Fan et al. (2007). Using the estimated hydraulic conductivity at 1.5m depth below the ground surface, each grid cell's horizontal hydraulic conductivity was estimated for the full depth of the model layer thickness, extending to 60 m below the groundwater level. This was done using the depth-dependent exponential relationship derived by Fan et al. (2007):

$$K = K_o \exp\left(-\frac{z'}{f}\right), \quad where: \quad f = \begin{cases} \frac{a}{1+b\,\beta}, & for\ \beta \le 0.16 \\ 5m, & for\ \beta > 0.16 \end{cases} \qquad (2)$$

where a and b are constants, z' is the depth where the hydraulic conductivity is calculated, β is the terrain slope, K_o is the hydraulic conductivity at 1.5m depth, and f is the e-folding depth which is the depth interval in which the exponential decline of hydraulic conductivity decreases by a factor of the natural logarithm constant e (e=2.71828). The constants a and b were taken from the best-fit in Fan et al. (2007), where a equals 120 m and b equals 150 m.

To calculate the terrain slope, β, and e-folding depth, f, additional data is required. The terrain slope β is calculated as the average slope of the 1.25 km pixels within the MODFLOW grid cell for both regolith (Fig. 1b) and bedrock. The regolith is a layer of sediment, such as dust, soil, sand, gravel, and loose rock, that covers a hard rock formation. The 1.25 km pixel slopes are calculated from the HydroSHEDS elevation data, first reprojected to the North American Albers Equal Area Conic projection, used in this MODFLOW set-up. The depth to bedrock data was downloaded from (http://www.soilinfo.psu.edu/index.cgi?soil_data&conus&data_cov&dtb&methods) which is available for a depth of 152 cm across the United States (Fan et al., 2007). Since this study focuses on the United States, the depth to the bedrock in Mexico and Canada was assumed to be zero to simplify the model. Because of the equations' form, this has only a minor impact on the estimated hydraulic conductivity. The assumption of 13 km grid size increases the error in the simulation due to the uncertainty related to averaging the depth to bedrock and the ground surface elevation over a relatively large grid cell especially in the heterogeneous terrain such as the southwest region.

The e-folding depth (f) was calculated for both bedrock and regolith at 1.5 m depth and depends on the topography (Fig. 1c). In Eq. (1), the hydraulic conductivity (K) decreases exponentially with depth z'. To calculate a depth-averaged hydraulic conductivity, the horizontal hydraulic conductivity was calculated for every 0.5 m depth interval from the ground surface elevation to the layer bottom of 60 m below water table, and the integrated equivalent horizontal hydraulic conductivity was calculated using the following equation (Todd and Mays, 2005):

$$K_x = \frac{\sum_{i=1}^{n} K_i\, dz_i}{H} \qquad (3)$$

where H is the model layer thickness, and K_x is the final horizontal hydraulic conductivity (Figure 1d) for the model layer thickness, dz_i is the 0.5m depth interval increment, K_i is the horizontal hydraulic conductivity at each 0.5 m interval increment within the model single layer.

As mentioned above, the model consists of one layer with a thickness of ~60m below the estimated observed groundwater level used to initialize the model; details of estimating the initial groundwater level are below. The exception is below rivers, where a minimum thickness of 20m below the riverbed is specified to allow for model convergence. In a few river grid cells, the river segment is below the average 13km layer bottom due to large topographic gradients, and this is not allowed in the model. Impermeable conditions were assumed below the model layer because the hydraulic conductivity values below 60 m are prohibitively small when using the exponentially decreasing hydraulic conductivity method in Eq. (1). The specific yield (Figure S1) was estimated from the aquifer sediment and rock type using Heath (1983) and Morris and Johnson (1967). The aquifer sediment and rock type map is taken from the USGS (Miller, 1990).

2.2. Groundwater observations and model initial conditions

U.S. Geological Survey (USGS) observation well data was used to specify the model initial conditions and to validate the modeled groundwater levels. There are 780,851 wells with at least one observation between January 1950 and December 2010 across the United States. Of these, wells were excluded if they met any of the following conditions: (1) the well was identified by the USGS as confined, mixed aquifer type, or undefined aquifer type, (2) the well elevation is unavailable, (3) the well elevation is below -10 meters to eliminate wells with missing data flags or unrealistic elevation values, (4) the well elevation is greater than 4000 meters above sea level, (4) the well is designated as pumping wells, injection wells, obstructed, damaged, plugged, dried, or flowing, (5) the well has an average observed water depth deeper than 300 m, implying it is in a deep aquifer outside this model's bounds. The latter condition was necessary to provide preliminary initial conditions of the water level. After the model was spun-up, the simulated water level was used to determine the model layer thickness, which is a minimum of 60m below the water table of each grid cell. After the layer thickness was determined, we verified that no wells are below the model's vertical extent and that 300m below the water table was a reasonable assumption for the model layer thickness. Applying these five criteria resulted in 642,839 wells that could be used for initializing the model. This approach was similar to that used in Fan et al. (2007).

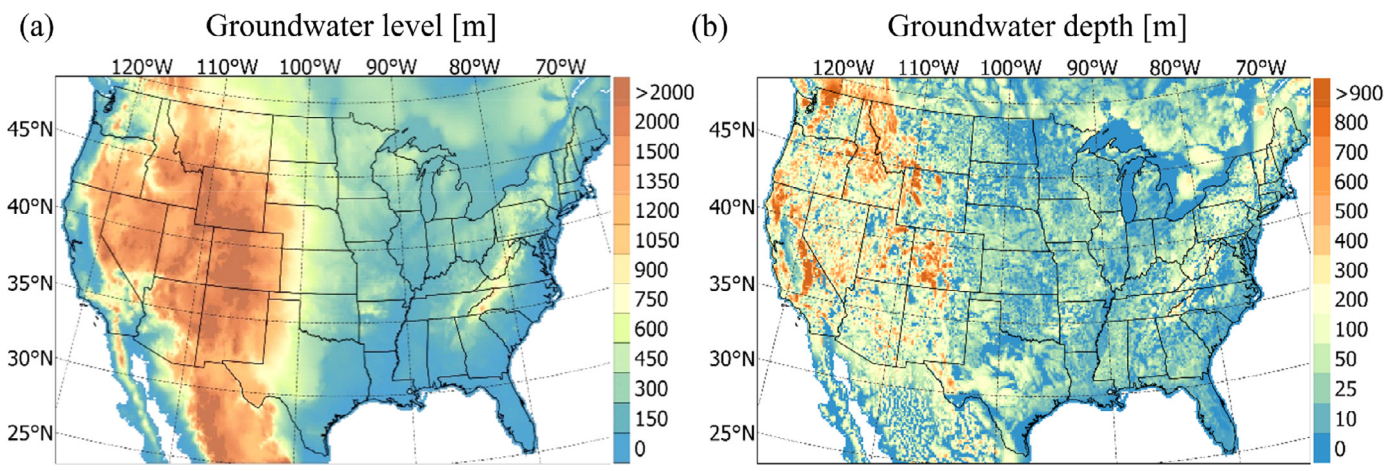

Fig. 2. (a) Interpolated groundwater level [m] above sea level used for the model initial conditions using a triangulation-based natural neighbor interpolation for the 642,839 selected USGS site observations over the United States. (b) Groundwater depth [m] below the ground surface over the United States.

Water depths from the selected wells were utilized to calculate the water level of the model initial conditions for January 1950. Water level observations were spatially interpolated using a triangulation-based natural neighbor interpolation, which is a mix between linear and cubic nonlinear interpolation using the MATLAB function griddata() with the natural interpolation option. The interpolated groundwater level is constrained to not exceed the average ground surface elevation for each grid cell. Because there are a limited number of wells with observations dating back to 1950, a strategy was developed to utilize well observations whenever they exist. For wells with observations beginning after January 1950, the water level for the earliest January observation is used; selecting one month, in this case January, eliminates the impact of seasonal variability on water levels. For wells with only one observation, the water level is assumed to be the same as January 1950. This assumption allowed for the inclusion of more wells, particularly in large areas of the country that were data scarce in the mid-20th century. The impact of this assumption on the model's performance is discussed in the next section. The water level in the ocean and lakes is assumed to be constant over time and is simulated using the General Head Boundary (GHB) module in MF-OWHM2. In the major lakes, the water level was estimated from the HydroSHEDS digital elevation map (Lehner et al., 2006). The ocean boundary condition water level is set equal to mean sea level. The initial conditions for the model consist of the interpolated January 1950 water level combined with the boundary condition for lakes and oceans (Fig. 2) and data for Canada and Mexico were linearly extrapolated and run under a spin up model as explained in detail in Section 3.

2.3. Groundwater pumping

Well pumping was estimated using the USGS Water Use dataset (https://water.usgs.gov/watuse/data/, last accessed on July 18, 2019), which is available on a county scale every five years since 1985. County groundwater withdrawals are uniformly distributed across the model grid cells located within the county. The available USGS groundwater withdrawal data do not specify aquifer type for withdrawals; all pumping, regardless of aquifer source, is assumed to be from the unconfined aquifer. Furthermore, because the USGS groundwater withdrawals data is only available for every five years, the pumping rate is assumed to be temporally constant between the available years. Pre-1985 years are assumed to be the same as 1985 withdrawals. This imposes errors, but in the absence of data for before 1985, it allows for incorporating human influence on the hydrologic cycle. To simplify the pumping simulation, each grid cell in the model contains one effective well in the center, which represents all the pumping wells within the grid cell.

2.4. Surface water routing

Streamflow is simulated using two different modules: the major rivers are routed using the streamflow routing (SFR) module, and the tributaries using the rivers (RIV) module (Fig. 3a). The river network shapefile was downloaded from https://www.naturalearthdata.com, which primarily derived the data from the U S. World Data Bank 2 (2006). They classified the main rivers from the "double-lined rivers" classification level, this data is used for the main rivers in this study. The small rivers and tributaries obtained from the same source for lower river classification levels. The SFR module calculates the water exchange in the river channel between surface water and groundwater and routes the river discharge downstream, accounting for surface runoff and baseflow inputs to the river for each river grid cell. The RIV module only calculates the water flux between the streamflow and baseflow from groundwater without routing, making it computationally more efficient than the SFR module. In the river validation, the baseflow to rivers calculated by the RIV module is added to the rivers simulated by the SFR module for each month. Because the RIV module does not incorporate surface runoff into the streamflow, all the surface runoff is included in the SFR routing.

There are 358 USGS streamflow gages selected for the SFR module that are located on the SFR river network and that have at least 50 years of continuous monthly observations from 1950 to 2010. The SFR river network consists of 5241 grid cells that are divided into 595 segments, where 358 of these segments have a USGS gage station at the start of the segment. Thus, the discharge at the upstream grid cell for each segment is equal to the observed discharge of that USGS gage station. During the simulation, the model SFR module routes the streamflow for each river segment downstream, accounting for interactions between surface water and groundwater via baseflow and for the surface runoff that runs off to the river segments. In this study, infrastructure such as dams and water transfers are neglected.

The model study area was divided into 183 water balance subregions (WBS, Fig. 3b), using the 1:250,000-scale Hydrologic Units (huc250k) (Steeves and Nebert, 1994), and watersheds without a SFR river segment were combined so that each water balance region has a river except Canada and Mexico. This allows the Farm Process (FMP) in MF-OWHM2 to distribute the surface runoff water from each water balance subregion over the number of SFR rivers grid cells located within each one uniformly.

To simplify the SFR model, some assumptions were made. First, due to a paucity of large-scale data on river bed hydraulic conductivity, the river bed hydraulic conductivity was assumed to be equal to the grid cell's horizontal hydraulic conductivity, this used by the model to cal-

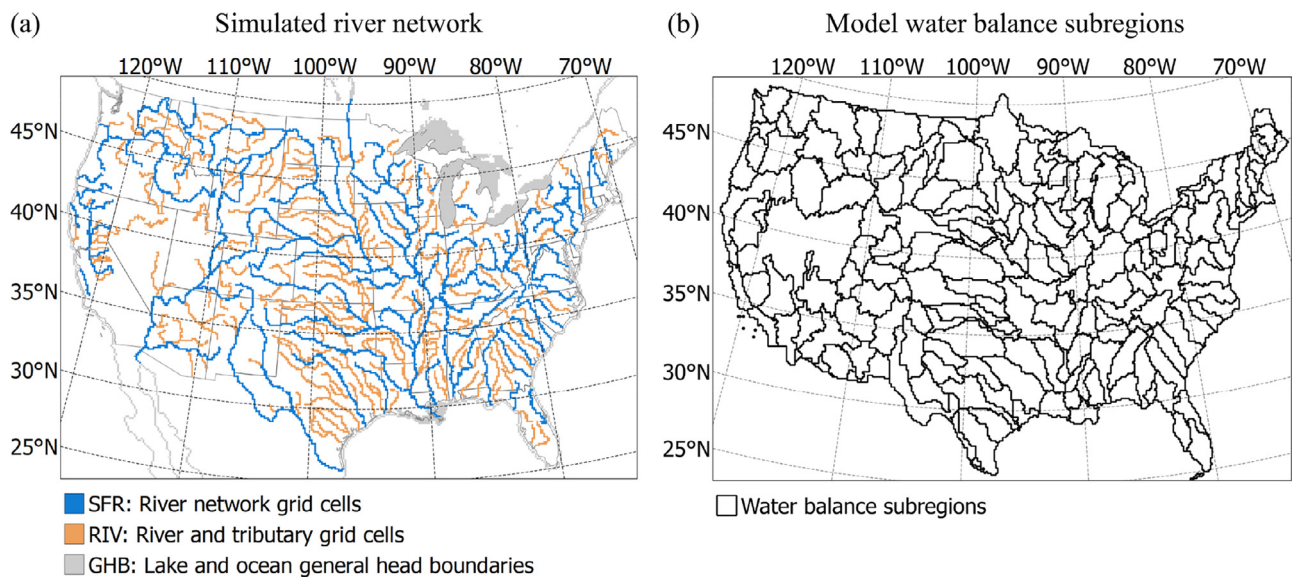

Fig. 3. (a) The simulated river network. Blue grid cells represent rivers simulated using the stream flow routing (SFR) module. Orange grid cells are rivers tributaries simulated using the river (RIV) module. Grey grid cells are lakes and oceans included in the general head boundary (GHB) module. (b) The simulated water balance subregions from Steeves and Nebert (1994), aggregated into 183 regions.

culate the river bed conductance. The river bed thickness is set to 1m, and the Manning's roughness coefficient is 0.04, which is typical for natural streams (Chow, 1959). If a stream gage was missing stream width or elevation data, it was taken from Google Earth or the USGS DEM, respectively. In this model, the stream width and elevation data linearly interpolated for grid cells between the gage stations by the model, this option can be changed to specify the stream width and elevation data for each grid cell. The model allows for precipitation and evaporation from the river water surface; this was assumed to be zero given the large-scale nature of the simulation and the lack of data with which to validate these processes.

The river module (RIV) was used for the smaller river tributaries. This tributary river network covers the United States with 4625 grid cells included in this model. The river elevation for RIV grid cells was estimated from the USGS digital elevation map. The water depth was assumed to be 3m for all RIV tributaries.

2.5. Land cover

Land cover data was taken from USDA's CropScape (USDA National Agricultural Statistics Service Cropland Data Layer, 2010) for 2010 and held constant over the modeling period. For Canada and Mexico, land cover is from MODIS (Channan et al., 2014). The land cover dataset was first aggregated into 23 types using the California Department of Water Resources (2000) classification system, then the dominant land cover within each model grid cell was selected (Fig. 4b). The land cover data is used to specify crop coefficients and calculate potential evapotranspiration. The crop coefficient values were estimated from previous studies (Allen et al., 1998; Faunt et al., 2009; Hanson et al., 2014a; Hanson et al., 2014b; Boyce et al., 2020) (Fig. 4a), and the crop coefficient of non-cropped land cover types is set equal to one by construct. For cropped land cover types, the crop coefficients are constant across the study domain.

2.6. Climate data

Climate conditions, such as precipitation, potential evaporation, and temperature, have a significant impact on water supply and demand for both surface and groundwater. Climate data such as precipitation and temperature are available on daily intervals but in this study climate data are simulated on a monthly time interval called a stress period. In this model, the monthly climate-related input data are precipitation, snowmelt, and reference evapotranspiration, then MF-OWHM2 simulates the impact of climate variability on both surface water and groundwater processes. MF-OWHM2 assumes all precipitation is rainfall, and therefore capable of being partitioned into runoff and infiltration, a portion of which can be consumed as evapotranspiration. However, much of the US has a significant snow season, and this assumption would be inappropriate in the northern and western US. To account for the snowmelt lag, the Variable Infiltration Capacity (VIC) hydrologic model (Liang et al., 1994; Liang et al., 1996) snowmelt outputs were used. VIC is a land surface model that solves for energy and water budget closure at the land surface and includes a snow model. VIC was implemented using the meteorological forcings of Maurer et al. (2002), updated to extend from 1950 to 2010, and the calibrated parameters from Troy et al. (2008). The VIC-simulated liquid input to the soil, essentially rainfall and snowmelt, is then passed to MF-OWHM2 to account for the time lag in precipitation that would occur due to the snowpack. Because MF-OWHM2 is on the North America Albers Equal Area Conic projection and the Maurer and VIC data are on the World Geodetic System of 1984 (WGS84) projection, the VIC and Maurer datasets are interpolated using a bilinear interpolation method.

The monthly reference evaporation was calculated by VIC for natural vegetation and tall and short reference crops, and the minimum potential evaporation for the MF-OWHM2 is constrained to be at least the actual evaporation from VIC model. Finally, MF-OWHM2 calculates the potential evapotranspiration by multiplying the reference evapotranspiration by the crop coefficient for each land cover type. For Canada and Mexico, precipitation and evapotranspiration data were taken from global VIC simulations (Sheffield and Wood, 2008) forced with the dataset of Sheffield et al. (2006). Further details on evapotranspiration calculations are in the following section. In this model, the irrigation withdrawals from streams were neglected and all irrigation water from pumping is assumed to be consumed. Thus, the simulated evapotranspiration is only calculated from the precipitation and groundwater. This may increase the potential simulation errors and neglecting the seasonality variation of the irrigated. In future studies, the irrigation water supply and demand from pumping and streams and its consumptive use could be included.

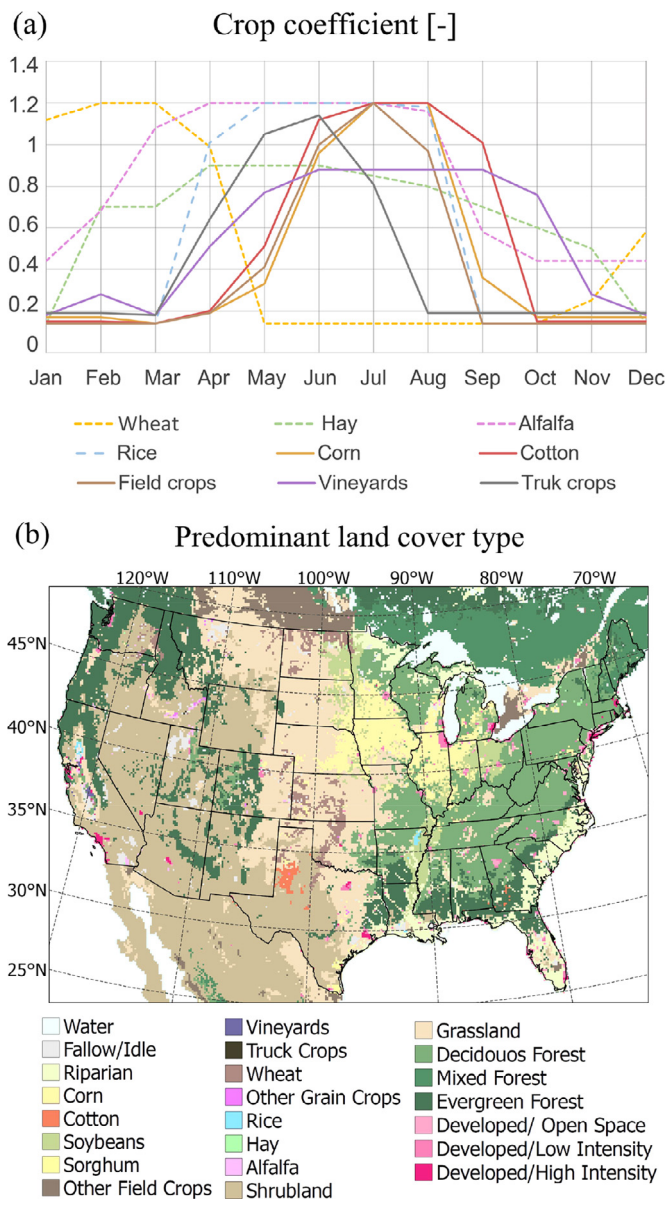

Fig. 4. (a) Crop coefficients for different crop land cover types (Allen et al., 1998; Faunt et al., 2009; Hanson et al., 2014a; Hanson et al., 2014b; Boyce et al., 2020). (b) Predominant land cover type. Over the United States land cover is from the USDA's CropScape (USDA National Agricultural Statistics Service Cropland Data Layer, 2010) (https://nassgeodata.gmu.edu/CropScape/). For Canada and Mexico, land cover is from MODIS (Channan et al., 2014).

3. Model spin-up and calibration

For calibration purposes, a simplified 12-month model—using the seasonal averaged climate variables from 1950 to 2010—was created for the model spin-up and calibrated to estimate the groundwater level in an equilibrium state and to calibrate the extinction depth of groundwater-sourced evapotranspiration using the Farm Process version 4 (FMP) of MODFLOW-OWHM2. The simplified, 12-month model was run with the seasonal climatology, allowing for errors in initial conditions to dissipate. Evaluation of the water levels in the model shows that changes in water levels exponentially decline across several months, such that water levels are stable by the end of the 12-month run. These errors can also be due to uncertainty in the estimated hydraulic conductivity used in this model. Errors in the hydraulic conductivity can lead to incorrect hydraulic gradients and lateral water movement. In addition, hydraulic gradients between rivers and the surrounding topography can lead to inaccurate baseflow estimates until equilibrium is reached. At the end of the 12-month model spin-up, the water levels are used as the initial conditions for both the calibration and the full study period simulations.

For modeling evapotranspiration, it is important to note that MF-OWHM2 does not separately simulate the near-surface soil moisture that is directly accessible to the plants in their rooting zone (Hanson et al., 2014b). Instead, the near-surface soil moisture is included in the groundwater budget as part of the water stored in the unconfined aquifer layer. Thus, the change in both soil moisture and groundwater level is included in the simulated change in groundwater level. In the FMP, evaporation and transpiration are modeled separately. One of evapotranspiration concepts employed in FMP is to linearly increase groundwater consumption as transpiration when the water table plus a capillary fringe length intersects the average root depth until the water table reaches the average root depth. When the water table intersects the average root depth, the transpiration consumption linearly decreases to represent anoxic conditions. Due to the size of this simulation domain, a positive feedback loop occurred when the water table resulted in too much anoxia causing a reduction in transpiration, which further increased the water table elevation and the amount of anoxia. To avoid this feedback loop, all evapotranspiration is modeled as evaporation, which does not include anoxia as part of its simulation. Because crop coefficients are used in conjunction with reference evaporation, this results in reasonable simulations of evapotranspiration within the model structure.

In this model, a linear relationship is used for the subsurface-sourced evapotranspiration calculations, where the calculated evapotranspiration will be equal to the potential evapotranspiration when the groundwater level reaches the ground surface elevation, and subsurface evapotranspiration (from groundwater and near-surface soil moisture) equal to zero when the groundwater level is lower than the groundwater evapotranspiration extinction depth. The model extinction depth was calibrated using the 12-month calibration model with the objective of matching the evapotranspiration calculated by the VIC model. Specifically, the spatial average of the absolute difference between MF-OWHM2 and VIC-simulated ET is minimized. Compared to the VIC-estimated ET, the MF-OWHM2 simulated evapotranspiration, for the 12-month model of the CONUS simulation domain, had an average absolute bias of 1.033 mm/month and a root mean square error of 3.701 mm/month. Because there are no accurate observations of evapotranspiration over the US, we are using VIC-estimated ET to calibrate the constant extinction depth for this model. For the 732-month model, the average absolute bias is 2.66 mm/month and the average root mean square error is 5.46 mm/month. These error statistics demonstrate that the calibration with the simplified model can be used with the full version, with a small increase in errors. This allows for a much faster calibration over such a large domain.

The surface runoff coefficient was calibrated using the full model simulation from 1950 to 2010. This calibration was conducted because sensitivity studies show that the cumulative change in groundwater level is very sensitive to the surface runoff coefficient parameter, as this controls infiltration to the subsurface. The calibration allows for the assumption of negligible change in groundwater storage from 1950 to 2010. The calibration error of groundwater level for the calibrated surface runoff coefficient is calculated as the spatial average of root mean square error, and it is equal to 0.135 m for the CONUS simulation domain. This calibrated surface runoff coefficient was used for the full monthly model simulation from 1950 to 2010, which also included groundwater pumping.

4. Model validation, results, and discussion

4.1. Groundwater validation

The full monthly model simulation from 1950 to 2010 with groundwater pumping is validated against the observed groundwater levels

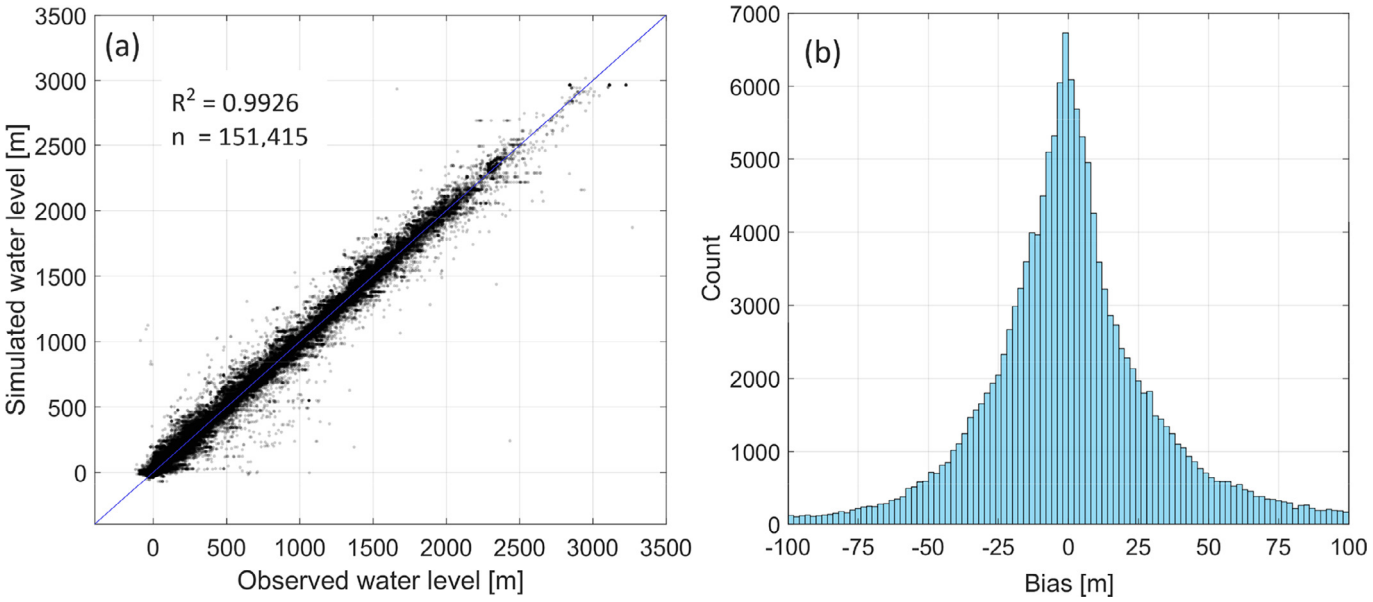

Fig. 5. (a) Validation between the average annual values of the reference observed water level from USGS wells (x-axis) and the model simulated water level (y-axis) from 1950 to 2010, with root mean square value of 3.82m. (b) Histogram of average monthly model bias of simulated water levels [m], the mean bias is 2.28m.

from the selected USGS wells for unconfined aquifers. The validation was implemented between USGS well measurements and the simulated groundwater level in the grid cell that contains that well for the months when a measurement was made. The criteria to select the USGS observation wells for the groundwater level validation is similar to the criteria used for the model initial condition, but wells with only one month of observations were removed. This results in 151,415 USGS wells with time series for the groundwater validation. Of the selected wells, 900 have at least 30 years of data and 16,054 have at least five years of data. Because the well observations have different record lengths, the statistics are calculated for each well for those months with observed groundwater levels. The statistics for grid cells with observations beginning in January 1950 and those grid cells with later data are similar, indicating that the assumption made to include wells in the initial conditions, regardless of data availability at the start of the model period, did not significantly impact the model performance.

Simulated and measured mean annual groundwater levels for 1950 to 2010 across the 151,415 USGS groundwater observation sites have an R^2 value of 0.9926 (Fig. 5a). This is similar to the R^2 of the estimated groundwater level by previous studies. For example, an R^2 value of 0.999 was calculated by Fan et al. (2007), where the water table was estimated as the equilibrium of long-term climatic forcing on groundwater level under a steady state model. Maxwell et al. (2015) simulated the water table using the ParFlow groundwater model (Maxwell et al., 2015) over most of CONUS using a high spatial resolution, steady state run with an R^2 value of 0.998. Another example of simulating the water table is a study by de Graaf and others (2017) using MODFLOW model globally, with R^2 value of 0.94.

Fig. 5b shows the histogram of model bias in groundwater level. The mean bias is 2.28m, with relatively few wells having a large bias. For example, 5.2% of the wells have a negative bias greater than -50m and 9.1% of the wells have a positive bias greater than 50m. The bias results for many reasons. There is uncertainty in model parameters, boundary conditions, and pumping observations. Errors are introduced due to the difference between the observed water level for each individual USGS well and the average groundwater level for wells located on a 13 × 13 km grid cell area. In other words, the elevation of the water level measurement can be very different from the average grid cell elevation. The largest error in groundwater level values occurs in three regions (Fig. 6 and Figures S2-S4): heterogeneous terrain, near rivers, and the regions with high pumping rates. In heterogeneous terrain, the errors occur for two reasons. First, significant errors are introduced due to the topography, as groundwater levels can vary significantly within a 13 × 13km grid cell; the error is calculated as an average throughout the grid cell and is therefore sensitive to where the individual USGS wells are located within a grid cell. The large, 13 km, grid cell may mischaracterize the deep arid zones in the west and increase the simulation error. In addition, this model includes only unconfined aquifer type wells following previous studies (Fan et al., 2007; Maxwell et al., 2015). Groundwater flow in the confined aquifer is neglected. This assumption increases the model uncertainty and misrepresents the aquifers system because groundwater in the confined aquifer has much lower storativity, therefore calibrating to groundwater head will lead to skewed parameterization. Second, there is simulation error due to the uncertainty in model parameters. For example, overestimating hydraulic conductivity (K) allows the groundwater to move too quickly from higher to lower groundwater levels, resulting in a milder hydraulic gradient than the hydraulic gradient of the interpolated USGS observation wells. Near rivers, the errors occur due to the uncertainty in river parameters and due to topography. For example, in the SFR module, the errors result from uncertainty in the river bed hydraulic conductivity, the difference between the groundwater level and water level in streams, and accumulated errors from river routing. Outside of heterogeneous regions and along rivers, the model has a high correlation with the USGS historical observed groundwater levels. In these areas, the model is able to simulate the mean water levels well, with larger errors in simulating the variability (Figure S5).

4.2. River routing and water budget validation

Results of both SFR and RIV modules were post-processed by adding the monthly simulated baseflow between the rivers and groundwater in the RIV module to the SFR module, where the RIV flow within each watershed was uniformly distributed across the SFR grid cells located within the same watershed. The routed streamflow was validated against the observed discharge from the available USGS gage stations

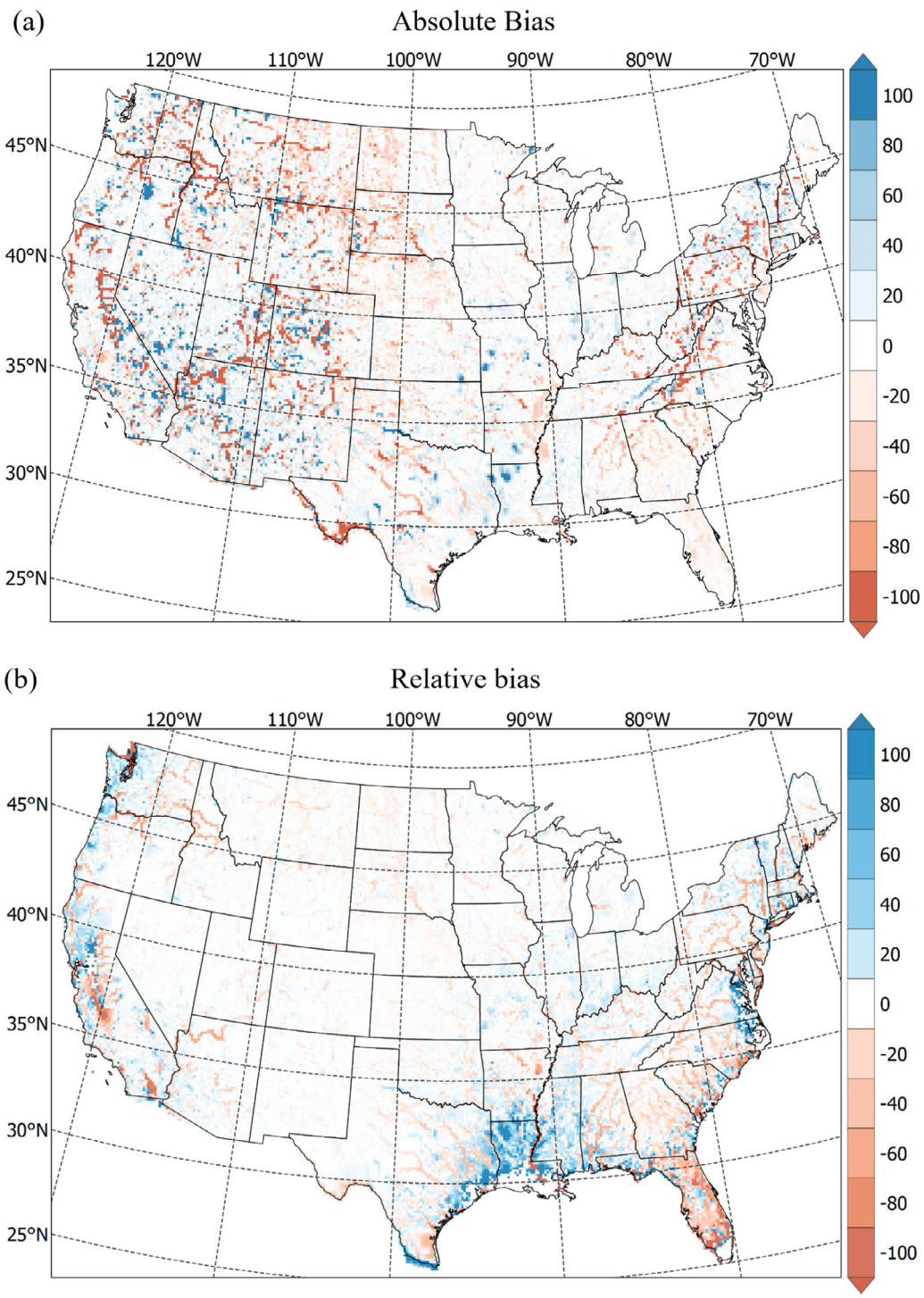

Fig. 6. Absolute bias (m, top) and relative bias (%, bottom) of the time-averaged water level [m] from 1950 to 2010, between the model simulated water level and the interpolated, grid cell averaged observed water level from USGS wells. A positive bias indicates the simulated water level is higher than the measured water level.

located on the SFR river network with at least 50 years of monthly streamflow observations from 1950 to 2010. The simulated routed outflow from each river segment downstream was validated with the USGS gage station located on the upstream of next river segment that connected with it. The USGS observed streamflow is used as the upstream boundary condition for each routed segment. Fig. 7a shows the scatter plot of the average observed versus simulated streamflow on a log scale, and Fig. 7b shows Nash-Sutcliffe Efficiency values which represents how well the model simulates both the mean and variability in streamflow. The results show a strong correlation between the simulated and observed discharge for each SFR river segment, with a R^2 equal to 0.95 for the actual values and 0.87 for the log values. The model overestimates streamflow in smaller basins; this occurs because the current version of the Farm Process (FMP) distributes surface runoff water equally over the SFR river grid cells that are located within the same water balance subregion. Time series of randomly selected gages are plotted in Figure S6. Surface water withdrawals for irrigation are not included in this model. Because USGS observations are used as upstream boundary conditions,

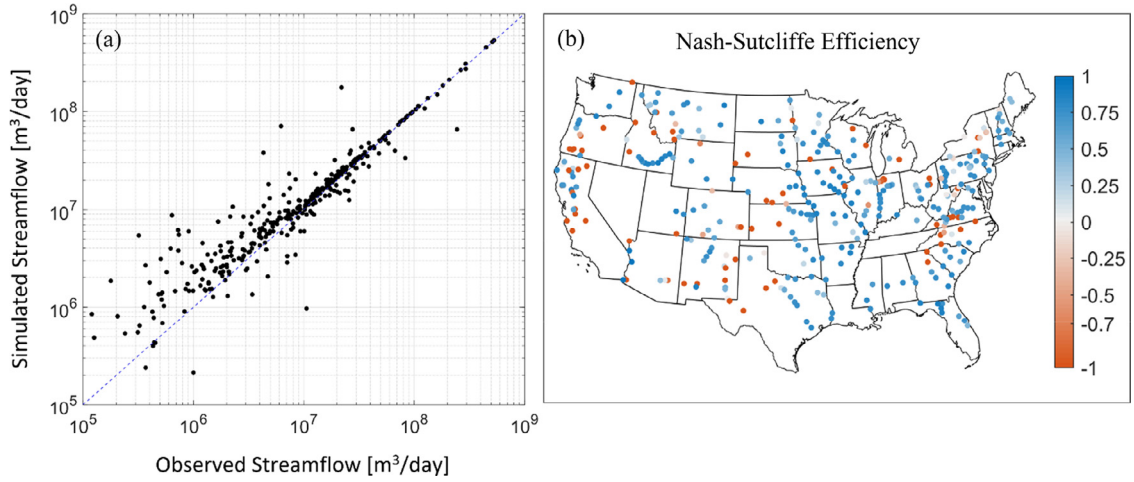

Fig. 7. (a) Average daily observed streamflow (x-axis) and average monthly simulated streamflow (y-axis). Each dot represents one of 358 USGS gage stations located on the SFR river network with at least 50 years of streamflow observations. The R^2 is 0.95 for the actual streamflow values and 0.87 for the log-transformed streamflow values and root mean square error of 0.277 [m^3/day]. (b) Nash-Sutcliffe Efficiency.

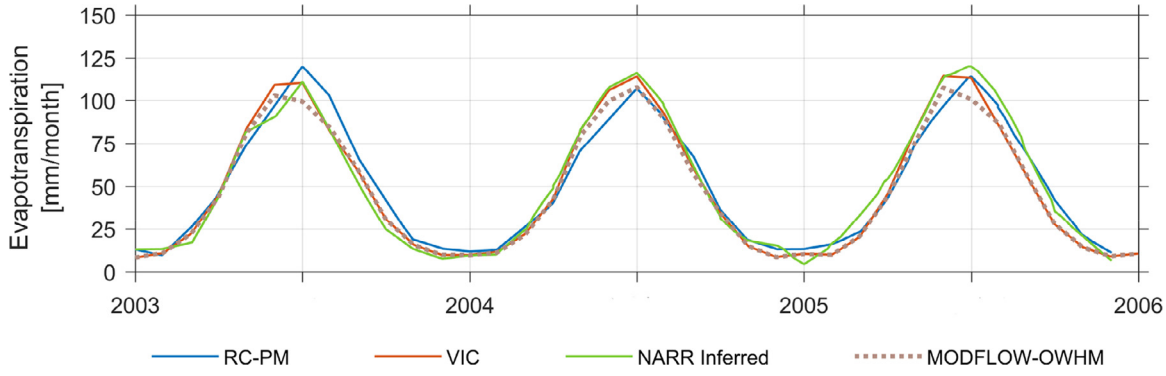

Fig. 8. Monthly evapotranspiration averaged over the Mississippi basin estimated from different methods from 2003 to 2006 [mm/month]. VIC (red) represents the model simulation used as boundary conditions for MF-OWHM2, which simulates evapotranspiration using the Farm Process version 4. The NARR inferred evapotranspiration was calculated by Sheffield et al. (2009) as the residual of the atmospheric water budget and the RC-PM (Ferguson et al., 2008; Sheffield et al., 2009) calculates evapotranspiration using remote sensing inputs with the Penman-Monteith equation.

this will not have a significant impact on the routed streamflow as it is implicitly accounted for in the observations.

Fig. 8 shows the time series of the simulated evapotranspiration against evapotranspiration estimated using different methods over the Mississippi basin presented by Sheffield et al. (2009); the data were digitized from the published paper. The simulated evapotranspiration using MF-OWHM2 is compared with three estimation methods of daily evapotranspiration: (1) Estimation of daily evapotranspiration using remote sensing based on the Penman-Monteith equation (Monteith, 1965). These remote sensing, Penman-Monteith (RC-PM) estimations of evapotranspiration data are the ensemble mean with varying inputs, which are calculated using a revised Penman-Monteith equation from Mu et al. (2007). (2) Evapotranspiration inferred as a residual from the North American Regional Reanalysis (NARR) (Mesinger et al., 2006) atmospheric water budget. (3) Estimation of daily evapotranspiration using the VIC hydrologic model set-up described in Section 2.6. In this MF-OWHM2 set-up, the extinction depth of groundwater evapotranspiration was calibrated to get simulated evapotranspiration similar to the estimated evapotranspiration using this VIC hydrologic model. The root mean square error between the MF-OWHM2 simulated evapotranspiration and the RC-PM, NARR Inferred, and VIC estimations are 8.57, 8.29, and 4.69 mm/month, respectively, over the Mississippi basin from 2003 to 2006. The evapotranspiration from near-surface soil moisture was assumed to be part of the groundwater budget, and both groundwater and near-surface soil moisture changes are accounted for in the simulated groundwater level change. This assumption was made because MF-OWHM2 does not explicitly account for the soil moisture term in its equations.

4.3. Water budget

Fig. 9 shows the seasonal cycle of the simulated water budget for various water resource regions across the United States. The water budget is averaged from 1950 to 2010, which includes the average monthly precipitation as input data, the simulated total evapotranspiration (ET) and the partitioning of ET from precipitation and groundwater sources, the simulated surface runoff, and the deep percolation, which is calculated as the residual of the surface water budget: precipitation minus the surface runoff minus the evapotranspiration from precipitation. Results show how the simulated water budget components vary spatially and seasonally over the model domain. Each region has a different seasonal behavior depending on its climatological and geological conditions. For example, California has a Mediterranean climate with precipitation occurring in the winter whereas the the Ohio and Missouri river basins have higher precipitation in the summer. Consequently, the evapotranspiration seasonality is weaker in California than other regions. New England demonstrates the importance of snowmelt on the water budget, whereas the Texas Gulf water resource region has the smallest ef-

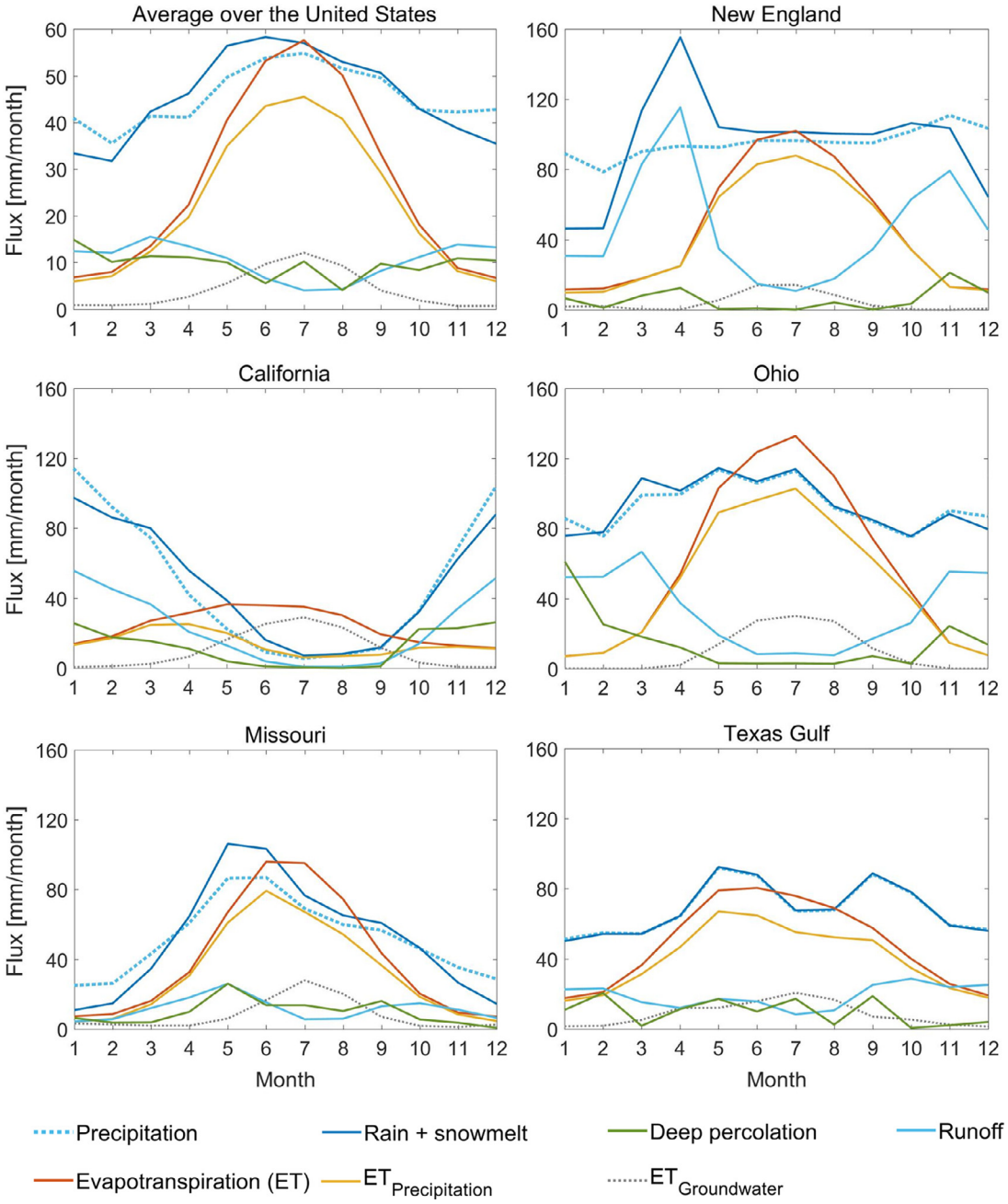

Fig. 9. Seasonal cycle of the modeled water budgets from 1950 to 2010 [mm/month]. The range of the vertical axes are the same except over the United States.

fect of snow. Incorporating snowmelt into MODFLOW affected the simulated seasonality of the water budget but not the long-term averages. Figure S7 plots the last decade of model results, showing that rainfall and runoff have the largest interannual variability. Variability in evapotranspiration is dominated by the seasonal cycle rather than interannual variability. In humid climates, like the Great Lakes and Ohio water resource regions, deep percolation primarily occurs in the winter months when ET demand is low. The same is true for California, but in this case the wet season coincides with low ET in the winter months, leading to stronger seasonality. The Great Basin, which is arid, shows a tighter relationship between precipitation and deep percolation than in the other water resource regions. Throughout the USGS water resource regions, evapotranspiration has the lowest coefficient of variation of the hydrologic fluxes analyzed here. Deep percolation and evapotranspiration had the largest, likely due to the model structure and that these integrate the impacts of both precipitation variability and ET variability. Drier regions have larger variation in precipitation, and this propagates into higher coefficients of variation in all the other hydrologic fluxes in these regions.

5. Conclusions

In this paper, the groundwater and surface water budgets were simulated on a monthly time step across the United States, using the transient MODFLOW-One-Water Hydrologic Model Version 2 (MF-OWHM2, Boyce et al. 2020, Boyce 2020). The model includes the impact of climate variability and human activities, specifically groundwater pumping, on the groundwater levels, including their seasonality. In addition, the lateral groundwater flow is simulated and interacts with surface wa-

ter. This model contributes to the literature by fully coupling the simulation of surface and groundwater processes over a large spatial scale in a publicly-available, USGS-supported model. In this paper, version 4 of the Farm Process (FMP) is used for evapotranspiration, surface runoff, and recharge estimations. The FMP module works simultaneously with the groundwater, streamflow routing, and wells pumping modules as one integrated hydrologic model.

Model validation shows that both the simulated groundwater levels and streamflow have a good correlation with the 151,415 selected observation wells and the 358 selected river gage stations, with a R^2 equal to 0.992 and 0.946, respectively. The simulated evapotranspiration (ET) also is well simulated compared with other estimation methods. The root mean square error values between the simulated evapotranspiration using MF-OWHM2 and RC-PM, NARR Inferred, and VIC estimations are 7.97, 7.23, and 3.27 mm/month respectively, over the Mississippi basin from 2003 to 2006. Based on validation of ET, streamflow, and water levels, the model is well-constrained and able to simulate the water budgets.

Overall the simulation results match well with observed head, streamflow, and VIC evapotranspiration estimates with the exceptions likely originating from model structural error, lack of sufficient input datasets, and exhaustive automatic calibration. The largest source of model structural error is the large grid cell size (13km). Each grid cell has uniform aquifer properties such as hydraulic conductivity, uniform land surface properties such as elevation and land use, and calculates a single water table elevation. The single water table elevation represents the hydraulic head at the center of the grid cell, which is used for determining stream-aquifer interaction and consumption of groundwater as evapotranspiration consumption. Due to this large grid cell size, most model error occurs in highly heterogenous terrain due to the variation in surface elevation and aquifer properties within a single grid cell. Another potential structural error is the monthly stress periods. Stress periods specify input stresses such as precipitation and potential evapotranspiration rates and assumes they are constant over the time steps in the stress period. Each stress period contained 2-time steps making the average time step length about two weeks. Within the time step the calculated evapotranspiration, surface runoff, infiltration, and stream flow are constant, which averages out the effect of synoptic events. Another source of error are the available input datasets for groundwater pumping and how accurate the pumping is reported or estimated. Lastly, given sufficient computing power and time a robust automatic calibration can improve the model simulation. Methods such as global and gradient based optimization schemes can refine the input parameters to improve the simulated results. Another model construction error is assuming the 2010 land use data does not vary in the simulation. That is the same data set is repeated for multiple years. This assumption affects the hydrological fluxes especially the evapotranspiration estimation and can be improved in the future work by including landcover change over time.

For streamflow simulations, we only included the major rivers for the stream flow routing module (SFR) for computational efficiency. Future studies can include the river tributaries to the routing (SFR) module instead of the river (RIV) module for improved representation of the physics. In this model set-up, reservoirs are not explicitly included, which may affect the streamflow results particularly where there is significant storage capacity. However, using USGS observed streamflow as an upstream boundary condition for river segments implicitly incorporates the impacts of surface water management.

This model does not simulate the water demand and irrigation supply from surface water. FMP can estimate unknown pumpage, but this option was beyond the scope of this study. A future study can include estimating irrigation pumpage for a better representation of water use in the United States. Further, this would improve the estimated stream flow from irrigation runoff and stream flow delivery as irrigation. As with any model, there are improvements that can be implemented to give more realistic simulations, but these improvements depend on data availability and the model's structural limitations. Hydraulic conductivity can be estimated from the available baseflow data, or it can be calibrated to better estimate baseflow. In addition, the specific yield can be calibrated to better estimate seasonal fluctuations in groundwater levels.

The land surface simulation with the Farm Process can be improved by specifying multiple crops per grid cell. This version of the model selected the dominant crop to be represented as the land use for each grid cell. In addition, we can include the historical land cover data change over time. This will allow us to analyze the impact the human activities, such as crops use for agriculture, and urbanizing expansion, and the impact of the natural land cover changes, such as deforestation and the natural vegetation change, on the hydrologic processes across the United States. With this addition, we can calculate the irrigation water supply and demand for each crop type and include the water irrigation from surface water and groundwater.

In conclusion, the model can simulate the groundwater levels and surface water budgets well across the large-scale region of the contiguous United States; the results show high correlation with observational data. Tradeoffs were made between model realism and computational efficiency to realistically simulate the contiguous US, while prioritizing the inclusions of the dominant hydrologic cycle dynamics. The model has the capacity of simulating the evapotranspiration, surface runoff, infiltration, groundwater pumping, three-dimensional groundwater fluxes, interaction between surface water and groundwater, and stream flow routing holistically on monthly time step across the large-scale region of the United States.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the support of NSF Water Sustainability and Climate Grant 1360446, "WSC-Category 3 Collaborative: America's Water - The Changing Landscape of Risk, Competing Demands and Climate" and the support of NSF Division of Mathematical Sciences 1720114. They thank Nathaniel Chaney for sharing the POLARIS dataset. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.advwatres.2020.103682.

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. In: Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements-FAO Irrigation and Drainage Paper 56, 300. Fao, Rome, p. D05109.

Boyce, S.E., Hanson, R.T., Ferguson, I., Schmid, W., Henson, W., Reimann, T., Mehl, S.M., Earll, M.M., 2020. One-Water Hydrologic Flow Model: A MODFLOW based conjunctive-use simulation software: U.S. Geological Survey Techniques and Methods 6–A60, p. 435. https://doi.org/10.3133/tm6A60.

Boyce, S.E., 2020. MODFLOW One-Water Hydrologic Flow Model (MF-OWHM) Conjunctive Use and Integrated Hydrologic Flow Modeling Software, version 2.0.0. U.S. Geological Survey Software Release. https://doi.org/10.5066/P9P8I8GS.

California Department of Water Resources, 2000. Explanations of Land Use Attributes Used in Database Files Associated with Shape Files: Land and Water Use Section, p. 11.

Condon, L.E., Maxwell, R.M., 2017. Systematic shifts in Budyko relationships caused by groundwater storage changes. Hydrol. Earth Syst. Sci. 21 (2).

Condon, L.E., Maxwell, R.M., 2019. Simulating the sensitivity of evapotranspiration and streamflow to large-scale groundwater depletion. Sci. Adv. 5 (6). https://doi.org/10.1126/sciadv.aav4574.

Chaney, N.W., Wood, E.F., McBratney, A.B., Hempel, J.W., Nauman, T.W., Brungard, C.W., Odgers, N.P., 2016. POLARIS: A 30-meter probabilistic soil series map of the contiguous United States. Geoderma 274, 54–67. https://doi.org/10.1016/j.geoderma.2016.03.025.

Channan, S., Collins, K., Emanuel, W.R., 2014. Global Mosaics of the Standard MODIS Land Cover Type Data. University of Maryland and the Pacific Northwest National Laboratory, College Park, Maryland, USA, p. 30.

Chow, V.T., 1959. Open-Channel Hydraulics: New York. McGraw-hill, p. 680.

de Graaf, I.E.M., Gleeson, T., van Beek, L.R., Sutanudjaja, E.H., Bierkens, M.F., 2019. Environmental flow limits to global groundwater pumping. Nature 574 (7776), 90–94. https://doi.org/10.1038/s41586-019-1594-4.

de Graaf, I.E.M., Sutanudjaja, S.H., Van Beek, L.P.H., Bierkens, M.F.P., 2015. A high-resolution global-scale groundwater model. Hydrol. Earth Syst. Sci. 19 (2), 823–837. https://doi.org/10.5194/hess-19-823-2015.

de Graaf, I.E.M., van Beek, R.L., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E.H., Bierkens, M.F., 2017. A global-scale two-layer transient groundwater model: development and application to groundwater depletion. Adv. Water Res. 102, 53–67. https://doi.org/10.1016/j.advwatres.2017.01.011.

Döll, P., Fritsche, M., Eicker, A., Schmied, H.M., 2014a. Seasonal water storage variations as impacted by water abstractions: comparing the output of a global hydrological model with GRACE and GPS observations. Surv. Geophys. 35 (6), 1311–1331.

Döll, P., Mueller Schmied, H, Schuh, C, Portmann, F.T, Eicker, A., 2014b. Global-scale assessment of groundwater depletion and related groundwater abstractions: combining hydrological modeling with information from well observations and GRACE satellites. Water Resour. Res. 50 (7), 5698–5720.

Fan, Y., Li, H., Miguez-Macho, G., 2013. Global patterns of groundwater table depth. Science 339 (6122), 940–943. https://doi.org/10.1126/science.1229881.

Fan, Y., Miguez-Macho, G., Weaver, C.P., Walko, R., Robock, A., 2007. Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations. J. Geophys. Res. 112 (D10125). https://doi.org/10.1029/2006JD008111.

Faunt, C.C., Hanson, R.T., Belitz, K., Schmid, W., Predmore, S.P., Rewis, D.L., McPherson, K., 2009. Groundwater availability of the Central Valley Aquifer, California. U.S. Geol. Surv. Prof. Pap. 1766, 225.

Ferguson, C.R., Sheffield, J., Wood, E.F., 2008. Evapotranspiration estimates from satellite remote sensing: are present-day products of sufficient accuracy? AGU 89 (53) fall meeting, abstract H43G-1092.

Flint, L.E., Flint, A.L., 2007. Regional analysis of ground-water recharge. In: Ground-Water Recharge in the Arid and Semiarid Southwestern United States, 1703, pp. 29–60.

Hanson, R.T., Flint, L.E., Faunt, C.C., Gibbs, D.R., Schmid, W., 2014a. Hydrologic Models and Analysis of Water Availability in Cuyama Valley, California (ver. 1.1, May 2015): U.S. Geological Survey Scientific Investigations Report, 2014–5150, p. 150. https://doi.org/10.3133/sir20145150.

Hanson, R.T., Boyce, S.E., Schmid, W., Hughes, J.D., Mehl, S.W., Leake, S.A., Maddock III, T, Niswonger, R.G., 2014b. One-water hydrologic flow model (MODFLOW-OWHM). US Geol. Surv. https://doi.org/10.3133/tm6A51.

Hanson, R.T., Ritchie, A.B., Boyce, S.E., Galanter, A.E., Ferguson, I.A., Flint, L.E., Flint, A., Henson, W.R., 2020. Rio Grande Transboundary Integrated Hydrologic Model and Water-Availability Analysis, New Mexico and Texas, United States, and northern Chihuahua, Mexico, p. 186, U.S. Geological Survey Scientific Investigations Report 2019–5120. https://doi.org/10.3133/sir20195120.

Harbaugh, A.W., Banta, E.R., Hill, M.C., McDonald, M.G., 2000. In: MODFLOW-2000, The U. S. Geological Survey Modular Ground-Water Model-User Guide to Modularization Concepts and the Ground-Water Flow Process, 92, p. 134 Open-file Report. U. S. Geological Survey.

Harbaugh, A.W., 2005. MODFLOW-2005, the U.S. Geological Survey Modular Ground-Water Model: The Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16.

Heath, R.C., 1983. In: Basic Ground-Water Hydrology: US Geological Survey, Water-Supply Paper 2220, 84, p. 2220.

Kollet, S.J., 2009. Influence of soil heterogeneity on evapotranspiration under shallow water table conditions: transient, stochastic simulations. Environ. Res. Lett. 4 (3), 035007.

Krakauer, N.Y., Li, H., Fan, Y., 2014. Groundwater flow across spatial scales: importance for climate modeling. Environ. Res. Lett. 9 (3), 034003.

Lehner, B., Verdin, K., Jarvis, A., 2006. HydroSHEDS Technical Documentation, Version 1.0. World Wildlife Fund US, Washington, DC, pp. 1–27.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J. Geophys. Res. 99 (D7), 14415–14428.

Liang, X., Wood, E.F., Lettenmaier, D.P., 1996. Surface soil moisture parameterization of the VIC-2L model: evaluation and modification. Global Planet. Change 13 (1–4), 195–206.

Maurer, E.P., Wood, A.W., Adam, J.C., Lettenmaier, D.P., Nijssen, B., 2002. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. J. Clim. 15 (22), 3237–3251 10.1175/1520-0442(2002)0152.0.CO;2.

Maxwell, R.M., Condon, L.E., Kollet, S.J., 2015. A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. Geosci. Model Dev. 8 (3), 923.

Maxwell, R.M., Condon, L.E., 2016. Connections between groundwater flow and transpiration partitioning. Science 353 (6297), 377–380.

Maxwell, R.M., Miller, N.L., 2005. Development of a coupled land surface and groundwater model. J. Hydrometeorol. 6 (3), 233–247.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jovic, D., Woollen, J., Rogers, E, Berbery, E.H., 2006. North American regional reanalysis. B. Am. Meteorol. Soc. 87, 343–360.

Miller, J.A., 1990. Ground water atlas of the United States. Segment 6, Alabama, Florida, Georgia, and South Carolina. Hydrologic Investigations Atlas (No. 730-G) US Geological Survey, Washington, DC https://water.usgs.gov/lookup/getspatial?aquifers_us .

Monteith, J.L., 1965. Evaporation and environment, in the state and movement of water in living organisms. In: Paper presented at the Symp. Soc. Exp. Biol., pp. 205–234.

Morris, D.A., Johnson, A.I., 1967. Summary of hydrologic and physical properties of rock and soil materials, as analyzed by the hydrologic laboratory of the US Geological Survey, 1948-60. In: Summary of Hydrologic and Physical Properties of Rock and Soil Materials, as Analyzed by the Hydrologic Laboratory of the US Geological Survey, pp. 1948–1960.

Mu, Q., Heinsch, F.A., Zhao, M., Running, S.W., 2007. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sens. Environ. 111 (4), 519–536.

Niswonger, R.G., Panday, S., Ibaraki, M., 2011. MODFLOW-NWT, a newton formulation for MODFLOW-2005. US Geol. Surv. Tech. Methods 6 (A37), 44. http://pubs.usgs.gov/tm/tm6a37/.

Peterson, S.M., Flynn, A.T., Traylor, J.P., 2016. Groundwater-Flow Model of the Northern High Plains Aquifer in Colorado, Kansas, Nebraska, South Dakota, and Wyoming (No. 2016-5153). US Geol. Surv..

Pokhrel, Y., Hanasaki, N., Koirala, S., Cho, J., Yeh, P.J., Kim, H., Kanae, S., Oki, T., 2012. Incorporating anthropogenic water regulation modules into a land surface model. J. Hydrometeorol. 13 (1), 255–269.

Prudic, D.E., Konikow, L.F., Banta, E.R., 2004. A New Streamflow-Routing (SFR1) Module to Simulate Stream-Aquifer Interaction with MODFLOW-2000 US Geological Survey open-file report 2004-1042.

Reinecke, R., Foglia, L., Mehl, S., Herman, J.D., Wachholz, A., Trautmann, T., Döll, P., 2019. Spatially distributed sensitivity of simulated global groundwater heads and flows to hydraulic conductivity, groundwater recharge, and surface water body parameterization. Hydrol. Earth Syst. Sci. 23 (11).

Scanlon, B.R., Faunt, C.C., Longuevergne, L., Reedy, R.C., Alley, W.M., McGuire, V.L., McMahon, P.B., 2012. Groundwater depletion and sustainability of irrigation in the US high Plains and Central Valley. Proc. Natl. Acad. Sci. 109 (24), 9320–9325.

Scanlon, B.R., Keese, K.E., Flint, A.L., Flint, L.E., Gaye, C.B., Edmunds, W.M., Simmers, I., 2006. Global synthesis of groundwater recharge in semiarid and arid regions. Hydrol. Process. 20 (15), 3335–3370.

Schaller, M.F., Fan, Y., 2009. River basins as groundwater exporters and importers: implications for water cycle and climate modeling. J. Geophys. Res. 114 (D4).

Schmid, W., Hanson, R.T., 2009. The Farm Process Version 2 (FMP2) for MODFLOW-2005 - Modifications and Upgrades to FMP1: U.S. Geological Survey Techniques in Water Resources Investigations, p. 102 book 6, chapter A32.

Sheffield, J., Ferguson, C.R., Troy, T.J., Wood, E.F., McCabe, M.F., 2009. Closing the terrestrial water budget from satellite remote sensing. Geophys. Res. Lett. 36 (7).

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. J. Clim. 19 (13), 3088–3111.

Sheffield, J., Wood, E.F., 2008. Global trends and variability in soil moisture and drought characteristics, 1950–2000, from observation-driven simulations of the terrestrial hydrologic cycle. J. Clim. 21 (3), 432–458.

Steeves, P., Nebert, D., 1994. 1:250,000-scale hydrologic units of the United States. 1: 250,000-Scale Hydrologic Units of the United States, 1994-0236. https://doi.org/10.3133/ofr19940236.

Tapley, B.D., Bettadpur, S., Ries, J.C., Thompson, P.F., Watkins, M.M., 2004. GRACE measurements of mass variability in the earth system. Science 305 (5683), 503–505.

Todd, D.K., Mays, L., 2005. Groundwater Hydrology (Third Edition). John Wiley and Sons, New York, p. 636.

Troy, T.J., Wood, E.F., Sheffield, J., 2008. An efficient calibration method for continental-scale land surface modeling. Water Resour. Res. 44 (9).

USDA National Agricultural Statistics Service Cropland Data Layer, 2010. Published Crop-Specific Data Layer [Online] available at https://Nassgeodata.gmu.edu/CropScape/ (accessed (7/1/2016); verified (7/18/2019)). USDA-NASS, Washington, DC.

Wada, Y., Wisser, D., Bierkens, M.F., 2014. Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. Earth Syst. Dyn. Discuss. 5 (1), 15–40.

Wada, Y., Van Beek, L., Bierkens, M.F., 2011. Modelling global water stress of the recent past: On the relative importance of trends in water demand and climate variability. Hydrol. Earth Syst. Sci. 15 (12), 3785–3805.

Wada, Y., van Beek, L.P., van Kempen, C.M., Reckman, J.W., Vasak, S., Bierkens, M.F., 2010. Global depletion of groundwater resources. Geophys. Res. Lett. 37 (20).