Model Estimates of China's Terrestrial Water Storage Variation Due To Reservoir Operation

Ningpeng Dong1,2, Jianhui Wei3, Mingxiang Yang1, Denghua Yan1,4, Chuango Yang2,4, Hongkai Gao5, Joel Arnault6, Patrick Laux3,6, Xuejun Liu7,8, Jun Niu9, Hejia Wang1, Hao Wang1, Harald Kunstmann3,6, and Zhongbo Yu2,4

1State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing, China, 2State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, China, 3Institute of Meteorology and Climate Research (IMK-IFU), Karlsruhe Institute of Technology, Campus Alpin, Garmisch-Partenkirchen, Germany, 4Joint International Research Laboratory of Global Change and Water Cycle, Hohai University, Nanjing, China, 5Key Laboratory of Geographic Information Science (Ministry of Education of China), East China Normal University, Shanghai, China, 6Institute of Geography, University of Augsburg, Augsburg, Germany, 7Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, 8University of Chinese Academy of Sciences, Beijing, China, 9Center for Agricultural Water Research in China, China Agricultural University, Beijing, China

Abstract
Understanding the role of reservoirs in the terrestrial water cycle is critical to support the sustainable management of water resources especially for China where reservoirs have been extensively built nationwide. However, this has been a scientific challenge due to the limited availability of continuous, long-term reservoir operation records at large scales, and a process-based modeling tool to accurately depict reservoirs as part of the terrestrial water cycle is still lacking. Here, we develop a continental-scale land surface-hydrologic model over the mainland China by explicitly representing 3,547 reservoirs in the model with a calibration-free conceptual operation scheme for ungaged reservoirs and a hydrodynamically based two-way coupled scheme. The model is spatially calibrated and then extensively validated against streamflow observations, reservoir storage observations and GRACE-based terrestrial water storage anomalies. A 30-year simulation is then performed to quantify the seasonal dynamics of China's reservoir water storage (RWS) and its role in China's terrestrial water storage (TWS) over recent decades. We estimate that, over a seasonal cycle, China's RWS variation is 15%, 16%, and 25% of TWS variation during 1981–1990, 1991–2000, and 2001–2010, respectively, and one-fifth of China's reservoir capacity is effectively used annually. In most regions, reservoirs play a growing role in modulating the water cycle over time. Despite that, an estimated 80 million people have faced increasing water resources challenges in the past decades due to the significantly weakened reservoir regulation of the water cycle. Our approaches and findings could help the government better address the water security challenges under environmental changes.

1. Introduction

During the last century, reservoirs were intensively built around the world for a variety of purposes, such as flood control, irrigation and water supply (Mulligan et al., 2020). By far, the global reservoir capacity has exceeded 8,000 km³, equivalent to one-sixth of the river discharge to oceans (Boulange et al., 2021). The massive water stored in the reservoirs can bring about a long-term alteration to the terrestrial water storage (TWS), a key component of the global water budget (Lv et al., 2019, 2021). Knowledge of the TWS variations under the reservoir impact can contribute to quantifying and predicting water-related hazards (Oki & Kanae, 2006; Xi et al., 2021), sea level changes (Chao et al., 2008; Lettenmaier & Milly, 2009; Pokhrel, Hanasaki, Yeh, et al., 2012), and carbon emission and uptake (Keller et al., 2021). Moreover, the fraction of reservoir storage dynamics to the TWS variation can serve as an indicative measure of the portion of water resources that are actively modulated by humans, which can provide implications for authorities and stakeholders to pursue sustainable water management policies. This is particularly important for China, as the country relies on its thousands of reservoirs to address the increasing water challenges for the well-being of 1.4 billion people (MWR, 2019).

Despite such importance of water storages, previous research efforts have focused mostly on the alteration of water fluxes induced by reservoirs (Arheimer et al., 2017; Gutenson et al., 2020; Pokhrel et al., 2016; Shin et al., 2019; Wada et al., 2011, 2016, 2017), and reports of reservoir impacts have been generally limited to alterations of...
the magnitude and timing of streamflow (Gudmundsson et al., 2021; Pokhrel, Hanasaki, Koirala, et al., 2012; Veldkamp et al., 2018; Zhao et al., 2016), mitigation of flood and drought hazards (Ehsani et al., 2017; Ngo et al., 2018; Wang et al., 2017), and increases of different magnitudes in the soil moisture (Wei et al., 2021), groundwater recharge (Dong et al., 2019), and evapotranspiration (Shah et al., 2019). In comparison with these water fluxes, the dynamics of the water stored in the reservoirs and its role in the land surface have been less studied, due to either the modeling challenges to explicitly parameterize human water management and all TWS components in hydrologic simulations (Gou et al., 2021; Liu et al., 2020; Zhang et al., 2014) or the lack of large-scale TWS component observations (Pokhrel et al., 2021). Hanasaki et al. (2006) and Zhou et al. (2016) have been two out of few studies quantifying reservoir storage at large scales, which simulated the storage variation of 456 and 166 reservoirs at global scales, respectively.

With the recent progress in the modeling approaches and observation techniques, it has been possible to quantify the interaction between the reservoir storage and other TWS components in a more precise way. Several state-of-art global hydrological models have explicitly incorporated reservoir management, surface water and groundwater processes (Pokhrel et al., 2021), and are reported to have a good performance in simulating the observed streamflow and the TWS variation in large parts of the world (Hanasaki et al., 2018; Müller Schmied et al., 2021; Sutanudjaja et al., 2018; Wada et al., 2016). The launch of the Gravity Recovery and Climate Experiment (GRACE) mission in 2002 also provides a promising way to quantify the TWS through direct observations (Tapley et al., 2004). Multiple studies have since then combined the GRACE data with the land surface modeling products (e.g., GLDAS) to derive the individual TWS storage variations for human impact assessments (Döll et al., 2014).

While above studies have greatly advanced the knowledge of TWS components globally, several critical issues remain unsolved for China. On one hand, global hydrological models, despite extensively validated at larger scales, were found with a relatively limited ability to reconstruct the TWS variation at regional scales over China (Scanlon et al., 2018; Xu et al., 2019), and most of these models consider only a small fraction of China's reservoirs due to the lack of reservoir information. On the other hand, the relatively short time period of GRACE has prevented the combination of GRACE and hydrologic models from deriving the long-term trends of TWS components, and subsequent findings were generally limited to recent years (Chen et al., 2019; Ferreira et al., 2020; Xie et al., 2019; Yin et al., 2021; Zhang et al., 2016).

Among all these issues, simulations of reservoirs without historic operation records (hereinafter ungauged reservoirs) remain one of the biggest challenges in large-scale hydrologic modeling. Currently, some studies adopted data-driven reservoir operation schemes to represent reservoirs in hydrologic models (Coerver et al., 2018; Ehsani et al., 2017; Yang et al., 2019), yet their application is generally limited in large-scale modeling due to the lack of reservoir operation records for training. In comparison, non-data-driven or conceptual reservoir operation schemes have been more widely used in hydrologic modeling. These schemes, despite with less data requirement, often introduce parameters that essentially need to be either calibrated against or extracted from historic operation records for individual reservoirs (Hanasaki et al., 2008; Wisser et al., 2010; Wu & Chen, 2012; Yassin et al., 2019; Zajac et al., 2017). To allow application of these schemes in large-scale modeling, often a single empirical parameter value is applied to all of the ungauged reservoirs. Such an approach has in some cases shown unsatisfactory accuracies (Ehsani et al., 2016), or lacked sufficient evaluation against in-situ observations. With all these unsolved issues, the water storage dynamics of China's reservoirs and its interaction with China's terrestrial water cycle over the long term have not been well understood.

Here, we aim at reducing this knowledge gap with a two-way coupling of a total of 3,547 reservoirs to the Coupled Land surface-Hydrologic Model System (CLHMS) over the mainland China, which account for over 90% of the combined capacity for China's reservoirs as of 2011. A calibration-free conceptual reservoir operation scheme for ungauged reservoirs is developed, validated, and applied to all of the reservoirs considered. The specific research question we aim to address is: What is the role of RWS in the seasonal variation of China's TWS across different time periods and spatial scales? To answer this question, the coupled model is extensively validated against streamflow observations, reservoir storage observations and in particular GRACE-based TWS anomalies at annual and seasonal scales. A multi-scale analysis focusing the RWS and TWS variation is then performed based on model simulations for a 30-year period of 1981–2010 and a spatial resolution of 10 km. To our knowledge, it is the first modeling effort on the TWS variation under anthropogenic impacts over the entire
mainland of China. Therefore, our approaches and the subsequent findings are expected to provide a novel contribution to understanding the evolution of the terrestrial water cycle under anthropogenic disturbances.

2. Study Area

China is a country located in East Asia with a total land area of approximately 9.6 million km$^2$, and the river systems nationwide are partitioned into the National Three-Level Hydrologic Divisions (NHD). The National Level-1 and Level-2 Hydrologic Divisions (NHD-1 and NHD-2) were formulated based on the natural basins and sub-basins of China’s major river systems, and the National Level-3 Hydrologic Divisions (NHD-3) were based on smaller river systems and further takes into account the integrity of the water resources system among water suppliers and users. Figure 1 depicts the NHD-1 that consists of 10 major river systems, that is, the Yangtze River, Yellow River, Pearl River, Songhuajiang River, Liaohe River, Haihe River, Northwest rivers, Southwest rivers, and Southeast rivers. The demarcation between the Northwest rivers and Southwest rivers is adjusted with the northern boundary of the Tibetan Plateau (TP). Illustrations of NHD-2 and NHD-3 are provided in Figure S1 in Supporting Information S1.

Reservoirs have been extensively built across China for the past decades. In this study, the location, dead capacity, conservation capacity, storage capacity, completion year, and/or other basic information of 3,547 reservoirs nationwide are collected and provided by the Ministry of Water Resources (MWR) and basin-level and
provincial-level water authorities. We then categorize those reservoirs with a capacity larger than 1.0 km$^3$ as extra large reservoirs, those with a capacity smaller than 0.1 km$^3$ as medium reservoirs, and those with a capacity in between as large reservoirs. As is shown in Figure 1, a total of 113 extra large reservoirs, 523 large reservoirs and 2,911 medium reservoirs are included in this study, and they have a combined capacity of 721 km$^3$, accounting for over 90% of that of China’s completed reservoirs as of 2011.

3. Methods

3.1. Coupled Land Surface-Hydrologic Model CLHMS

We use the CLHMS to depict the terrestrial water cycle of China. The CLHMS is a fully coupled modeling system of the land surface scheme of LSX and the physically based distributed hydrological model of Hydrologic Model System (HMS) (Yu et al., 1999, 2006). Note that the HMS here is different from the HEC-HMS model developed by the U.S. Army Corps of Engineers. The LSX consists primarily of a six-layer soil module, a two-layer vegetation module, a two-layer snow module, and a glacier/ice sheet module, and is capable of solving the water and energy balance in the snow-vegetation-soil continuum and on the glacier areas, thus providing the HMS model with runoff, infiltration, evapotranspiration, and other hydrologic components over each grid cell. The HMS derives the surface runoff routing with the diffusion wave equation and is fully coupled with a two-dimensional groundwater routing model, which allows the model to compute the groundwater dynamics on a raster grid basis. In particular, the HMS is able to explicitly simulate the area and storage dynamics of floodplains and natural lakes by solving the two-dimensional diffusion wave equations in the river-lake-floodplain continuum. An introduction to the key structure and components of the CLHMS model is given in Appendix A, and a detailed model description can be found in Pollard and Thompson (1995), Thompson and Pollard (1995) for LSX, and Yu et al. (2006) for HMS and the coupling mechanism.

Multiple studies using the CLHMS model have reported good agreement between the simulation and observation of multiple hydrological components, including streamflow, soil moisture content, lake surface area, groundwater levels, and evapotranspiration (Yang et al., 2010, 2012, 2013; Wagner et al., 2016). The meteorological data needed to drive the CLHMS consist of the CN05.1 dataset for precipitation (Wu & Gao, 2013) and the NCEP/NCAR reanalysis data for temperature, wind speed, radiation, pressure, and humidity (Kalnay et al., 1996). The CN05.1 dataset published by China Meteorological Administration is a 0.25° × 0.25° gridded precipitation data set interpolated from the observed precipitation of ∼2,400 rain gauges nationwide. The specific yield of aquifer and the aquifer thickness required for the groundwater module is collected from the China National Geologic Survey Dataset, which was originally published by MGMR (1990) and further processed by Yang et al. (2010). The model is run at a spatial resolution of 10 km (∼0.09° at the equator) in this study.

3.2. A Calibration-Free Reservoir Operation Scheme for Ungauged Reservoirs

Each of the 3,547 reservoirs across China is fully coupled to the CLHMS model but is not activated during a model simulation until the simulation reaches the time point where that reservoir was completed in the real world. A calibration-free conceptual operation scheme for ungauged reservoirs is developed in this study to estimate the release and storage variation of reservoirs in the model. The release of reservoirs, $Q_t$, is calculated according to the current filling level compared with a few designed water storages (Figure 2) in the following form originally formulated in our previous study (Dong et al., 2019),

$$Q_t = \begin{cases} 
\min \left( \frac{Q_{\min}}{\Delta t} \right) & (V_t \leq V_d) \\
\max \left( \frac{Q_{\min}}{\Delta t}, r \cdot U_t \right) & (V_d < V_t \leq V_c) \\
r \cdot U_t + \left( \frac{Q_{\min}}{\Delta t} - r \cdot U_t \right) \cdot \left( \frac{V_t - V_c}{V_f - V_c} \right)^k & (V_c < V_t \leq V_f) \\
\max \left( \frac{Q_{\min}}{\Delta t}, \frac{V_t - V_f}{\Delta t} \right) & (V_t > V_f)
\end{cases}$$
where $V_i$, $V_d$, $V_c$, and $V_f$ are the water storage of reservoirs at the model time step $t$, at the dead level, conservation level, and high flood level, respectively. The conservation level ($L_c$) is often fluctuating that is higher around the dry season to store water for irrigation, water supply, etc. (i.e., the maximum normal operating level), and is lower in the wet season to create more space for flood control (i.e., the flood limit level) (Figure 2). In our scheme, the $L_c$ is taken as the maximum normal operating level for the dry season and the ending month of the wet season, and is taken as the flood limit level for the rest of the wet season. The wet (dry) season refers to the months where the average inflow is higher (lower) than the mean annual inflow. The variation in $L_c$ values is evenly distributed over each day of the previous month to avoid sudden changes in the $L_c$. The high flood level ($L_f$) is the upper water level for flood control, and is in many cases at the same level as the dry-season $L_c$ (maximum normal operating level) for China's reservoirs. The $Q_{min}$ is the minimum release; $Q$ is the maximum acceptable release for flood control purposes; $k (k \leq 1)$ is a flood indicator equal to the ratio of $Q_f$ to the inflow.

The $V_d$, $V_c$, and $V_f$ are crucial to reservoir simulations and are collected in this study, as these are designed information that is much more widely shared. Zajac et al. (2017) and others have otherwise suggested empirical values of these designed water storages. On the other hand, parameters such as $Q_{min}$, $Q$, $U_i$, and $r$ need to be determined. These parameters can be calibrated or derived in a variety of forms if historical reservoir operation records are available (e.g., inflow, release, storage). However, this is not the case for most of the reservoirs. Therefore, a novel estimation strategy with no historical operation record required is proposed as follows to determine these parameters for all of the 3,547 reservoirs considered in this study, despite operation records are available for some of the reservoirs (Section 4.2).

Specifically, $Q_{min}$ and $Q$ are estimated as the 10th and 99th percentile of non-exceedance probabilities of simulated streamflow in the local grid cell, respectively. The parameter $r$ is determined at a monthly scale based on the difference between the current storage and the target storage, a time-dependent storage level often prescribed in the operation rules or expected in the actual reservoir operation (Neitsch et al., 2011). Due to the data scarcity of the observed/designated target storage, $V_d$ is taken as the target storage at the starting month of the wet season; $V_c$ for the dry season (corresponding to the maximum normal operating level), hereinafter $V_{cd}$ is taken as the target storage at the starting month of the dry season, and the monthly target storages in between are derived from a linear interpolation between $V_d$ and $V_{cd}$. The starting month of the wet (dry) season is defined as the time point where the average monthly inflow becomes higher (lower) than the mean annual inflow for the first time during a year. Then, $r$ at the month $m$ is expressed as,

$$r_m = \left(1 + \frac{V_i - V_{tar,m}}{V_{cd} - V_d}\right)^c$$

with

$$c = \min \left(\frac{L_c}{3 \cdot (V_{cd} - V_d)}, 1\right)$$
where $V_{cd}$ is the water storage at the dry-season conservation level, that is, the maximum normal operating level as noted earlier; $V_{mp,m}$ is the target storage at the month $m$; $I_e$ is the mean annual inflow, and is derived by model simulations in this study; $c$ is a parameter to avoid excessive outflow variations for multi-year reservoirs, that is, reservoirs with the mean annual inflow $I_e$ smaller than triple the conservation storage $V_{cd} - V_d$ (designed for water supply, irrigation, etc.).

The human water demand of reservoirs $U_i$ is determined at a monthly scale (i.e., a constant value is applied for each day within a month) in the following steps:

1. Distributing the multi-year averaged monthly water demand of the local basin (i.e., the local NHD-3 division in this study) to each reservoir of the same basin in proportion to the mean annual inflow $I_e$. The average monthly water demand of NHD-3 divisions is derived from public datasets and is described in Section 3.4.

2. Classifying reservoirs to water supply/irrigation ones and non-water supply/irrigation ones. Given that many reservoirs are multi-functional which makes it difficult to differentiate a water supply/irrigation reservoir from a non-water supply/irrigation one, here we define a reservoir as non-water supply/irrigation if the distributed maximum 6-month human water demand in Step 1) is lower than the simulated minimum 6-month inflow of that reservoir. In all, 48% of the reservoirs are classified as non-water supply/irrigation, with a combined capacity 76% of that of all reservoirs, indicating larger reservoirs are more likely to be classified as non-water supply/irrigation. We also note that the classification results are not very sensitive to our defined 6-month criteria—replacing 6-month with 3-month would lead to a different classification only for less than 100 reservoirs.

3. Preliminarily estimating the human water demand according to the type of reservoirs as a first-guess value. For water supply/irrigation reservoirs, $U_i$ varies with month and is estimated as the distributed multi-year averaged monthly water demand in Step 1); for non-water supply/irrigation ones, $U_i$ is estimated as a time-constant value of the mean annual inflow $I_e$.

4. Correcting the estimated human water demand. The first-guess human water demand in Step 3) is multiplied by a time-constant correction parameter $a$ specific to each reservoir, which is determined to avoid the long-term storage overfilling above $V_e$ or storage depletion below $V_d$ in reservoir simulations due to the error in the first-guess human water demand of that reservoir. Here, $a$ is solved in an iterative form to make $I_e$ and $U_e$ equal in the reservoir operation simulation, that is,

$$a_{i+1} = a_i \cdot \frac{I_e(a_i)}{U_e(a_i)}, \quad i = 0, 1, 2, \ldots$$

where $I_e$ and $U_e$ are the inflow and the human water demand averaged over the period during which the simulated storage $V_e$ is within the range of $V_d \leq V_e \leq V_c$; $I_e$ and $U_e$ change with $a_i$ for each iteration (simulation) because a new value of $a$ would change the human water demand and the period of $V_d \leq V_e \leq V_c$, over which the inflow and human water demand are averaged. Experiments suggest that an initial value of $a_0 = I_e / U_e$ can guarantee the convergence, where $U_e$ is the mean annual first-guess water demand, and five iterations are enough to reach a tolerance of 0.03. The source code of our reservoir operation scheme with a few reservoir examples is provided in https://doi.org/10.5281/zenodo.6371014.

### 3.3. Two-Way Coupling of Reservoirs in the CLHMS Model

The CLHMS model depicts the area and storage dynamics of extra large reservoirs in a similar way to the floodplains and lakes in the model. For example, the impounding of extra-large reservoirs can lead to the surface expansion of reservoirs and variations of the upstream water level over its river-reservoir-floodplain continuum, further allowing two-way feedbacks with hydrological components such as the evaporation, seepage, precipitation over the reservoir water area. Often, an accurate depiction of the reservoir-river-floodplain continuum requires detailed reservoir bathymetry information (Fleischmann et al., 2021; Shin et al., 2019). Due to the lack of such information, we employ HydroSHEDs as our DEM source in this study, which we consider as alternative approximations of the underwater bathymetry. This is because HydroSHEDs is based on the satellite images taken in February, in much coincidence with the end of dry season where most of the reservoirs in China approaching their smallest water area during a year. During this period, the reservoir terrain, which is usually

---

DONG ET AL.
covered by water during most time of a year, is visible from satellites. A more detailed description on the coupling structure is given in Appendix B.

3.4. Representation of Multi-Sectoral Water Withdrawal and Irrigation

China experiences rapidly growing water demands for the past decades and currently withdraws over 600 km$^3$ water annually from rivers, aquifers, reservoirs, etc, which needs to be explicitly represented in the CLHMS model for an accurate description of the water cycle. To derive the amount of water withdrawal, we collect the global 0.5° gridded domestic, industrial and irrigation water withdrawal data from Huang, Hejazi, et al. (2018), in which the irrigation withdrawal was reconstructed from ensemble hydrologic simulations. On its basis, we collect the 10 km national population and Gross Domestic Product (GDP) datasets in the year of 1990, 2000, and 2010 from Xu (2017a, 2017b), and the global 10 km irrigation area dataset in 2005 from the Food and Agriculture Organization (FAO). The population and GDP data for years in between are derived based on a linear interpolation, and the data before 1990 are assumed same as that in 1990. The irrigation area is assumed unchanged over time due to the lack of data. The collected domestic, industrial, and irrigation water withdrawals are then downscaled to the model resolution in proportion to the population, GDP and irrigation area, respectively, and are eventually bias corrected against the published national statistics (2001–2010) at the provincial level (Figure S2 in Supporting Information S1). The derived water withdrawal is considered equivalent to the water demand, which is then summarized at multi-year averaged monthly scales for each NHD-3 division to determine the human water demand $U_t$ of reservoirs in Section 3.2.

In the CLHMS model, the water is abstracted from the system in the following manner: (a) the water volume is subtracted, within a same NHD-3 division, from the releases of local and upstream reservoirs determined in Section 3.2 and the flows of local and upstream rivers based on the distance; and (b) if the reservoir releases and river flows are smaller than the current water demand, the remaining unsatisfied water demand is abstracted from the aquifer at the local grid cell. Here, we set no maximum groundwater depth for abstraction, and groundwater abstraction can continue until the aquifer dries out. This is to allow for groundwater overexploitation such as in the North China Plain, where water has been pumped from aquifers at depths of tens to hundreds of meters.

The abstracted water for domestic and industrial use is first removed from the system, and partially returned to the river channel in the next model timestep after the net water consumption is subtracted. The net domestic and industrial water consumption is calculated by applying a water withdrawal-to-consumption ratio at the provincial level, which is collected from national statistics averaged over 2001–2010 and is assumed constant over time. The abstracted irrigation water, on the other hand, is treated as the next-step precipitation on the vegetation layer in the irrigation grid cell and partially returns to rivers and aquifers through soil drainage as part of the model physics.

3.5. Model Set-Up and Calibration

The CLHMS requires a long spin-up period for the groundwater dynamics to reach the equilibrium state. A 5000-year model simulation is therefore carried out beforehand using the average climate conditions for 1961–1980 to produce the initial groundwater, soil moisture, and snow state for subsequent model runs. Regarding the calibration of the CLHMS, model simulations without taking account of human water management practices (reservoirs, water withdrawal, and irrigation) are performed to identify the model parameters best suitable to China’s natural hydrologic regimes. The calibration process is a two-step procedure: the first step is focused on the grid-cell water balance to ensure a low bias of the simulated total runoff, and the second step is focused on the runoff routing to ensure a satisfactory accuracy of the simulated streamflow.

In the first step, we obtain the runoff coefficient of each model grid cell from the 0.1° rasterized natural runoff coefficient dataset of China and neighboring countries (Yan et al., 2019) and take it as the reference runoff coefficient. This dataset was originally developed as part of the results of the second national water resources survey based on the precipitation and natural streamflow of around 13,600 rain gauges and 3,100 hydrologic stations, which provides the most reliable runoff conditions across the country. Aiming at accurately simulating the runoff coefficients, and thus, the runoff conditions, the parameter $\text{dcor}$ to which the runoff generation process is highly sensitive is calibrated for each grid cell. $\text{dcor}$ specifies the sub-grid partitioning of precipitation into fast direct runoff and infiltration, and is one of the three sensitive parameters in the CLHMS model. The calibration is
carried out iteratively and stops at the point where the discrepancy between the reference and simulated runoff coefficient becomes no larger than 0.05 for each grid cell.

The second step is to calibrate the runoff routing parameters of the CLHMS model, which includes the channel roughness $r_o$ and the groundwater-surface water exchange coefficient $C$ in the hydrologic part. These parameters are calibrated against the observed streamflow at 53 hydrologic stations nationwide. To identify the optimal parameter values for the natural flow conditions, the observed streamflow records we use for parameter calibration and validation mostly start from early 1960s where neither large-scale dams nor water withdrawal projects were built and thus can be considered natural (Table S1 in Supporting Information S1). For regions not covered by any of these hydrologic stations, a spatial distance similarity-based regionalization approach (DSR) is adopted, in which the routing parameters in ungauged catchments are directly transferred from their geographic neighbors. The DSR is selected because it was reported to have the potential to achieve comparably reasonable results in China among a few other regionalization methods, such as the physical similarity-based regionalization method (PSR) (Gou et al., 2021). A detailed comparison of different regionalization methods is beyond the scope of this study.

3.6. Model TWS Calculation

The TWS consists of five storage components in the CLHMS model and is calculated based on the following equation:

$$\text{TWS} = S_{\text{reservoir}} + S_{\text{snow}} + S_{\text{soil}} + S_{\text{surface}} + S_{\text{groundwater}}$$

with

$$S_{\text{surface}} = S_{\text{river}} + S_{\text{lake}} + S_{\text{glacier}}$$

where $TWS$ is the terrestrial water storage; $S_{\text{reservoir}}$, $S_{\text{snow}}$, $S_{\text{soil}}$, $S_{\text{surface}}$ and $S_{\text{groundwater}}$ are the reservoir water, snow water, soil water, natural surface water and groundwater storage, respectively; $S_{\text{river}}$, $S_{\text{lake}}$ and $S_{\text{glacier}}$ are the water storage of the rivers, natural lakes (including floodplains) and glaciers, respectively.

In this study, we define the seasonal variation of a storage component, $\Delta S_i$, as the difference between the maximum and minimum storage during a calendar year, that is,

$$\Delta S_i = S_{i,\text{max}} - S_{i,\text{min}}$$

where $S_{i,\text{max}}$ and $S_{i,\text{min}}$ are the maximum and minimum monthly value of any storage component in each calendar year, respectively.

4. Results

4.1. Model Evaluation Against Streamflow Observations

The observed streamflow at the 53 hydrologic stations is first compared with the streamflow simulated by the model without taking account of the human water management for the calibration and validation period. The relative bias (PBIAS) and the Nash–Sutcliffe efficiency (NSE) are employed to evaluate the streamflow simulations (Figure 3), with detailed information provided in the Table S1 in Supporting Information S1. It is noted that, for the combined calibration and validation period, the absolute values of the PBIAS of the simulated streamflow is less than 20% for over 85% of the stations and less than 30% for all of the stations, suggesting that our runoff coefficient-based spatial calibration is effective in reducing the bias in the water balance across the entire domain and hence in the inflow of reservoirs across the country. Likewise, over 85% of the stations see a monthly NSE value larger than 0.7, suggesting the model is able to reasonably reconstruct the natural streamflow in different parts of the country.

With the rapid socio-economic development of China, the streamflow simulation in lack of human water management starts to deviate from the observations for the recent decades (Figure 4). This is especially the case for water-scarce basins such as the Yellow River Basin and Haihe River Basin, where the flow regime is highly impacted by reservoir operation, irrigation and water withdrawal. For example, the monthly NSE values is 0.86
for the Huayuankou station during 1961–1968 but decreases to −0.02 during 1971–1990 and further decreases to −7.17 during 1991–2006 (Figure 4). On the other hand, for water-abundant basins such as the Yangtze River Basin and the Pearl River Basin, water management practices play a less significant role in the water cycle, and no remarkable accuracy degradation is observed for the entire period. By running the model taking account of water

Figure 3. The absolute values of PBIAS (left) and the Nash–Sutcliffe efficiency (NSE) values (right) of the simulated streamflow at 53 hydrologic stations for the combined calibration and validation period.

Figure 4. The observed streamflow, the simulated natural streamflow (nat), and the simulated streamflow considering human water management (hum) at the control hydrologic stations of China’s NHD-1 major river systems. PBIAS: relative bias. NSE: Nash–Sutcliffe Efficiency.
management practices, the streamflow simulation sees a notable improvement in the bias and NSE for all of the highly impacted basins. For example, the monthly NSE values increases from −1.28 to 0.67 for the Huayuankou station during 1971–2006 (Figure 4). This suggests our CLHMS model coupled with water management practices can well capture the impacted streamflow across different basins nationwide.

4.2. Model Evaluation Against Reservoir Storage Observations

In this study, the in-situ observed water storage of 16 reservoirs is collected. The satellite-derived water level data of another two reservoirs, namely Xiaowan and Nuozhadu, are also collected from the Global Reservoirs and Lakes Monitor (G-REALM) database and then combined with their respective in-situ storage-level correlations to calculate their historic water storage. In all, 18 reservoirs are used to validate the model-simulated reservoir storage with our calibration-free operation scheme for ungauged reservoirs. These reservoirs have a combined capacity equivalent to over 30% of China's total reservoir capacity as of 2011, and are distributed in different NHD-1 divisions, designed for distinct purposes (hydropower, irrigation, flood control, etc.), and with a capacity ranging from 0.14 to 45.07 km$^3$ (Figure 1 and Table 1).

Table 1
Summary of Validation Reservoirs

<table>
<thead>
<tr>
<th>Reservoir</th>
<th>Basin</th>
<th>Capacity (km$^3$)</th>
<th>Inflow (km$^3$/a)</th>
<th>Residence time (a)</th>
<th>Regulation ability</th>
<th>Designed purposes$^a$</th>
<th>Classified as$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three Gorges</td>
<td>Yangtze</td>
<td>45.1</td>
<td>410</td>
<td>0.11</td>
<td>within-year</td>
<td>F, H, W</td>
<td>NWI</td>
</tr>
<tr>
<td>Danjiangkou</td>
<td>Yangtze</td>
<td>21.0$^c$</td>
<td>39.0</td>
<td>0.54</td>
<td>multi-year</td>
<td>F, W</td>
<td>NWI</td>
</tr>
<tr>
<td>Longyangxia</td>
<td>Yellow</td>
<td>27.4</td>
<td>17.6</td>
<td>1.56</td>
<td>multi-year</td>
<td>F, H</td>
<td>NWI</td>
</tr>
<tr>
<td>Xiaowan</td>
<td>Mekong</td>
<td>15.0</td>
<td>35.0</td>
<td>0.43</td>
<td>multi-year</td>
<td>F, H</td>
<td>NWI</td>
</tr>
<tr>
<td>Fengman</td>
<td>Songhuajiang</td>
<td>10.8</td>
<td>11.9</td>
<td>0.91</td>
<td>multi-year</td>
<td>F, I, W</td>
<td>NWI</td>
</tr>
<tr>
<td>Xiaolangdi</td>
<td>Yellow</td>
<td>12.7</td>
<td>42.3</td>
<td>0.30</td>
<td>within-year</td>
<td>F, I, W</td>
<td>NWI</td>
</tr>
<tr>
<td>Baishi</td>
<td>Liaohe</td>
<td>1.6</td>
<td>0.3</td>
<td>5.47</td>
<td>multi-year</td>
<td>I, W</td>
<td>WI</td>
</tr>
<tr>
<td>Jinping I</td>
<td>Yangtze</td>
<td>8.0</td>
<td>38.7</td>
<td>0.21</td>
<td>within-year</td>
<td>F, H</td>
<td>NWI</td>
</tr>
<tr>
<td>Ertan</td>
<td>Yangtze</td>
<td>6.1</td>
<td>45.4</td>
<td>0.14</td>
<td>within-year</td>
<td>H</td>
<td>NWI</td>
</tr>
<tr>
<td>Huanglongtan</td>
<td>Yangtze</td>
<td>1.2</td>
<td>5.4</td>
<td>0.22</td>
<td>within-year</td>
<td>H, I</td>
<td>NWI</td>
</tr>
<tr>
<td>Tuanjie</td>
<td>Yangtze</td>
<td>0.14</td>
<td>0.5</td>
<td>0.28</td>
<td>within-year</td>
<td>F, I, W</td>
<td>WI</td>
</tr>
<tr>
<td>Hongmen</td>
<td>Yangtze</td>
<td>1.2</td>
<td>2.3</td>
<td>0.53</td>
<td>within-year</td>
<td>F, I, H</td>
<td>NWI</td>
</tr>
<tr>
<td>Pubugou</td>
<td>Yangtze</td>
<td>5.3</td>
<td>25.2</td>
<td>0.21</td>
<td>within-year</td>
<td>F, H</td>
<td>NWI</td>
</tr>
<tr>
<td>Hongjiadu</td>
<td>Yangtze</td>
<td>5.0</td>
<td>4.9</td>
<td>1.02</td>
<td>multi-year</td>
<td>F, H</td>
<td>NWI</td>
</tr>
<tr>
<td>Liujiaxia</td>
<td>Yellow</td>
<td>5.7</td>
<td>22</td>
<td>0.26</td>
<td>within-year</td>
<td>H, I</td>
<td>NWI</td>
</tr>
<tr>
<td>Nuozhadu</td>
<td>Mekong</td>
<td>23.7</td>
<td>53.9</td>
<td>0.44</td>
<td>multi-year</td>
<td>F, H</td>
<td>NWI</td>
</tr>
<tr>
<td>Xiluodu</td>
<td>Yangtze</td>
<td>12.7</td>
<td>137</td>
<td>0.09</td>
<td>within-year</td>
<td>F, H</td>
<td>NWI</td>
</tr>
<tr>
<td>Meishan</td>
<td>Huaihe</td>
<td>2.3</td>
<td>2.2</td>
<td>1.04</td>
<td>multi-year</td>
<td>F, I, W</td>
<td>WI</td>
</tr>
</tbody>
</table>

$^a$In alphabetical order. F, Flood control; H, Hydropower; I, Irrigation; W, Water supply. $^b$Classification results in Section 3.2. NWI, Non-water supply/irrigation; WI, Water supply/irrigation. $^c$The capacity of the Danjiangkou reservoir increased to 33.9 km$^3$ in 2013.

Figure 5 compares the observed and simulated storage of the 18 reservoirs and quantifies the simulation results with the correlation coefficient $R$ and the NSE. Here, the NSE is calculated with the storage normalized by subtracting the mean value in particular to reflect the accuracy of storage variations rather than its absolute values. Results indicate that more than 70% of the reservoirs see a $R > 0.7$ over the entire period and a $R > 0.85$ over a seasonal cycle, and around 80% of the reservoirs see a $NSE > 0.4$. A few multi-year reservoirs, for example, Baishi, show a relatively less desirable simulation result during a multi-year period, whereas it remains satisfactory at seasonal scales. Among within-year reservoirs (i.e., reservoirs with the mean annual inflow $I_a$ larger than triple the conservation storage $V_{sd} - V_d$), the model shows a limited ability in reconstructing the storage variation of Liujiaxia, possibly because the Liujiaxia is jointly operated with upstream reservoirs to meet...
Figure 5. Reservoir storage simulations with our calibration-free reservoir operation scheme for ungauged reservoirs, as compared with observations. $R$ is Pearson’s correlation coefficient. Nash–Sutcliffe efficiency (NSE) (normalized) is calculated with storage series normalized by subtracting the mean value.
the fluctuating energy demands, which is difficult to be precisely captured by the model due to the lack of a joint operation scheme for cascade reservoirs.

In general, the seasonal cycle of the simulated water storage is in good agreement with observations for most of the reservoirs, which suggests that our proposed operation scheme for ungauged reservoirs has the potential to reproduce the reservoir operation processes without historical operation data, and that the coupled model can be used for subsequent analysis of RWS.

### 4.3. Model Evaluation Against GRACE-Based TWS Anomalies

To further evaluate the model capability of simulating the TWS across the country, we collect the TWS anomalies (TWSA) with a baseline period of 2004–2009 from the GRACE JPL and CSR mascons solutions, and take their average as the reference TWSA for comparison with the model simulations for the period of 2003–2010.

Figure 6 presents the GRACE-based TWSA and the model-estimated TWSA with and without taking account of water management practices. During a multi-year period, results indicate that the TWSA simulations without water management practices show no clear trends for the Haihe River and Yellow River and an increasing trend for the Northwest Rivers, which are in poor agreement with the GRACE data. By coupling reservoirs and other water management practices, the TWSA simulation shows an improvement especially for the Haihe River, Yellow River, and other highly impacted basins. The TWS of these basins shows a more downward trend as compared to the natural conditions, primarily due to the considerable amount of surface water and groundwater withdrawal relative to their limited water resources. During a seasonal cycle, the simulated TWS variations also match well with the GRACE data for most of the basins, especially for simulations taking account of human water management practices. The Northwest Rivers generally see a larger discrepancy between simulations and GRACE, which could be due to the sparsely distributed rain gauges and hence low quality of precipitation fields there. In general, both the multi-year and within-year TWS simulations taking account of water management practices show good agreement with the GRACE data for most of the basins, which suggests that the model can be used for subsequent analysis of TWS.

### 4.4. Seasonal Dynamics of China’s RWS Averaged Over 1981–2010

On the national average, our simulation results indicate that the RWS sees a seasonal variation of 8.5 mm or 79 km$^3$ (equivalent to 19% of China’s TWS variation) over the period of 1981–2010. At basin scales, the seasonal variation of RWS is 0.6 mm or 1.3 km$^3$ (equivalent to 3% of the TWS variation) for the Northwest Rivers, 0.6 mm or 0.9 km$^3$ (equivalent to 1% of the TWS variation) for the Southwest Rivers. These regions are sparsely populated and hence reservoirs were not densely built, making the RWS variation only a very small fraction compared to the TWS variation. The RWS variation is larger for the Huaihe River (6.9 mm or 2.3 km$^3$), Liaohe River (7.7 mm or 2.4 km$^3$), Songhua Jiang River (10.0 mm or 9.1 km$^3$), Haihe River (9.0 mm or 3.0 km$^3$), and Southeast Rivers (23.4 mm or 4.6 km$^3$), which are equivalent to 8%, 15%, 21%, 23%, 24% of the TWS variations during 1981–2010, respectively. The RWS variation of the Yangtze River (16.7 mm or 30.1 km$^3$) and the Pearl River Basin (22.0 mm or 12.7 km$^3$) is one of the largest among all the basins and represents 18% and 16% of their TWS variation, respectively. The Yellow River sees an RWS variation of 15.5 mm or 12.3 km$^3$, which is equivalent to 43% of the TWS variation.

During a seasonal cycle, the monthly peak reservoir storage occurs between September and October for most of the basins, in coincidence with the end of the wet season for most parts of China (Figure S3 in Supporting Information S1). This reveals a common reservoir management policy of keeping a high water level at the end of the wet season to meet the water demand for the forthcoming dry season. On the other hand, the minimum monthly reservoir storage occurs between March and April for Southern China but around June for Northern China. For example, the Yangtze River and the Pearl River see their smallest RWS in April, which is 2 months earlier than the Haihe River, Yellow River, Liaohe River, and Songhua Jiang River in Northern China. The earlier impounding of reservoirs in Southern China is possibly due to the earlier onset of the India monsoon and East Asia monsoon and hence the earlier wet season in Southern China than Northern China.
4.5. Role of RWS in the TWS at Regional and Local Scales

Here we further present the seasonal variation of the RWS as a fraction of that of the TWS for each of the NHD-3 divisions averaged over 1981–2010 (Figure 7). It is noted that the role of RWS in the TWS vary significantly for different regions. The RWS variation represents less than 10% of the TWS variation for 130 of the 209 NHD-3 divisions, especially in Western China and North China Plain. On the other hand, the RWS variation represents over 50% of the TWS variation for 14 out of 209 NHD-3 divisions. In these regions, the hydrologic regime is often regulated by mega dams or cascade reservoirs. For example, the only two divisions where the RWS variation represents over 90% of the TWS variation lie in the middle and upper Yellow River Basin, where four mega dams, namely Xiaolangdi, Sanmenxia, Liujiaxia, and Longyangxia are located. Three of these four reservoirs rank among the largest 10 reservoirs in China.

Figure 6. The GRACE-based TWSA, the simulated TWSA neglecting human water management (nature), and the simulated TWSA considering human water management (human) for China's NHD-1 major river systems during 2003 and 2010.
Figure 8 presents the seasonal variation of the RWS as a fraction of that of the TWS averaged over 2006–2010 at the local scale. Notably, the magnitude of the RWS variation relative to the TWS variation is likely related to the reservoir capacity. The water storage of most of the extra large reservoirs and large reservoirs is equivalent to over 80% of the TWS variation at local grid cells. In particular, the water storage and floodplains of extra large reservoirs can straddle over a few grid cells upstream due to their massive capacities. This is the case for several China's largest reservoirs, for example, the Three Gorges, Danjiangkou, Longyangxia, Tianshengqiao, and Ertan Reservoir (Figure 8). The maximum extent of these reservoirs simulated by the model can stretch upstream by dozens to hundreds of kilometers, which are generally in good agreement with the HydroLakes, a digital map repository providing the shoreline polygons of global lakes (Messager et al., 2016). The storage variations of these reservoirs are almost equivalent to the TWS variations over its floodplain, indicating a dominant role of reservoirs in the local hydrologic cycle. Medium reservoirs, on the other hand, mostly experience a storage variation equivalent to less than 60% of the TWS variation in the local grid cell due to their relatively small size.

4.6. Trends of Seasonal Dynamics of RWS Relative to TWS

Given that China did rapid dam construction for the past decades (Figure 9a), we further compare the fraction of the RWS variation to the TWS variation during three sub-periods, that is, 1981–1990 (1980s), 1991–2000 (1990s), and 2001–2010 (2000s). For the national average, results indicate that the seasonal variation of RWS is 6.5, 7.5, and 11.6 mm for 1980s, 1990s and 2000s, representing 15%, 16%, and 25% of the TWS variation, respectively (Figure 9b), and the seasonal variation of the RWS increases from 1980s to 2000s for most of the NHD-1
This is an expected result because China's reservoir capacity in 2000s is 120 km$^3$ larger than in 1990s and 180 km$^3$ larger than in 1980s. This leads to the RWS variation representing a larger portion of the TWS variation in 2000s than in other sub-periods, especially for the Yangtze River, the Songhuajiang River, and the Northwest Rivers (Figures 9b and 9c).

To characterize the variation of the role of reservoirs over the recent decades at regional scales, we present the trend of the RWS variation relative to the TWS variation in each of the NHD-3 divisions for the period of 1981–2010 (Figure 10). Notably, for most of the NHD-3 divisions, the RWS plays an increasingly more important role in the TWS variation over time, especially in Southern China where reservoirs were intensively built in recent decades. Despite that, a significant decrease in the RWS variation relative to the TWS variation is observed for some of the NHD-3 divisions that are mostly located in the Yellow River Basin and the Liaohe River Basin. To investigate its possible cause, we further calculate the trend of RWS variation and TWS variation separately. As shown in Figures S4 and S5 in Supporting Information S1, most of these NHD-3 divisions experience a significantly decreased RWS variation, but experience no increasing trend of TWS variation. This suggests that the decreasing role of reservoirs for these regions is unlikely to be caused by the TWS variation, but more likely to be caused by the decreasing RWS variation. The decreased RWS variation could be a result of the increasing upstream water demands that reduce the inflow and diminish the functionality of reservoirs, plus no large-scale reservoir was constructed for the past decades for these regions.

Figure 8. The fraction (−) of the simulated seasonal variation of reservoir water storage (RWS) to that of the terrestrial water storage (TWS) at grid cell scales during 2006–2010, and the simulated extent of the selected five extra large reservoirs overlaid by the HydroLakes polygon.
5. Discussion

5.1. Reservoir Storage Effectively Used During a Seasonal Cycle

At the national scale, we estimate that the seasonal variation of RWS is approximately 21% of the total capacity of reservoirs during 1981–2010, and this value varies between 15% and 25% in different years (Figure 9a). Despite a seemingly small portion, this is a higher estimate than previous estimates at global scales. For example, Zhou et al. (2016) simulated 166 of the largest reservoirs around the globe, and estimated the magnitude of the reservoir storage variation is about 10% of their combined reservoir capacity. Aside from the different reservoirs considered in both studies, this could be due to the fact that most part of China is controlled by the monsoon climate, which causes a large variation of precipitation and water resources among different seasons, leading to a larger reservoir storage variation to attenuate the fluctuated water resources.

5.2. Population Exposure to a Less Modulated Water Cycle

China was historically prone to frequent flood and drought hazards, and reservoirs were extensively built across the nation to increase the water security for one-fifth of world’s population. To investigate the likely challenge China’s population and economy may face in terms of the water resources management, we employ the fraction of the RWS variation to the TWS variation as a quantitative indicator. Specifically, a low fraction of RWS variation to the TWS variation indicates an untamed water cycle, and a decreasing trend of RWS variation over TWS variation suggests that the water cycle is becoming less modulated by the reservoirs during a seasonal cycle over time. In both cases, the water system is likely to suffer water-related hazards to a comparably greater extent, and vice versa.

Figure 9. (a) The combined reservoir capacities of China considered in our study, derived based on the completion year of the 3,547 reservoirs collected from the Ministry of Water Resources and local authorities, and the simulated seasonal variation of RWS over time. (b) The fraction (−) of the simulated seasonal variation of RWS to that of the TWS for NHD-1 diversions, and (c) the national averaged seasonal variation of RWS during 1980s, 1990s, and 2000s.
Our results in Figure 7 show that 570 million people or 41% of China’s population live in regions where the RWS variation represents less than 10% of the TWS variation. For some of these regions (e.g., the Huaihe River Basin), large-scale dam construction is not possible due to their flat/complex terrain and dense population, making them relatively prone to water-related hazards. For several other regions, the considerable water resources availability and variability hinder reservoirs from effectively guaranteeing the water security, albeit a massive reservoir capacity in these regions (Figure 7). This is the case for the Yunnan Province located around the boundary of the Southwest Rivers, Yangtze River and Pearl River, which have suffered increasing economic losses from intensified floods and droughts over the recent decades.

With the rapid construction of reservoirs nationwide for the past decades, an increase in the RWS variation over TWS variation at the national level is found with our simulations (Figure 9), indicating an overall stronger ability of reservoirs to manage the water-related hazards. This is in line with the national statistics on the flood and drought disasters, which record a significantly decreasing trend of both the death toll and the affected area due to floods and droughts during 1981–2010 (MWR, 2019). Despite that, our results in Figure 10 show that there is a significantly decreasing trend of RWS variation over TWS variation for some of the NHD-3 divisions, for example, in the Haihe River Basin, Songhuajiang River Basin and Liaohe River Basin (Figure 10). With a total population of around 80 million, these areas are comparably more vulnerable to water security issues under the changing environment as a result of a significantly weakened ability of reservoirs to regulate the water cycle over time. This is consistent with several basin-level reports, which note a rise in the loss of life and property caused by droughts in the Haihe River Basin and Songhuajiang River Basin over recent decades, despite an overall decreased loss at the national level.

![Figure 10. The trend of the fraction of the simulated seasonal variation of RWS to that of the TWS for NHD-3 divisions during 1981–2010. The line hatch denotes the RWS variation over the TWS variation is larger than 0.1. Decrease (orange) denotes a decreasing trend with a negative \(Z_a\), and Increase (blue) denotes an increasing trend with a positive \(Z_a\), both using the Mann–Kendall test. Significantly decrease (red) and significantly increase (dark blue) denote the trend is significant at the 0.05 level, that is, the null hypothesis of no trend is rejected at \(a = 0.05\), using the Mann–Kendall test.](image)
In such a context, effective policy interventions should be considered to manage these issues for the identified vulnerable areas, such as planning of new water infrastructures, and adaption of water management practices to the availability and variability of water resources. The application of virtual water trades and water right trades can also play an important role in improving the resilience of the water system through water demand reduction in these areas (Krueger et al., 2019).

5.3. Sources of Uncertainties

In this study, the model exhibits a satisfactory accuracy in reconstructing observed streamflow, reservoir storage and GRACE estimates, and our results are generally in good agreement with the findings of existing literature. However, our simulation is subject to a few sources of uncertainties that should be taken care of.

First, our proposed operation scheme for ungauged reservoirs, despite extensively evaluated, is nevertheless a simplification of actual reservoir operation. Real-time decision-making on dam releases usually take account of hydrologic forecasts, fluctuating water demands, expert experiences and other frequently updated information, which can induce uncertainties in simulation of reservoir operation (Wu et al., 2018; Yang et al., 2019). It is noted that recent development in satellite altimetry, such as the launch of Ice, Cloud and land Elevation Satellite (ICESat-2), has the potential to derive the target storage in our operation scheme remotely to further improve the simulation of reservoir operation (Bonnema & Hossain, 2017, 2019; Cooley et al., 2021; Fan et al., 2020; Getirana et al., 2020; Han et al., 2020; Huang, Long, et al., 2018; Li et al., 2019), and this will be one of our future research directions. In addition, the irrigation water demand is collected from the results of ensemble hydrologic simulations in this study, which may be subject to uncertainties induced by the model structure. Previous studies have shown the irrigation water demand is related to crop types, irrigation techniques, irrigated areas, water availability, climate conditions, etc. (Yin et al., 2020), and discrepancies exist in the simulated amount of irrigation water demand between different models (Pokhrel et al., 2016). Despite that, we believe that this uncertainty is minor as the collected water withdrawal is bias corrected to reduce the uncertainty and error as much as possible.

Moreover, the meteorological forcing could add to the model uncertainties. For example, the input precipitation forcing of the model was derived from ∼2,400 rain gauges nationwide and has been proven very reliable for most parts of China where there are a high density of rain gauges (Shang et al., 2021). However, for regions where the rain gauges are sparsely distributed, for example, in Northwest China, the real precipitation fields of the region can be difficult to capture precisely, which can further induce errors in hydrologic simulations (Ehsan Bhuiyan et al., 2019).

Regardless of these uncertainties, the model validation results suggest that our land surface-hydrologic model coupled with human water management can generally serve as a useful tool to reproduce the reservoir storage and the altered terrestrial water cycle across China’s landmass, and subsequent findings on the decomposition of the TWS variation can provide insights into the interaction between the human society and the natural hydrological system of China.

6. Conclusions

To better understand the role of reservoirs in China's terrestrial water cycle, we explicitly incorporate a total of 3,547 reservoirs into the CLHMS land surface-hydrologic model over the mainland China by proposing a calibration-free operation scheme for ungauged reservoirs. Validation results suggest that our proposed scheme can well reproduce the storage variations of China’s major reservoirs, and the coupled model exhibits a satisfactory accuracy against the gauged streamflow, observed reservoir storage and GRACE-estimated TWS anomalies. Our major findings are highlighted as follows:

1. During the period of 1981–2010, the average seasonal variation of China's RWS is approximately 8.5 mm or 79 km³, as large as 19% of China’s TWS variation. The RWS variation accounts for approximately one-fifth of the total reservoir capacity, which is higher than previous estimates at the global scale.
2. The role of reservoir water storage in the TWS varies with basins. For the Yellow River Basin, the RWS variation is equivalent to up to 43% of the TWS variation, while it is nearly zero for Northwest Rivers and Southwest Rivers. For most of the other major river basins, the RWS variation represents around 10%–30% of the TWS. For all of the NHD-1 basins, the reservoir storage variations increase in the past decades.
3. Reservoir water storage dominates the TWS in the local grid cell where the dam is located. Extra large reservoirs often exhibit a large storage variation almost equivalent to the TWS variation in the river-reservoir-floodplain continuum over its maximum extent. Many of the smaller reservoirs experience storage variations of less than 60% of the TWS, with their water storage generally limited within a grid cell.

4. In most regions, reservoirs play a growing role in modulating the water cycle over time. The RWS represents 15%–16% of TWS variation during 1980–1999, and increases to 25% during 2000s on the national scale. Despite that, it is estimated that 80 million people live in regions where the water cycle became significantly less controlled by reservoirs during a seasonal cycle over the recent decades.

This study focuses on the seasonal variability of the RWS, and future work intend to further improve our proposed reservoir operation scheme with the possible use of satellite altimetry and moves on to the annual variability of the storage components at multiple scales, aiming at providing a broader view of the co-evolution between reservoir operation and the hydrologic cycle in a changing environment.

Appendix A: Brief Introduction to the CLHMS Model

The CLHMS is a fully coupled modeling system of the land surface scheme of LSX and the physically based distributed hydrological model of Hydrologic Model System (Yu et al., 1999, 2006).

The LSX is the land surface scheme of the GENESIS global climate model, and simulates a single vertical column of vegetation, snow, soil, ocean and glacier at each grid cell. The vegetation module consists of an upper layer of trees and a lower layer of grasses. Radiative fluxes and turbulent fluxes of momentum, sensible heat, and water vapor through the vegetation to the soil surface are calculated at each time step, including processes such as transpiration and interception of precipitation by the vegetation canopy. The soil module has six layers over the upper 4.25 m, with temperature and soil moisture (ice and liquid) predicted for each layer. Physical processes in the vertical soil column include heat diffusion, unsaturated liquid water transport and saturated gravitational drainage, local surface runoff and bottom drainage, uptake of liquid water by plant roots for transpiration, and explicit freezing and thawing of soil ice.

The HMS is a physically based hydrologic model originally developed from the Basin-Scale Hydrologic Model (BSHM), and has been coupled with several land surface models such as LSX and Noah (Wagner et al., 2016). When coupled with land surface models, the HMS consists of a river routing module with one-dimensional diffusion wave equations, a lake/floodplain hydrodynamic module with two-dimensional diffusion wave equations, and a groundwater routing module with two-dimensional Boussinesq equations. The lake/floodplain hydrodynamic module is automatically activated only if the simulated river stage of a local grid cell rises above the riverbed, causing the river inundating the entire grid cell.

Traditional land surface models employ free drainage boundary at the bottom of the soil column and neglects the flux exchange with groundwater. This free drainage boundary was replaced in the CLHMS by a computationally efficient parameterization that approximates the solution of the Darcy–Buckingham equation with respect to the LSX bottom layer moisture content and the groundwater depth, assuming a quasi-steady state moisture profile. Further details can be found in Fersch et al. (2013).

The LSX runs at a temporal resolution of 1 hr and aggregates the hourly runoff and the bottom two-way flux to daily values before passing these variables to the HMS model. The HMS takes in these variables at a daily basis for surface and groundwater routing. To ensure numerical convergence, a routing step of 20 min is set in the model for river routing, groundwater routing and reservoir simulation. In our study, all variables are output at a daily basis.

Appendix B: Description of the Reservoir Modeling

B1. Reservoir Water Balance

The reservoir water storage at the model time step \( t \), \( V_t \), is depicted with the water balance equation, that is,

\[
V_t = V_{t-1} + \Delta t \cdot (I_t - Q_t + A_t \cdot P - A_t \cdot E - A_t \cdot D)
\]
where $V_{t-1}$ is the reservoir storage at the last time step; $\Delta t$ is the model time step; $I$ and $Q_s$ are the inflow and the release of the reservoir, respectively; $A$ is the water surface area at the current time step; $E$ is the evaporation loss, which is calculated from the Penman Equation; $P$ is the precipitation on the surface of reservoirs; $D$ is the reservoir-groundwater flux, which is calculated using the Darcy’s Law.

**B2. Coupling Equations**

Figure 2 presents a schematic diagram depicting the reservoir coupling mechanism within the CLHMS model. The hydrodynamic process upstream of the dam is depicted with the two-dimensional diffusion wave equation, that is,

$$\frac{dh}{dt} + \frac{1}{w} \left( \frac{dQ_x}{dx} + \frac{dQ_y}{dy} \right) = (1 - f_h - f_x) R + (f_x + f_s)(P - E) - f_s C - f_x D$$

where $h_i$ is the surface water level, $w$ is the width, $Q_x$ and $Q_y$ discharge at $x$ and $y$ directions, respectively. $f_h$ and $f_x$ are area fractions of the riverbed and reservoir; $R$, $P$, $E$, $C$, $D$ are runoff, precipitation, evapotranspiration, river-groundwater flux and reservoir-groundwater flux, respectively.

Similarly, the hydrodynamic process downstream of the dam is depicted as,

$$\frac{dh}{dt} + \frac{1}{w} \cdot \frac{dQ_x}{dx} = (1 - f_h) R + f_s(P - C) + \frac{Q_s}{f_s L^2}$$

where $L$ is the grid cell length.

The reservoir-groundwater flux exchange is accounted for in the groundwater hydrodynamic process with the two-dimensional Boussinesq equation, that is,

$$\frac{dV_g}{dt} = \frac{d}{dx} \left[ K(h_g - b) \frac{dh_g}{dx} \right] + \frac{d}{dy} \left[ K(h_g - b) \frac{dh_g}{dy} \right] + (1 - f_h - f_x) I + f_s C + f_x D$$

where $V_g$ is groundwater volume over unit area, $K$ is the saturated hydraulic conductivity, $h_g$ is the groundwater level, $I$ is the infiltration.

The reservoir evaporation (latent heat) is added to the model evapotranspiration following Wei et al. (2021), that is,

$$E_m = f \cdot E + (1 - f) (E_d + E_t)$$

where $E_m$ is the evapotranspiration flux of the grid cell, $f$ is the water area fraction, $E_d$ and $E_t$ are direct evaporation flux and transpiration flux, respectively.

**Data Availability Statement**

The source code of our validation-free reservoir operation scheme with a few reservoir examples is provided in https://doi.org/10.5281/zenodo.6371014. The downscaled and bias corrected water withdrawal of China is provided in https://doi.org/10.5281/zenodo.6365598. The Global Reservoirs and Lakes Monitor (G-REALM) database is available at https://globalreservoirs.appspot.com; the NCEP reanalysis data are available at https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html; the global irrigation area is available at https://podaac.jpl.nasa.gov/dataset/TELLUS_GRAC-GRFO_MASCON_CRI_GRID_RLO6_V2; the GRACE CSR mascon product is available at https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html; the global reconstructed water withdrawal is available at https://zenodo.org/record/1209296; the rasterized runoff coefficient dataset of China and neighboring countries is available at https://zenodo.org/record/1209296; the HydroLakes data is available at https://www.hydrosheds.org/page/hydrolakes. Additional data support is attributed to the National Earth System Science Data Center, National Science & Technology Infrastructure of China.