Wiley Interdisciplinary Reviews: Water

WILEY

WIREs WATER

ADVANCED REVIEW OPEN ACCESS

The Unexploited Treasures of Hydrological Observations Beyond Streamflow for Catchment Modeling

Paul D. Wagner[1,2] | Doris Duethmann[3,4] | Jens Kiesel[1,5] | Sandra Pool[6,7] | Markus Hrachowitz[8] | Serena Ceola[9] | Anna Herzog[10] | Tobias Houska[11,12] | Ralf Loritz[13] | Diana Spieler[14,15] | Maria Staudinger[16] | Larisa Tarasova[17] | Stephan Thober[18] | Nicola Fohrer[1] | Doerthe Tetzlaff[3,19] | Thorsten Wagener[10] | Björn Guse[1]

[1]Department of Hydrology and Water Resources Management, Kiel University, Kiel, Germany | [2]Applied Physical Geography - Environmental Hydrology and Resource Management, Department of Earth Sciences, Freie Universität Berlin, Berlin, Germany | [3]Department Ecohydrology and Biogeochemistry, Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany | [4]State Office for the Environment Rhineland-Palatinate, Mainz, Germany | [5]Stone Environmental, Montpelier, Vermont, USA | [6]Department of Infrastructure Engineering, University of Melbourne, Melbourne, Victoria, Australia | [7]Department Water Resources and Drinking Water, Eawag - Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland | [8]Department of Watermanagement, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands | [9]Department of Civil, Chemical, Environmental and Materials Engineering, Alma Mater Studiorum Università di Bologna, Bologna, Italy | [10]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany | [11]Department of Landscape Ecology and Resources Management, Justus Liebig University, Gießen, Germany | [12]Institute of Soil Science and Site Ecology, TUD Dresden University of Technology, Dresden, Germany | [13]Institute of Water and Environment, Karlsruhe Institute of Technology (KIT), Germany | [14]Institute of Hydrology and Meteorology, TUD Dresden University of Technology, Germany | [15]Schulich School of Engineering, University of Calgary, Canada | [16]Department of Geography, University of Zurich, Zurich, Switzerland | [17]Department Catchment Hydrology, Helmholtz Centre for Environmental Research – UFZ, Halle, Germany | [18]Department Computational Hydrosystems, Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany | [19]Department of Geography, Humboldt University Berlin, Berlin, Germany

Correspondence: Doris Duethmann (doris.duethmann@mailbox.org)

Received: 24 May 2024 | **Revised:** 24 February 2025 | **Accepted:** 7 March 2025

Associate Editor: Jason Leach | **Editor-in-Chief:** Wendy Jepson

Funding: We acknowledge funding by the German Research Foundation for the scientific network on Identification and analysis of process limitations in hydrological model structures (IMPRO, Project number 471280762).

Keywords: Catchment Hydrology | Hydrological Modeling | In-situ data | Multi-variable Calibration | Satellite-derived data

ABSTRACT

While measured streamflow is commonly used for hydrological model evaluation and calibration, an increasing amount of data on additional hydrological variables is available. These data have the potential to improve process consistency in hydrological modeling and consequently for predictions under change, as well as in data-scarce or ungauged regions. Here, we show how these hydrological data beyond streamflow are currently used for model evaluation and calibration. We consider storage and flux variables, namely snow, soil moisture, groundwater level, terrestrial water storage, evapotranspiration, and altimetric water level. We aim at summarizing the state-of-the-art and providing guidance for the use of additional hydrological variables for model evaluation and calibration. Based on a review of the current literature, we summarize observation methods and uncertainties of currently available data sets, challenges regarding their implementation, and benefits for model consistency. The focus is on catchment modeling studies with study areas ranging from a few km^2 to ~500,000 km^2. We discuss challenges for implementing alternative variables that are related to differences in the spatio-temporal resolution of observations and models, as well as to variable-specific features, for example, discrepancy between observed and simulated variables. We further discuss

Paul D. Wagner and Doris Duethmann contributed equally to this work and the order of the author names was randomised.

advancements required to deal with uncertainties of the hydrological data and to integrate multiple, potentially inconsistent datasets. The increased model consistency and improvement shown by most reviewed studies regarding the additional variables often come at the cost of a slight decrease in streamflow model performance.

1 | Introduction

Catchment hydrological models serve as important tools for research and for water management, where they are applied for water resources planning, climate change impact assessments, flood warning, or drought risk assessments. The evaluation and calibration of these models mostly rely on measured streamflow data (Beven 2011). However, numerous studies have now shown that good model performance for streamflow does not necessarily translate into reliable simulations of hydrological storages or fluxes beyond streamflow (e.g., Bouaziz et al. 2021; Duethmann et al. 2014; Mei et al. 2023; Pool et al. 2024; Rajib et al. 2016; Rakovec et al. 2016; Wu et al. 2023). Consequently, multi-variable evaluation and calibration have been a way to go forward in improving process consistency and reducing parameter uncertainty. Process consistency is particularly important if the models are intended for process understanding or for predictions under changed conditions. In regions and time periods in which the lack of streamflow measurements challenges model applications, other hydrological variables can be used for model evaluation and thus also enable predictions in ungauged basins.

In recent years, an increasing amount of data on hydrological variables has become available, including data from remote sensing (Lettenmaier et al. 2015) and in situ measurements, for example, shared in networks such as Fluxnet (Pastorello et al. 2020), the International Soil Moisture Network (ISMN; Dorigo et al. 2021) or Sapfluxnet (Poyatos et al. 2021) (for a list of data sets, see Table A1). While these data are available and the advantages of multi-variable evaluation and calibration are well known, many studies still rely on single-variable calibration against streamflow. We assume that a key reason for this reliance on single-variable approaches is due to a lack of guidance on how to deal with the various challenges when integrating alternative hydrological variables into model evaluation and calibration processes.

There are several comprehensive overviews of data on additional hydrological variables available, particularly on those derived from remote sensing. Lettenmaier et al. (2015) provide an overview of the derivation of data on hydrological variables from remote sensing over the past 50 years. Moreover, Sheffield et al. (2018) review satellite data with respect to its use for water resources management in data-scarce regions, including its application for assimilation in hydrological models. Xu et al. (2014) and Jiang and Wang (2019) provide an overview of studies that used remote-sensing-derived data in hydrological modeling and their use for data assimilation (state estimation) and model calibration. However, none of these reviews comprehensively address the challenges related to the implementation of different hydrological variables and provide guidance on how to use satellite-derived and in situ data for model evaluation and calibration.

Therefore, our aim is to show how modeled and observed data can be compared for the different hydrological variables beyond streamflow to improve process representation in hydrological modeling. Particularly, we highlight crucial steps that need to be considered when comparing observed and modeled data. To this end, we report the state-of-the-art observation methods and their inherent uncertainties and describe specific challenges when using these data for model evaluation and calibration. We further discuss the benefits and future directions of including hydrological variables beyond streamflow for model evaluation. Thus, this review serves the hydrological community through state-of-the-art guidance for modelers who would like to use additional hydrological variables for model evaluation and calibration. By outlining necessary processing steps and pitfalls, implementation-related constraints of multi-variable model evaluation and calibration should be reduced.

We review catchment modeling studies for study areas of a few km^2 to around 500,000 km^2. Our focus is on model evaluation and calibration, excluding studies that update model variables in a data assimilation framework. With regard to variables, our focus is on hydrological storages and fluxes, including snow, soil moisture, groundwater level, terrestrial water storage (TWS), evapotranspiration (ET), and altimetric water level (Figure 1).

We consider hydrological observations from in situ measurements or derived via remote sensing, focusing on a large number of commonly used international and freely available data sets without the claim of being comprehensive. We exclude reanalysis data, since they are essentially simulation data, as well as data that have so far rarely been used for catchment model evaluation or calibration (e.g., ground-based gravimetry).

2 | General Considerations on the Use of Hydrological Observations Beyond Streamflow for Hydrological Model Evaluation and Calibration

2.1 | Selection of Data Sets

Before using additional hydrological variables for model evaluation and calibration, it is necessary to get an overview of the available data and select appropriate variables with respect to the study aims. A list of suitable data sets that are considered in this review is provided in the Appendix (Table A1). Based on our review of studies, some general advice should be considered when choosing a product from this list for comparison to model output, including (i) the temporal availability during the modeling period, (ii) the spatial resolution with regard to the research area and question, and (iii) the plausibility of the product based on the reported applications or evaluation with in situ data (i.e., for satellite-based products).

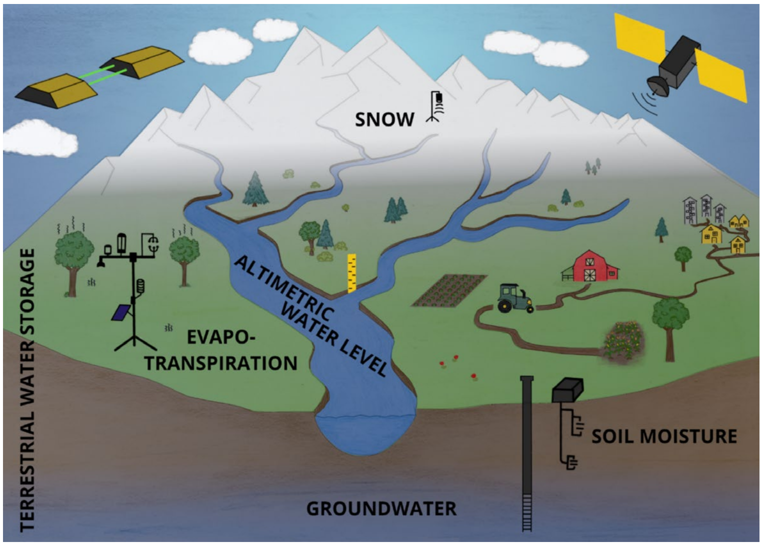

FIGURE 1 | In situ and satellite-based observations of hydrological storages and fluxes in a catchment (modified from Fisher 2005; CC BY-SA 4.0).

2.2 | Evaluation of Spatio-Temporal Model Performance With Hydrological Variables Beyond Streamflow

While streamflow integrates hydrological processes across space and time into one time series, such an integrating measure does not exist for other hydrological variables. Therefore, a key consideration when using variables beyond streamflow is the spatial representation at which the comparison should be carried out. We identify three spatial representations in models: lumped (Demirel et al. 2019; Nijzink et al. 2018), semi-distributed—like subcatchments (Alemayehu et al. 2022; Bennett et al. 2019; Hattermann et al. 2004), elevation zones (Duethmann et al. 2014) or hydrologic response units (Rajib et al. 2016)—and fully distributed, that is, grid-based pixel to pixel or pixel to observation comparisons (Corbari et al. 2022; Roy et al. 2010). A lumped comparison focuses only on the temporal performance, typically by averaging all measurements in a catchment for each time step. However, in the case of spatially sparse point measurements, it should be considered that these might not be representative for the entire catchment. For instance, while it might be a straightforward procedure for gridded ET values, spatial averaging of groundwater observations to the lumped model scale can be challenging (Seibert 2000; Staudinger et al. 2021). In principle, the same performance metrics that are applied for a single-variable calibration against streamflow can be used—unless variable-specific characteristics or biases require other metrics. For example, satellite-derived soil moisture products are considered more reliable for relative dynamics than absolute values so that the correlation coefficient is frequently used (Boeing et al. 2022; Parajka et al. 2009; Sutanudjaja et al. 2014). Specific measures may be used for binary snow cover information, and a large set of indices is available for groundwater dynamics to specifically evaluate structure (e.g., seasonality, flashiness), distribution, and shape (Heudorfer et al. 2019). In general, mostly well-known and standard hydrological performance criteria such as Kling-Gupta efficiency, Nash-Sutcliffe efficiency, or root mean square error are used (Gupta et al. 2009; Kreye et al. 2019; Nash and Sutcliffe 1970).

To account for temporal and spatial model performance, semi-distributed or fully distributed spatial representations are used. For a semi-distributed comparison, all measurements within a model unit (subcatchment, hydrologic response unit or elevation zone) are averaged, resulting in as many time series as the considered model units. When compared to streamflow calibration, this approach is comparable to a multi-gauge calibration, so that approaches commonly applied for multi-gauge calibration can be adopted. While the semi-distributed approach accounts for spatial heterogeneity to some extent, fully distributed approaches are needed to take the spatial heterogeneity fully into account. Specific performance criteria are applied that are specifically designed for comparing spatial patterns, including the SPAtial EFficiency metric (SPAEF; Koch et al. 2018; Gomis-Cebolla et al. 2022), which comprises correlation, coefficient of variation, and histogram overlap as three equally weighted components, and the Spatial Pattern Efficiency Metric (E_{SP}; Dembélé et al. 2020b). In addition, Gaur et al. (2022) use joint empirical orthogonal functions (EOF) and the fractional skill score (FSS). Koch et al. (2017) also employ variograms and assess the connectivity (CON) and multiple-point compatibility (MPC).

Moreover, a decision has to be taken on whether the data are temporally averaged before applying these metrics or if the metrics are applied for each time step and are combined or weighted afterwards, that is, assessing the temporal model performance in a spatially distributed approach and spatial model performance in a temporally distributed approach to detect where or when processes are poorly represented (Dembélé et al. 2020a; Nguyen et al. 2022).

A prerequisite for a proper evaluation of spatial performance is to account for scale differences between observed and modeled data, namely the observed or simulated data need to be upscaled or disaggregated to match the spatial resolution. To account for uncertainties in individual grid cells, a coarser resolution may be used for comparison (e.g., at subcatchment scale), assuming that over- and underestimations balance each other when averaged. Depending on the available gridded product and the spatial resolution of the model, comparisons at grid scale are carried out at different resolutions ranging from 90 m (Nguyen et al. 2022), over 1–10 km (Demirel et al. 2018; Gaur et al. 2022; Ruhoff et al. 2013) to 0.25° (Dembélé et al. 2020b) up to 3° (Yang et al. 2020). The spatial unit of comparison can often be linked to the spatial unit used by the hydrological model (e.g., subcatchment and hydrologic response unit for SWAT; grid comparison for MIKE SHE or mHM).

Parameter regionalization schemes are used to derive spatially consistent sets of model parameters (Samaniego et al. 2017). For the integration and full utilization of spatially distributed observations in multi-variable calibration, parameter regionalization plays an important role (Soltani et al. 2021). With regard to the temporal evaluation, it should be noted that hydrological variables beyond streamflow are not necessarily available at daily time steps, so that evaluations are carried out at different time scales ranging from daily (e.g., Rajib et al. 2018), 8-day averages (e.g., Ruhoff et al. 2013), monthly (e.g., Odusanya et al. 2019; Jing et al. 2018), seasonal (e.g., Bouaziz et al. 2021) to mean annual values (e.g., Bouaziz et al. 2018). Also, while streamflow data from several years are usually included in model calibration and evaluation, the length of the observation period for other hydrological variables could vary considerably from 1 year (Ruhoff et al. 2013: 8-day averages) to several years (Sirisena et al. 2020: monthly data for 13 years). The highest temporal resolution available is typically used, particularly if spatial patterns are not considered (e.g., daily: Ala-aho et al. 2017; daily and monthly: Zhang et al. 2009). Also, the evaluation of temporal dynamics can be focused on time periods in which these variables show a strong variability; for example, snow is evaluated for days with snowfall (Széles et al. 2020), the snow melt season (Finger et al. 2015), or the entire year (Corbari et al. 2022; Tong et al. 2022).

2.3 | Multi-Variable Model Calibration

Calibrating a model involves determining how the model can be constrained across multiple variables. One of the approaches commonly employed considers step-wise calibration, where the model is first constrained by one variable and then further constrained by one or more other variables (e.g., Széles et al. 2021). Alternatively, the model may be calibrated simultaneously to different hydrological variables. In this case, different objective functions may be combined into a single objective function, for example, by applying a weighted average to the individual functions for the different variables. Subsequently, calibration algorithms for single objective functions can be used (e.g., Rajib et al. 2016; Dembélé et al. 2020b; Odusanya et al. 2019; Parajka et al. 2006; Pool et al. 2024). A disadvantage is the subjectivity in setting the weights (Efstratiadis and Koutsoyiannis 2010). Several authors show the importance of carefully choosing the weights assigned to the individual error metrics (Ala-aho et al. 2017; Tarasova et al. 2016; Tong et al. 2022; Werth and Güntner 2010) and it is recommended to perform prior sensitivity analyses with varying weights. While trade-offs between the different objectives can be analyzed to some extent by varying the weights, this is not always done, and it is also not a very efficient way to approximate the Pareto front. Optimization algorithms that sample multiple objective functions individually are particularly suited to explicitly analyze the Pareto front. For sampling the Pareto front, two groups of approaches exist. The first group comprises approaches based on Monte Carlo simulations (e.g., Silvestro et al. 2015; Nguyen et al. 2022; Kuppel et al. 2018), where the parameter space is randomly sampled and parameter sets are ranked based on model performance with regard to pre-defined objective functions. Advantages of the Monte Carlo method are its simple implementation, its ability to explore the parameter space for many different objective functions, and that parameter uncertainties can easily be accounted for. However, the Monte Carlo approach is not very efficient and, compared to iterative optimization algorithms, it may result in lower model performances despite a higher number of model evaluations (Duethmann et al. 2014). The second group of approaches is multiobjective algorithms. These algorithms return a set of solutions approximating the Pareto front with one optimization run and are thus very well suited to investigate trade-offs between different objective functions. Multiobjective algorithms used for multi-variable calibration include, for example, NSGAII (Nondominated Sorted Genetic Algorithm II, Deb et al. 2002; for example, in Bai et al. 2018; Werth and Güntner 2010), ε-NSGAII (Kollat and Reed 2006; for example, in Duethmann et al. 2014), SPEA-II (Strength Pareto Evolutionary Algorithm II; Zitzler et al. 2001), or PADDS (Pareto-Archived Dynamically Dimensioned Search, Asadzadeh and Tolson 2013; for example, in Mei et al. 2023). Using approaches that analyze the Pareto front allows for a more informed decision on how to select parameter sets from the Pareto optimal set. In the case of trade-offs, a single best parameter set does not exist, and, if only one parameter set can be selected, this choice should depend on the modeling objectives. By selecting an ensemble of parameter sets that perform well for different objectives, parameter uncertainty can be accounted for in an efficient way.

Irrespective of which algorithm and approach is used for multi-variable calibration, it is advisable to perform multi-variable sensitivity analysis prior to the calibration (Pianosi et al. 2016; Rosolem et al. 2012). Two important aspects regarding the parameters need to be considered. First, the range of parameters needs to be selected so that it covers the optimum for all objective functions for which the parameter is sensitive (Guse et al. 2020). For parameters that are sensitive for more than one objective, this is extremely important in order not to pre-select which objective function will be simulated well. Second, different numbers of parameters will be sensitive for different objective functions (e.g., Rosolem et al. 2012; Yassin et al. 2017). This is directly related to the parametrization of hydrological models that simulate different compartments with different levels of detail. For example, hydrological models typically use fewer parameters for interception than for runoff (e.g., Samaniego et al. 2010; Seibert and Bergström 2022). If a variance-based sensitivity analysis is used (Sobol 2001), then the percentage of variance in an objective function will be quantified for individual parameters or groups of parameters (Mai et al. 2020). Subsequently, parameters can be selected so that the explained variance in all objective functions is covered. This will result in a more efficient calibration compared with a selection that over- or underrepresents objective functions.

3 | Variable-Specific Considerations on Using Hydrological Observations Beyond Streamflow for Hydrological Model Evaluation and Calibration

3.1 | Snow

3.1.1 | Observation Methods, Uncertainties and Data Products

A large number of hydrological modeling studies recognize the pivotal role of snow for streamflow generation, especially in mountainous and cold regions, and thus explicitly consider snow cover (SC), snow depth (SD), or snow water equivalent (SWE) data from in situ observations, airborne surveys, or satellites for model evaluation or calibration. SC data can be extracted from optical satellite data, such as MODIS (e.g., applied in Bennett et al. 2019; Nijzink et al. 2018; Tarasova et al. 2016; Tong et al. 2022) or AVHRR (e.g., Dressler et al. 2006; Duethmann et al. 2014). SC data have further been retrieved from in situ field cameras (e.g., Schöber et al. 2010; Széles et al. 2020). While readily available SC products exist for MODIS (provided by the National Snow and Ice Data Center), data from AVHRR or Landsat need to be processed, which involves the calculation of radiances and brightness temperatures, corrections for atmospheric conditions and sensor degradation, classification, and georeferencing (Dressler et al. 2006; Duethmann et al. 2014; Ragettli et al. 2015; Schöber et al. 2010). Classification algorithms typically employ variables including the normalized difference snow index (NDSI), normalized difference vegetation index (NDVI), surface temperature, and albedo. SC data are generally accurate, with the mean absolute error of fractional SC derived from MODIS generally < 10% (Appel 2018). However, uncertainties can be higher in forested areas and under overcast skies (Parajka et al. 2012; Gascoin et al. 2015). Space-borne radar systems (i.e., Synthetic Aperture Radar, SAR) are a further data source for SC data, providing data at high spatial resolution. However, the application of SAR data for hydrological modeling is very limited due to its complex processing and low temporal resolution (Tsai et al. 2019).

SD may be retrieved from in situ snow observations (e.g., Bongio et al. 2016; Corbari et al. 2022; Tuo et al. 2018), airborne ground penetrating radar (e.g., Huss et al. 2014) or microwave-based satellite observations (Xiao et al. 2020), including Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), Advanced Microwave Scanning Radiometer-2 (AMSR2) or Global Snow Monitoring for Climate Research (GlobSnow), where the latter is based on satellite microwave data and station data. Uncertainties of satellite-derived SD data can be considerable, with generally best performance in plain areas and worst in forested mountain regions (Xiao et al. 2020). Uncertainties further vary depending on the observed SD. More details on the advantages and limitations of remote sensing and in situ observations of snow can be found in the review by Dong (2018).

SWE is typically measured in situ, for example, manually with snow tubes or automatically with snow pillows or gamma/cosmic ray sensors. An extensive network of 890 snow monitoring sites (SNOTEL) with SWE observations from snow pillows of up to 60 years is available in the US (Fleming et al. 2023) but SWE measurements are often only sparse in other parts of the world. Snow pillows have a precision of approximately 2.5 mm (Fassnacht et al. 2017), and measurement errors are typically between 5% and 15% (Serreze et al. 2001), but can exceed 40% when the base of the snow cover is at the melting temperature (Johnson and Marks 2004). While automatic SWE measurements with stationary gamma ray sensors are costly and have very limited spatial extent (Pirazzini et al. 2018), there is a huge potential for using airborne gamma radiation snow surveys (Cho et al. 2023). This technique is based on the difference of the radioisotope rates between bare and snow-covered land surfaces (Cho et al. 2020) and is able to provide highly accurate SWE estimates at a large scale even for forested areas and wet snow conditions (i.e., melting periods) known to be particularly challenging for other remote sensing tools (Tuttle et al. 2018). Such surveys have been available throughout the United States and Canadian provinces since 1979 (Carroll 2001).

3.1.2 | Challenges When Comparing Simulated and Observed Data

Comparing observed SC data to hydrological model data is not straightforward since many hydrological models only simulate SWE and not SC. Simulated SWE is therefore often converted to SC by assuming that a modeling entity is entirely snow covered if SWE is larger than a user-defined threshold (Table 1). These thresholds typically range from 0.5 to 10 mm (Gyawali and Bárdossy 2022; Peker and Sorman 2021; Széles et al. 2020; Di Marco et al. 2021; Tong et al. 2022), though larger values are also used (50 mm; Corbari et al. 2022). SC data may further come in the form of fractional SC data, also as a result

TABLE 1 | Practical considerations for evaluating simulated with observation-based snow data.

Data source	Considerations
Satellite-derived	Comparison of satellite-derived snow cover (SC) to simulated snow water equivalent (SWE): • Conversion of simulated SWE to SC requires definition of a threshold (0.5–10 mm) above which a modeling entity is considered as snow covered. Dealing with clouds: • Definition of a minimum threshold of cloud-free fraction of a modeling entity above which the data are considered • (Partly) removing clouds from the data set before model evaluation
In situ	Comparison of observed snow depth (SD) to simulated SWE • Conversion of observed SD to SWE using empirical functions • Conversion of observed SD and simulated SWE to SC through a user-defined threshold above which a modeling entity is considered as snow covered

of summarizing binary SC data to model entities. As most models neglect the spatial variability of SWE within model entities, a threshold is often introduced that defines when a pixel or model entity is considered snow covered (Parajka and Blöschl 2008a). Reported values ranged from 10% (Bouaziz et al. 2021) to 25% (Tong et al. 2022). Alternatively, the variability of SWE within a model entity may be described by a separate subgrid snow distribution parameterisation (e.g., Liston 2004), thus enabling a direct comparison between simulated and observed fractional SC (Duethmann et al. 2014). Since optical sensors cannot record snow under clouds, uncertainties increase with cloud coverage (Table 1). Cloud-cover thresholds above which data for a model unit are excluded vary widely, ranging from 10%–60% (Di Marco et al. 2021; Finger et al. 2015; Duethmann et al. 2014; Gyawali and Bárdossy 2022; Tong et al. 2022; Nijzink et al. 2018; Tarasova et al. 2016). Since clouds can reduce data availability considerably, cloud removal procedures that make use of merging MODIS Aqua and Terra and spatial and temporal filters (Gafurov and Bardossy 2009; Parajka and Blöschl 2008b) are sometimes applied (Parajka and Blöschl 2008a; Peker and Sorman 2021; Gyawali and Bárdossy 2022). Recent developments in cloud removal approaches are reviewed by Li et al. (2019b). To analyze uncertainties, sensitivity analyses have been carried out for the threshold for converting simulated SWE to SC (Gyawali and Bárdossy 2022; Peker and Sorman 2021; Sleziak et al. 2020; Széles et al. 2020; Tarasova et al. 2016) or the cloud cover threshold (Gyawali and Bárdossy 2022; Parajka and Blöschl 2008a; Tarasova et al. 2016).

A challenge related to point-based in situ SD observations is their poor representativeness resulting from high spatial variability (López-Moreno et al. 2011; Miller et al. 2022). To enable a comparison to simulated data, SD data are typically converted to SWE. However, this conversion can be very uncertain (Hill et al. 2019). Usually, empirical relationships are employed (Bongio et al. 2016; Tuo et al. 2018), though one could also use snow density models, which are common in snow research (Winkler et al. 2021; Fontrodona-Bach et al. 2023). An alternative way of using in situ SD observations is conversion to SC after spatial interpolation so that they can be treated as SC observations (Parajka et al. 2007).

3.2 | Soil Moisture

3.2.1 | Observation Methods, Uncertainties and Data Products

Since soil moisture controls runoff generation and ET, model evaluation against observed soil moisture can improve process consistency. In situ soil moisture measurements at the point scale can provide temporally highly resolved data. Measuring techniques range from traditional gravimetric methods to soil water sensors, such as time domain reflectometry (TDR) and frequency domain reflectometry (FDR) (Graeff et al. 2010; Robinson et al. 2008). Although these methods are among the only techniques that allow soil moisture measurements in deeper soil layers at high accuracy (gravimetric method: errors of $<1\%$; soil water sensors: accuracies of ~$0.02\,m^3\,m^{-3}$, but offsets up to $0.2\,m^3\,m^{-3}$, depending on sensor type, soil and land-use properties; Robinson et al. 2008; Jackisch et al. 2020), their signals are only representative across a small volume of a few cubic centimeters to decimeters. Labor-intensive installation of in situ soil moisture measurement networks and long-term maintenance demands significant economic and human resources (Dorigo et al. 2013). Moreover, measurements are often only available at research sites. The International Soil Moisture Network (ISMN, Dorigo et al. 2021, Table A1) maintains a global database of in situ soil moisture data, with high station density in the US, lower density in Europe and Asia, and currently very low density in South America and Africa. Recently, in situ data have also been used to generate gridded soil moisture products through deep learning techniques (SoMo.ml, Orth 2021, Table A1).

At field scale (>100 m), Cosmic-Ray Neutron Sensors (CRNS) offer a new technique that allows continuous measurements of the average soil water content over a horizontal footprint of a few hundred meters to a depth of ~70 cm without soil disturbance. The method is based on detecting neutrons in the soil generated by high-energy cosmic radiation, where the intensity of neutrons is dependent on the soil water content (e.g., Iwema et al. 2015). CRNS soil moisture agrees well with in situ soil moisture data, with RMSE values around 0.02 to $0.06\,m^3\,m^{-3}$ and uncertainties around 10% (Baroni et al. 2018). CRNS data have been made available by networks for the U.S. (Zreda et al. 2012), Europe (Bogena et al. 2022), or Australia (Hawdon et al. 2014), and can also be retrieved from the ISMN.

Satellite-derived soil moisture based on active and passive microwave sensors provides data at large spatial scales of a few kilometers (see Table A1). Microwave remote sensing of soil moisture uses the sensitivity of surface emissivity and backscattering properties of the soil for microwaves to soil dielectric properties and thus the soil water content. The main limitation is the low penetration depth of the microwave signals, so that only near-surface soil moisture can be derived. Furthermore, while absorption and scattering of microwave radiation by vegetation can partly be accounted for, soil moisture retrieval under dense vegetation such as forests remains challenging. In addition to soil moisture products based on single sensors, products that combine multiple sensors have been developed. This enables covering a longer time period (e.g., the ESA-CCI soil moisture product; Gruber et al. 2019) or downscaling coarse-scale data to a finer resolution by fusion with higher resolution SAR data (e.g., the SCATSAR-SWI product; Bauer-Marschallinger et al. 2018; Peng et al. 2017). Criteria for selecting products should include their quality in terms of performance against observed data, which may vary by region, climate, vegetation type, or topography (Beck et al. 2021). For example, SMAP data outperformed ESA-CCI, AMSR2, ASCAT, and SMOS data in a comparison to in situ measurements from 826 sensors located mostly in the USA and Europe (Beck et al. 2021). Available in situ data further enable testing the performance of various satellite products for a specific study region before their application in hydrological modeling (Duethmann et al. 2022). Comparison studies of satellite-derived near-surface soil moisture products to in situ data over different parts of the world report average Pearson correlation coefficients of 0.4–0.8 and average RMSE values of $0.04–0.13\,m^3\,m^{-3}$ (Al-Yaari et al. 2019; Beck et al. 2021; Fang et al. 2016).

3.2.2 | Challenges When Comparing Simulated and Observed Data

In situ point scale soil moisture measurements have been widely employed in modeling catchment hydrology for over three decades, primarily for model and process evaluation (e.g., Bronstert and Plate 1997), but also as a direct modeling objective (Loritz et al. 2022). Data from sensors at various soil depths and different locations are either compared directly to simulated data—if the model incorporates a spatially and temporally explicit soil routine, common in more process-based models (e.g., Ebel et al. 2007)—or in a normalized, aggregated manner in more conceptual models (e.g., Boeing et al. 2022). In situ soil moisture measurements can exhibit large biases, even when taken in close proximity, the same soil depth, and in similar hydro-pedological conditions. Data aggregation from several soil moisture sensors can compensate for small-scale variabilities. Additionally, it is often advised to evaluate the model's performance specifically with respect to its dynamics, which can be achieved by scaling the data (e.g., through normalization), or by focusing on exceedance probabilities (e.g., Loritz et al. 2017; Mälicke et al. 2020) as for instance done in the German drought monitor (Zink et al. 2016). Due to the larger footprint, CRNS soil moisture data are less affected by small-scale variability. However, the dependency of the footprint radius and measurement depth of soil moisture, air humidity, and vegetation (Köhli et al. 2015) can be a difficulty.

Comparing satellite-derived and simulated soil moisture is frequently challenged by a mismatch in vertical scale, as only near-surface soil moisture can be inferred from the satellite signal. Satellite-derived near-surface soil moisture can directly be compared with simulated surface soil moisture if the model already contains (López López et al. 2017; Rajib et al. 2016) or is complemented with a surface soil layer of only a few centimeters (Parajka et al. 2009; Kundu et al. 2017). Alternatively, near-surface soil moisture may be converted to a root zone soil water index (SWI) by temporal smoothing with an exponential filter (Wagner et al. 1999; Table 2). Calculating the SWI is equivalent to a simplified two-layer water balance model where the upper layer represents the remotely sensed topsoil layer and the soil moisture variation in the lower layer is a linear function of the soil moisture difference. This model has one parameter, the characteristic time length T, which depends on the depth of the soil layer and soil hydraulic properties (Wagner et al. 1999; for guidance see Paulik et al. 2014 and Bouaziz et al. 2020). It should be noted that the SWI approach relates to a simplified model that may be inappropriate in some cases, such as in the case of groundwater influences or when surface soil moisture evaporates back into the atmosphere without infiltrating to the root zone layer (Bauer-Marschallinger and Paulik 2019).

Scale differences between satellite-derived and simulated soil moisture also occur at the horizontal scale, since many microwave-based soil moisture products are provided at grid spacings of 25 km or more, which are usually larger than model units in catchment hydrological models. As for in situ data, systematic biases between simulated and satellite-derived soil moisture are frequently observed, and the data are often scaled to focus on the temporal dynamics (Table 2). Approaches used include z-scores or mean–variance scaling (Cammalleri et al. 2015; López López et al. 2017; Rajib et al. 2016; Duethmann et al. 2022), min-max scaling (Hrachowitz et al. 2021) or cumulative density function matching (Sutanudjaja et al. 2014). Further pre-processing steps may require conversion to the same unit, for example, by dividing the simulated soil water volume by the maximum storage (Parajka et al. 2009), or by converting satellite-derived volumetric soil moisture to plant available water (Rajib et al. 2016).

TABLE 2 | Practical considerations for evaluating simulated with observation-based soil moisture data.

Data source	Considerations
Satellite-derived	Using surface soil moisture versus a root zone soil water index Using surface soil moisture: • May require model modifications (in case a near surface soil moisture layer does not exist in the model) and additional parameters Using soil water index data: • Root zone soil moisture more relevant for runoff generation and ET than surface soil moisture • Involves additional assumptions (soil moisture in root zone as linear function of difference between surface and root zone soil moisture) • Requires selection of the parameter T (characteristic time length; dependent on soil depth and soil hydraulic properties)
Satellite-derived and in situ	Scaling of simulated and observed soil moisture helps dealing with uncertainties in absolute soil moisture (min-max scaling, mean–variance scaling, cumulative density function matching, focus on exceedance probabilities)

3.3 | Groundwater Level

3.3.1 | Observation Methods, Uncertainties, and Data Products

Groundwater is conceptualized as the deepest part of the subsurface in hydrological models and is most commonly characterized by smaller fluctuations than the other components of the hydrological cycle. Because of its accessibility, groundwater is difficult and complex to measure.

Groundwater level data stem from groundwater wells or, more precisely, from the measurement of hydraulic groundwater heads (de Graaf et al. 2019). Outside of experimental catchments, groundwater level data are measured non-uniformly in space and may only be available at weekly or even lower temporal resolution (Reinecke et al. 2024). However, the low-resolution data can still be useful for modeling due to the relatively small fluctuations in groundwater, as Jing et al. (2018) show for monthly data. The representativity of groundwater level measurements is influenced by outliers, site-specific aquifer properties, and the connectivity of groundwater beyond the catchment boundary (Liu et al. 2020). Additionally, observations of groundwater level may be influenced by groundwater extraction, groundwater injection, and sewer leakage. However, data regarding such direct anthropogenic impacts are only available in some regions (Gleeson et al. 2021).

3.3.2 | Challenges in the Comparison of Simulated and Observed Data

Whereas observations of the groundwater level are usually provided as hydraulic heads, many hydrological models that represent groundwater in a simplified way only generate time series of groundwater storage (Gleeson et al. 2021). Thus, unless the dynamics of groundwater heads are used directly, groundwater level data need to be transformed into groundwater storage information or vice versa simulated storage into groundwater levels (e.g., through linear regression, Seibert 2000). Moreover, groundwater well observations are point-based and do not always represent the actual spatial heterogeneity or the average groundwater level of the catchment, whereas model results are provided for continuous space, that is, at the spatial model resolution (Table 3).

The degree of complexity for representing groundwater in hydrological models is generally lower compared to other hydrological components. Information on geological structures, for instance from geophysical data or derived from geological maps, are usually not directly included in hydrological models due to the missing or only weak link between geological structures and the conceptual groundwater components (Staudinger et al. 2019). Groundwater storage is often highly simplified and simulated as one bucket or a series of buckets (McMillan et al. 2023; Staudinger et al. 2019; Zipper et al. 2023) and optimal

TABLE 3 | Practical considerations for evaluating simulations with observation-based groundwater data.

Data source	Considerations
In situ	• Spatial distribution of point measurements varies and they might not always represent the groundwater heterogeneity of the system • Coarse temporal resolution of the observed data (e.g., monthly) is still useful, due to small variations • Model structure is crucial for the comparison to in situ groundwater level data, e.g., simplified bucket representation, undefined or fixed depth of groundwater component • Groundwater head measurements at groundwater wells need to be converted to groundwater storage unless the dynamics can be used directly.

groundwater model structures have been found to be aquifer-specific (Stoelzle et al. 2015). However, some hydrological models represent groundwater with more complexity (e.g., SHE, Abbott et al. 1986) by using a more detailed spatial discretization. Many challenges are related to how groundwater is represented in hydrological models:

1. In catchment models with a conceptual groundwater storage, the groundwater component is not assigned to an absolute depth (distance to the surface) (Pfannerstill et al. 2014; Wu et al. 2022). Therefore, groundwater dynamics are often only considered in relative rather than absolute values (Kuczera and Mroczkowski 1998; Seibert 2000; Fenicia et al. 2008). Groundwater level data cannot be directly included in the model if the model only considers root zone depth (Kreye et al. 2019). A further challenge is that the boundary between a soil (or unsaturated) and a groundwater (or saturated) zone is dynamically changing, whereas, in hydrological model structures, it is mostly fixed.
2. In addition, the interaction of groundwater and surface water is not accounted for in most bucket-type hydrological models (Staudinger et al. 2021). One exception is the WALRUS model that includes groundwater-surface water interactions (Brauer et al. 2014) and this concept has also been adopted for the HBV model (Staudinger et al. 2021) to allow for back-infiltration of water from the river to the aquifer. In addition, coupled modeling approaches are available for representing groundwater-surface water interactions (SWAT-MODFLOW, Bailey et al. 2016; Tigabu et al. 2020; SWAT+ gwflow, Bailey et al. 2020).
3. Most hydrological models assume that the spatial extent of the catchment borders at the surface is identical to the aquifer extent, which is rarely the case (Bouaziz et al. 2018; Liu et al. 2020; Gleeson et al. 2021). Interbasin groundwater flow from and to adjacent catchments as well as leakage below the gauging station (Käser and Hunkeler 2016) are not covered in most hydrological models (Merz et al. 2020). Le Moine et al. (2007) model interbasin groundwater flow via subsurface catchment boundaries that differ from the ones derived from surface topography.

There are also studies that take streamflow data as the alternative data to constrain models by using streamflow data to constrain water table depth for wetlands in the Penn State Integrated Hydrologic Model (PIHM; Yu et al. 2015), streamflow intermittency data to better estimate hydraulic conductivity and porosity (Abhervé et al. 2024), and baseflow indices derived from streamflow observations to facilitate a more prominent representation of groundwater in Earth system models (Xie et al. 2024).

3.4 | Terrestrial Water Storage

3.4.1 | Observation Methods, Uncertainties and Data Products

Terrestrial water storage (TWS) is the volume of water stored within a catchment, and its distribution in ice, snow, lakes, reservoirs, rivers, soil, and groundwater ultimately characterizes the state of the hydrological system (McNamara et al. 2011). Unfortunately, TWS is inherently difficult to measure. Moreover, different methods to determine the depth of stored water through different interpretations of so-called active storage (i.e., referring to zones that fill and release water on time scales relevant to annual input and output fluxes) versus total storage depth result partly in similar (Rodhe et al. 1996) and sometimes in different (Soulsby et al. 2009) estimates of storage. When using storage estimates for model evaluation, these definitions need to be considered.

Stable water isotopes are useful tracer data to constrain water storage in landscapes, particularly when used in conjunction with soil moisture and groundwater level data. The passage of conservative environmental tracers (e.g., water isotopes, chloride) through the landscape can be monitored to quantify how precipitation input signals are damped and lagged by internal mixing processes and connections between different spatial units, including deeper groundwater (Tetzlaff et al. 2014). Thus, tracers provide measures of the internal storage in terms of the water volume needed to dampen the tracer input signal (Soulsby et al. 2011). However, tracer-aided hydrological models are required for model evaluation with tracer data (Tetzlaff et al. 2015, 2024). Stadnyk and Holmes (2023) provide a thorough review of such models and their applications.

As another data source to estimate TWS, the gravity recovery and climate experiment (GRACE) and GRACE-follow on (GRACE-FO) twin satellites provide information on temporal changes in the Earth's gravity field. The mass redistribution in the Earth system can be attributed—besides other geophysical processes like post-glacial land uplift in northern Europe, air mass changes, tides, and sea level fluctuations (Güntner et al. 2023)—to changes in the TWS on and in the Earth's crust. Multiple reviews provide overviews of using GRACE data in hydrology (Frappart and Ramillien 2018; Li et al. 2019a; Chen et al. 2022) including background information, recommendations, pitfalls, and examples (Humphrey et al. 2023). GRACE data are available from 2002 with a gap between July 2017 (when the first GRACE satellites were decommissioned) and May 2018 (when the GRACE- FO satellites became operational). At best, GRACE-derived TWS changes have a resolution of 300 km × 300 km on a monthly time scale. According to Güntner et al. (2023), the accuracy of GRACE data is in the range of 10 to 20 mm water equivalents, but individual assessments are required depending on the study characteristics. Storage changes derived from GRACE may be compared with simulated storage changes and data from groundwater wells to assess uncertainties in GRACE storage changes, though such a comparison to in situ data is challenging due to scale differences and conceptual differences in the compared variables (Schumacher et al. 2018).

Two further approaches to evaluate TWS are geophysical methods such as seismic and resistivity measurements (Wiederhold et al. 2021) and in situ gravimetry measurements (Pool and Eychaner 1995). Since these methods are not frequently applied and are only rarely used for model calibration (Creutzfeldt et al. 2010), they are not further considered in this review.

TABLE 4 | Practical considerations for evaluating simulated with observation-based terrestrial water storage (TWS) data.

Data source	Considerations
Satellite-derived	• Compare sufficiently large areas (ideally > 200,000 km^2 approximating a circular or squared area) • GRACE data provide changes in TWS to a baseline and the hydrological model baseline should match that baseline period or be corrected accordingly. • TWS includes storages from all components and the hydrological model must be capable of simulating these components. • Consider application of scaling factors provided with Level 3 products for reducing the leakage effect
Stable water isotopes	• Requires tracer-aided hydrological model • Assessment based on one output isotope variable is possible (i.e., isotopes from streamwater, groundwater, or soil water). • Streamwater isotopes integrate the different storages (i.e., isotopic signals) in the landscape and different temporal resolutions (daily, weekly, monthly) can be employed • Isotope data in precipitation and rivers are available from international data archives

3.4.2 | Challenges in the Comparison of Simulated and Observed Data

All components of TWS (e.g., deep groundwater) and processes that impact water storages (such as deep groundwater extraction) should be included in the hydrological model (Table 4) for comparison of storage dynamics and trends, particularly over longer time periods (Stampoulis et al. 2019; Huang et al. 2019). To disentangle the individual components of TWS, these can be combined with other remote sensing-based data sets for snow, soil moisture, and surface water storages (Werth and Güntner 2010).

Simulated tracer concentrations are compared to time series of monitored tracer data of different landscape compartments. Only one output isotope variable is required for model assessments, that is, isotopes from streamwater, groundwater, or soil water. A major challenge is whether the correct source water has been sampled (Birkel and Soulsby 2015). To ensure this, knowledge about catchment functioning, in terms of which and how spatially distributed and temporally dynamic source waters mix and integrate, is crucial to quantifying TWS. Due to technical and cost constraints, tracer data at daily resolution for prolonged time periods are still limited to a few research sites (Tetzlaff et al. 2015). However, even weekly or monthly isotope data are still extremely insightful and comparably easy and cheap to sample and analyze. In addition, isotope data in precipitation (Global Network of Isotopes in Precipitation, GNIP) and rivers (Global Network of Isotopes in Rivers, GNIR) are available from international data archives (Table A1).

Challenges using GRACE data are manifold (Chen et al. 2022; Humphrey et al. 2023):

1. GRACE data are usually expressed as anomalies to the long-term mean over a baseline period. This period must agree with the baseline in the hydrological model for comparing time series of anomalies (Humphrey et al. 2023) (Table 4).
2. Mass changes derived from GRACE are generally based on two approaches, spherical harmonics and mass concentration blocks (mascons), and are affected by a "leakage" of the signal, causing the observed mass changes to be blurred, scattered, and smeared over the true location due to spatial smoothing that is required for noise reduction (Tripathi et al. 2022). The leakage effect is particularly relevant for TWS data derived with spherical harmonics solutions. Therefore, the hydrological model domain should be $> 200{,}000\,\text{km}^2$ (ideally approximating a square or circle). In contrast, results from the mascons processing are less affected by leakage effects and, therefore, hydrological model results can be more directly compared within the 3° native mascon resolution by aligning the hydrological model domain and resolution (Humphrey et al. 2023) (Table 4).
3. GRACE products are provided by several processing centers (Table A1), resulting in differences in the products with specific advantages and disadvantages that need to be taken into account (Güntner et al. 2023; Humphrey et al. 2023). Data from different GRACE processing centers can be used to assess uncertainties (Werth and Güntner 2010; Schumacher et al. 2018).
4. GRACE data are provided at different processing levels (see Behzadpour 2022). For modelers without specific expertise in GRACE data processing, level 3 data, which only contain TWS anomalies, are most suitable and may directly be used for evaluating TWS variations in hydrological models. Level 3 data can contain scaling factors that may reduce leakage effects (Humphrey et al. 2023). Level 2 products require considerably more processing, including correction for signal losses (Bai et al. 2018), spatial smoothing with filters (Milzow et al. 2011; Schumacher et al. 2018; Werth and Güntner 2010), error propagation, signal separation, and applying scaling factors (Schumacher et al. 2018; Werth and Güntner 2010). It is important to test different smoothing filter techniques and parameterisations of those, such as the radius that defines the spatial or temporal range over which data is averaged or smoothed, since they can lead to significantly different results (Lo et al. 2010; Güntner et al. 2023).
5. The applicability of GRACE data in small areas and for short time periods is limited due to its coarse spatial and temporal resolution, with most studies located in basins $> 100.000\,\text{km}^2$ and only a few in smaller mesoscale catchments (Nijzink et al. 2018; Bouaziz et al. 2021; Bai et al. 2018). Most commonly, the average of the GRACE grids is compared with the complete model domain following a lumped approach (e.g., Milzow et al. 2011; Lo et al. 2010; Hulsman et al. 2021a; 2021b; Bouaziz et al. 2021; Dembélé et al. 2020a, 2020b; Nijzink et al. 2018). However, not only (spatially averaged) TWS data but also long-term changes in TWS or other statistical properties (such as minimum (dry) or maximum (wet) TWS) are evaluated.

3.5 | Evapotranspiration

3.5.1 | Observation Methods, Uncertainties, and Data Products

Evapotranspiration (ET) is one of the most important water balance components. Compared with the other components, ET and its components, interception, transpiration, and soil evaporation are difficult to measure directly. Lysimeters can be used to indirectly determine ET by measuring the other components of the water balance equation. Sap flow measurements provide a measure of transpiration but need correct estimates of stand characteristics. Moreover, micrometeorological measurement techniques (Bowen ratio-energy balance, scintillometers, eddy covariance) are used to measure ET. Due to their respective footprints, these measurements are often land use and land cover specific. Further, water stable isotopes sampled in open surfaces but also plant xylem waters are useful to quantify evaporation and transpiration processes—as only mixing and evaporative fractionation change the isotopic signals of landscape compartments (Kleine et al. 2020; Luo et al. 2024; Sprenger et al. 2022). In situ measurements of ET are far more expensive than measurements of streamflow, and the spatial density of these measurements is therefore rather sparse.

Measured ET fluxes from eddy covariance stations (Kunnath-Poovakka et al. 2018; Lin et al. 2018; Stisen et al. 2018; Pan et al. 2022; Széles et al. 2020) and transpiration of forests from sap flow measurements (Kuppel et al. 2018; Smith et al. 2021) have been employed to evaluate or calibrate (eco-)hydrological models. Stable isotopes are used in tracer-aided (eco)-hydrological models and allow temporal and spatial dynamics of evaporation and transpiration fluxes to be quantified (e.g., Neill et al. 2021; Yang et al. 2023). Kite and Droogers (2000) underline that the choice of observational ET data depends on the study purpose; for example, in situ measurements are more suitable for small-scale studies and remote sensing-based ET for large-scale studies.

Still, spatially distributed products derived from satellite data are most commonly used. Several ET retrieval algorithms to estimate ET based on remote sensing data like land surface temperature and landscape characteristics have been developed (see Lettenmaier et al. (2015) for details). For example, surface energy balance algorithms (SEBAL, SEBS) can be applied to satellite data in individual studies (e.g., Immerzeel and Droogers 2008; Rientjes et al. 2013). Other retrieval methods incorporate the Penman-Monteith (MOD16 ET, Mu et al. 2011) and Priestley–Taylor (GLEAM, Miralles et al. 2011a) formulas to estimate potential ET and use different approaches to derive actual ET from multi-spectral (MODIS ET, Table A1) and microwave (GLEAM, Table A1) remote sensing data. Particularly, the processed, readily available products from MODIS (e.g., Gaur et al. 2022; Jiang et al. 2020; Koch et al. 2022) and GLEAM (e.g., Bouaziz et al. 2021; Odusanya et al. 2019; Sirisena et al. 2020) are most commonly used for hydrological model evaluation against ET data.

Annual estimates of global terrestrial ET sum up to 62.8×10^3 km^3 for MODIS ET (Mu et al. 2011) and to 67.9×10^3 km^3 for GLEAM ET (Miralles et al. 2011a), slightly underestimating and overestimating other reported estimates of $65.5\times10^3 km^3$ (Oki and Kanae 2006). When compared to eddy covariance measurements, MOD16 ET shows a mean absolute error of 0.33 mm/day for daily ET estimates (Mu et al. 2011) whereas performance differs with regard to land cover (best performance for forest for an evaluation over Asia, Kim et al. 2012). GLEAM ET has an average correlation with eddy covariance measurements of 0.83 on a daily and 0.90 on a monthly basis (43 stations, 1 year of measurements, Miralles et al. 2011b). SEBS performs overall well for different site conditions with lower performance over tall and heterogeneous canopies (20 stations, 1.5–10 years of measurements, Ershadi et al. 2014). On average, remote sensing estimates of daily ET differ from eddy covariance measurements by 1.18 mm/d (mean RMSE, review of 348 studies by Tran et al. 2023). However, eddy covariance measurements may also be greatly biased (Miralles et al. 2011b), for example, due to footprint heterogeneity and non-closure of the energy balance (Klosterhalfen et al. 2023), which is, with up to 10%–20%, a major source of uncertainty for these measurements (Stisen et al. 2018). Some authors provide plausibility checks of the data used, including a comparison of satellite-derived ET to in situ pan-evaporation (Pan et al. 2022), to eddy covariance data (Ruhoff et al. 2013), or in the absence of measured data to other ET products (Odusanya et al. 2019; Nguyen et al. 2022) or to (regional) evaluations of the product or algorithms in the literature (Poméon et al. 2018; Rientjes et al. 2013).

Actual ET, that is, the combined evaporative fluxes from vegetation, soil, and water, is most commonly used for evaluation and calibration of catchment models (e.g., Dembélé et al. 2020b; Hulsman et al. 2021a; Jiang et al. 2020). Bouaziz et al. (2021) also compare interception evaporation to GLEAM data. ET calculation in models is mostly based on one approach (often Penman–Monteith, Monteith 1965). However, Odusanya et al. (2019) have evaluated different methods to calculate potential ET (Hargreaves, Priestley–Taylor, Penman–Monteith) with regard to the best fit of simulated and observed actual ET.

3.5.2 | Challenges in the Comparison of Simulated and Observed Data

Due to the representativeness of point measurements and their comparatively sparse spatial coverage, evaluation of simulated ET is either rather limited or further processing is required (Table 5). For instance, eddy covariance measurements have been further processed to allow for a better spatial representation through a combination with MODIS data (Pan et al. 2022) or land use-specific adjustments (Széles et al. 2020).

TABLE 5 | Practical considerations for evaluating simulated with observation-based evapotranspiration data.

Data source	Considerations
Satellite-derived	• Choice between readily available products and application of an algorithm • Clouds restrict the availability of products based on optical satellite data • Be aware of seasonal (rainy or dry season) and regional biases in the data
In situ	• More suitable for small catchments or model units for which the measurements are representative, e.g., considering the footprint of eddy covariance measurements

As with the retrieval of SC, cloud cover is a major challenge for the ET retrieval from satellite data (Ruhoff et al. 2013, Table 5), particularly for those approaches that are based on optical satellite imagery (e.g., MOD16 algorithm). In monsoon regions, this may result in missing data for the entire rainy season (Immerzeel and Droogers 2008). Sometimes, climate station data and the Penman-Monteith equation are used to fill data gaps due to cloud cover (Nesru et al. 2020; Rientjes et al. 2013). Wang et al. (2023) point out that increased cloud coverage also affects microwave-based products like GLEAM, resulting in an overestimation of ET. More generally, Jiang et al. (2020) argue that MODIS products have positive and negative biases in different locations. For the comparison of modeled and satellite-derived ET, seasonal misfits are reported: for dry conditions between the HBV model and SEBS (Rientjes et al. 2013), in summer attributed to biases of MODIS for heterogeneous vegetation covers and in complex terrain (Ala-aho et al. 2017) and on the contrary, in the wet season for MOD16 (Ruhoff et al. 2013). Moreover, for a study in New South Wales, the errors in the observed water balance derived from MODIS ET reach values of up to 30% of total precipitation (Figure 5 in Vervoort et al. 2014), preventing any meaningful calibration of a hydrological model. A further challenge regarding the spatial information provided by remotely sensed ET products is that ET in general follows a climate gradient. Koch et al. (2022) show for the Senegal river basin that normalizing the ET pattern to remove a precipitation-induced trend in the data before using it for model calibration improves model performance with respect to spatial patterns of ET.

Regarding the choice of the ET product, the readily available ET products from MODIS and GLEAM are more commonly used, as the application of algorithms like SEBS (Nesru et al. 2020) or SEBAL (Immerzeel and Droogers 2008) requires more knowledge and effort. In addition, further processing is sometimes necessary, for example, Immerzeel and Droogers (2008) adjust solar radiation inputs to make reference ET in SEBAL and SWAT comparable. However, in a study by Odusanya et al. (2019) GLEAM and MODIS ET differ significantly from each other, underlining that the choice of the ET product affects the results. Lettenmaier et al. (2015) highlight the need to reduce uncertainties in parameters like albedo, aerodynamic temperatures, and surface emissivity to improve ET retrieval from satellite data.

3.6 | Altimetric Water Level

3.6.1 | Observation Methods, Uncertainties and Data Products

Satellite-borne radar altimeters provide estimates of surface elevation. Wherever the track of a specific satellite passes over an open water surface, the obtained surface elevation reflects the water level at the time and location of the overpass, in literature referred to as "virtual station". Despite various sources of uncertainty arising from, among others, errors in radial orbits, instrumental bias, or the choice of the tracker (e.g., Birkett 1998), many studies have demonstrated that altimeters can estimate water levels of inland water bodies "within errors of ~10–50 cm" (e.g., Alsdorf et al. 2007; Seyler et al. 2013; Schneider et al. 2018; Nielsen et al. 2022). With early altimeters, this was possible only for water bodies "larger than ~1–2 km in width" (e.g., Birkett 1998), but retrievals from more recent satellite missions have been routinely and successfully used to generate water level estimates for rivers and lakes with widths of 100–200 m and in some cases even as small as 50 m (Sulistioadi et al. 2015). Depending on the satellite mission, the estimates are available at temporal resolutions of 10–35 days. A comprehensive list of past and current satellite missions and their characteristics is provided in Table A2. Several open access databases (see Table A1; Abdalla et al. 2021; Birkett et al. 2009; Crétaux et al. 2011; Schwatke et al. 2015) have facilitated the easy usability of processed and ready-to-use altimetry data for applications in terrestrial hydrology.

3.6.2 | Challenges in the Comparison of Simulated and Observed Data

While altimetry water levels are routinely used for data assimilation in hydrodynamic models (e.g., Andreadis et al. 2007; Biancamaria et al. 2009; Tourian et al. 2017; Larnier et al. 2021), they are used much less for hydrological model evaluation (e.g., Coe et al. 2002; Getirana 2010; Zhang et al. 2023). This seems to be due to the fact that hydrological models typically only produce estimates of streamflow rather than river water levels as output. They therefore require the definition of stage–discharge relationships for individual virtual stations to map modeled streamflow onto altimetry water level estimates.

Challenges when using water surface elevation for hydrological model evaluation can generally be related to the inherent characteristics of altimetry estimates on the one hand and the definition of stage-discharge relationships on the other hand. With respect to the challenges associated with the characteristics of altimetry estimates, high inter- and along-track distances between individual signal retrievals lead to the situation that a considerable fraction of surface waters, located between individual satellite tracks, is missed by satellite overpasses (Alsdorf et al. 2007). In addition, and as a major source of uncertainty, the wide footprint of the radar beam on the earth's surface, which may be several kilometers wide, may result in a "contaminated" signal over heterogeneous areas composed of water and land surfaces (Calmant et al. 2008). The long repeat periods of 10–35 days lead to rather low temporal resolutions at individual virtual stations. Virtual stations from multiple missions can be combined to increase the number of data points and ensure robust model calibration (e.g., Getirana 2010; Sun et al. 2012; Hulsman et al. 2020). However, uncertainties of combined series should be accounted for carefully, since each mission is characterized by different types of errors that need to be considered before combining the products (Calmant et al. 2008).

Challenges associated with the definition of stage-discharge relationships are mainly related to uncertainties in estimates of channel roughness (i.e., Manning coefficient), geometry, and slope. Different strategies have been outlined to relate altimetry water levels to streamflow, as summarized in Table 6, based on (i) coupling past in situ water level and streamflow observations with a functional relationship between in situ and altimetry water levels (e.g., Roux et al. 2008; Coe and Birkett 2004; Tarpanelli et al. 2017), (ii) coupling altimetry water levels with in situ streamflow estimates (Kouraev et al. 2004; Zakharova et al. 2020; Tarpanelli et al. 2013; Tourian et al. 2017), or (iii) merging altimetry water levels with information on cross-section geometry, channel width, slope, and flow velocity (Bjerklie et al. 2003; Birkinshaw et al. 2014; Zakharova et al. 2020). In the frequent absence of local data, values for river depth, cross-sectional area, and depth-averaged velocity have to be either calibrated or inferred from other data sources, such as optical

TABLE 6 | Practical considerations for evaluating simulated hydrological variables with water levels derived from altimetry data.

Data source	Considerations
Satellite-derived	Strategies to relate altimetry water levels to available in situ streamflow data through a stage-discharge relationship • Coupling past in situ water level and streamflow observations with a functional relationship between in situ and altimetry water levels • Coupling altimetry water levels with in situ streamflow estimates Strategies to simulate streamflow data through a stage-discharge relationship when in situ observations are unavailable • Conversion of simulated streamflow into water levels by means of stage-discharge relationships based on river cross section geometry, roughness, and slope. • Additional remotely sensed data may be employed to obtain more reliable estimates of river cross section geometry

sensors for river width (e.g., Landsat) or terrain elevation data, which, however, frequently do not offer the required spatial resolution and accuracy. Similarly, robust and reliable values for channel roughness need to be estimated, which are typically assumed constant in time and space in models, whereas they can be highly variable in reality. Several studies highlight that the main source of uncertainty when using altimetry data is due to the simplified definition of the channel cross sections at the virtual stations (Getirana et al. 2013). Indeed, the lower model performances obtained from calibration to altimetry data arise mostly from the uncertainty of the rating curve parameters in addition to the low temporal resolution of the altimetry data used (Sun et al. 2012).

Looking ahead, the surface water and ocean topography (SWOT) mission, launched in December 2022, will provide global water level observations for rivers wider than 50–100 m (Biancamaria et al. 2016). Instead of using the profiling technique of conventional altimeters, this mission makes use of wide swath altimetry, which provides spatially distributed water surface levels, river widths, and slopes (Prigent et al. 2016; Tarpanelli et al. 2021), addressing some of the above challenges (Durand et al. 2016, 2023; Yoon et al. 2016; Sikder et al. 2021) with the potential of establishing streamflow estimates from remotely sensed water levels and geometry complementary to river gauges (Fekete et al. 2012; Gleason and Durand 2020). For more detailed information on satellite altimetry water levels and their potential, we refer the reader to synthesis and review papers by Alsdorf et al. (2007), Gleason and Durand (2020), Abdalla et al. (2021) and Tarpanelli et al. (2021).

4 | Benefits and Future Directions of Including Hydrological Observations Beyond Streamflow for Hydrological Model Evaluation and Calibration

4.1 | Benefits of Additional Hydrological Variables in Model Calibration

One of the main motivations for including additional hydrological variables besides streamflow in model calibration is to enhance the representation of these variables and to improve process realism (e.g., Hrachowitz et al. 2021; Kreye et al. 2019; Werth et al. 2009; Schumacher et al. 2018). Including additional hydrological variables can lead to more realistic parameter estimates (e.g., Széles et al. 2020) and reduced parameter uncertainties of the related model components (Parajka et al. 2009; Duethmann et al. 2014; Kelleher et al. 2017). Furthermore, using additional data besides streamflow may result in better performance of other hydrological variables, known as cross-benefit. For example, authors report an improved simulation of ET after calibrating the model with in situ root zone soil moisture (Rajib et al. 2016), machine learning-based soil moisture data (Mei et al. 2023), or TWS (Dembélé et al. 2020b; Pool et al. 2024) in addition to streamflow.

Obviously, adding a further variable to model calibration in addition to streamflow is not likely to increase streamflow model performance during calibration. While many studies show only minor streamflow performance decreases (e.g., Nijzink et al. 2018; Rajib et al. 2016; Bai et al. 2018; Milzow et al. 2011), other studies reveal stronger trade-offs with a clear decrease in streamflow performance (e.g., Lo et al. 2010; Dembélé et al. 2020b). Such trade-offs can indicate data problems (e.g., Duethmann et al. 2014) or issues with the model structure that can motivate model structural improvements (Hulsman et al. 2021a; Meyer Oliveira et al. 2021). Improved process consistency may, however, increase streamflow performance during evaluation (e.g., Li et al. 2018; Gomis-Cebolla et al. 2022; Parajka and Blöschl 2008b). Furthermore, the spatial information introduced into the calibration process by satellite-derived data can result in better streamflow simulation at gauges that were not used during calibration, as Li et al. (2018) show for soil moisture.

Recently, an increased number of studies include more than one hydrological variable besides streamflow, for example, combining streamflow, surface water level, TWS, and soil moisture (Milzow et al. 2011), streamflow, TWS, soil moisture, and ET (Dembélé et al. 2020b; Pool et al. 2024), or streamflow, groundwater level, and ET (Ala-aho et al. 2017). Nijzink et al. (2018) compare the value of various remote-sensing products for constraining model parameters of lumped conceptual models applied to catchments in Europe. Their results indicate that AMSRE-E and ASCAT soil moisture, as well as GRACE storage variations, are particularly useful, while LSAF and MODIS ET products contribute relatively little. Széles et al. (2020) include various in situ data to calibrate a model for a small catchment in Austria and find that soil moisture and ET data had the largest influence on runoff. Additional hydrological variables may also be used for further evaluation, particularly with regard to spatial patterns, as shown by Gaur et al. (2022), who calibrate a model against streamflow and groundwater levels and evaluate its spatial performance against ET and soil moisture patterns.

A big advantage of additional hydrological variables is that they may serve for model calibration in catchments where in situ streamflow measurements are scarce or absent. For large rivers, altimetry data have often been applied for this purpose, and the majority of the reviewed studies show that streamflow is effectively estimated with very good to satisfactory performance, even though with slightly lower performance values compared to the classical calibration with in situ streamflow data (Getirana 2010; Sun et al. 2012; Huang et al. 2020; Liu et al. 2015). ET data have also often been applied in data-scarce regions (e.g., Odusanya et al. 2019). Immerzeel and Droogers (2008) and Meyer Oliveira et al. (2021) showed that calibration with ET improved streamflow model performance when compared to an uncalibrated model, underlining the benefit of hydrological variables beyond streamflow for predictions in ungauged basins. Calibration in data-scarce regions can further benefit from including several hydrological variables, for example, by combining altimetry water levels with GRACE TWS (Hulsman et al. 2020).

4.2 | Recommendations and Future Directions

There has been a growing emphasis on advancing multi-variable calibration approaches to improve the accuracy and reliability of hydrological models. We see several key aspects to further advance multivariable calibration and evaluation of hydrological models. Open data platforms, such as those for various satellite-derived products, in situ data (Table A1), or the CAMELS data sets (Newman et al. 2015) for catchment streamflow and meteorological data, are extremely useful to facilitate data access. However, there is much more data available that is not currently shared in large-scale databases, and further data sharing would greatly improve data accessibility. Furthermore, the provision of ready-to-use products, such as those from MODIS for remote sensing data, reduces the need for specific expertise in remote sensing technologies and has greatly increased the use of satellite-derived data for hydrological model calibration. Additionally, the development and sharing of pre-processing tools and scripts can simplify processing workflows.

Systematically accounting for uncertainties represents a crucial step in advancing multi-variable hydrological model calibration. While it is now common practice to consider parameter uncertainty in model calibration, uncertainties in the data used for model calibration, particularly beyond streamflow, are often not explicitly addressed. Examples of exceptions are the consideration of uncertainties in the study by Jiang et al. (2020) or a discussion of uncertainties, for example, by Ruhoff et al. (2013). However, disregarding data uncertainties during model calibration can result in over-constraining the model, resulting in overconfidence in the model and reduced process consistency. Moreover, underestimation of uncertainties in hydrological data may result in unjustified model rejection when used for model evaluation.

Unfortunately, information on data uncertainties is not always reported with the data, although this would be desirable. Since information on data uncertainties is often only vague, for example when estimated based on expert opinion, informal approaches may be the most appropriate to account for uncertainties. For example, following the Limits of Acceptability approach, uncertainty estimates can be used to define the limits of acceptability (Mackay et al. 2018). Alternatively, objective functions may be expressed as fuzzy functions to reflect vague information on uncertainties (Duethmann et al. 2015). Moreover, soft indicators may be generated based on satellite-derived data of hydrological variables together with uncertainty estimates and expert knowledge (Bouaziz et al. 2021). Accounting for uncertainties will particularly prevent cases where integrating uncertain data of an additional hydrological variable deteriorates a model. It will enable adding data from different hydrological variables with different levels of uncertainty, so that the model can then serve to integrate the knowledge on the hydrology of the system obtained from different data sources. The integration of various data sets in a model framework may also reveal potential inconsistencies in specific data sets. Ways to convey such information from data users to data providers should be established and used.

5 | Conclusion

The reviewed studies use a variety of alternative hydrological variables beyond streamflow observations for model evaluation and calibration in catchment hydrology. Consistent findings from these studies highlight that:

1. The use of multiple hydrological variables during model calibration considerably improves process representation in hydrological models. Mostly, additional variables are tested in combination with streamflow, but the use of a single variable itself (e.g., ET) has also proven useful, underlining the suitability for model applications in data-scarce regions.
2. The inclusion of an additional variable in an otherwise streamflow-based calibration often results in a slight decrease in performance with regard to streamflow. This trade-off is generally accepted, given the gain in process representation.
3. Comparability of observed and modeled data needs to be accounted for to facilitate the use of additional data. Differences may be due to a different variable definition in the data product or the model (e.g., groundwater in conceptual models) as well as differences in spatial scales and biases in observed and modeled data. These differences result in challenges and substantial pre-processing efforts when comparing observed and modeled data.
4. Uncertainties are often not considered. However, integrating biased or noisy data of an additional hydrological variable may deteriorate model performance and decrease process consistency, though such cases are hardly reported. There is a need for better knowledge of uncertainties in observational data sets, particularly in remote sensing-derived data, as well as for multi-variable calibration frameworks that enable the consideration of uncertainties.
5. Objective functions for multi-variable calibration often include subjective weights. More transparent and well-documented approaches for evaluating performance are recommended to facilitate compatibility between different studies.

As the reviewed studies show, the use of additional hydrological variables has become state-of-the-art in model evaluation. However, given the amount of hydrological data beyond streamflow observations, we conclude that these data treasures still remain largely unexploited. This review should encourage the use of additional hydrological variables in hydrological model evaluation and ultimately contribute to improving process representation in models. This is especially important given the increased relevance of accurate process representation when models are applied under changed conditions such as in climate or land use change impact studies.

Author Contributions

Paul D. Wagner: conceptualization (lead), writing – original draft (lead), writing – review and editing (lead). **Doris Duethmann:** conceptualization (lead), writing – original draft (lead), writing – review and editing (lead). **Jens Kiesel:** conceptualization (equal), writing – original draft (equal), writing – review and editing (equal). **Sandra Pool:** conceptualization (equal), writing – original draft (equal), writing – review and editing (equal). **Markus Hrachowitz:** conceptualization (supporting), writing – original draft (equal), writing – review and editing (equal). **Serena Ceola:** conceptualization (supporting), writing – original draft (supporting), writing – review and editing (equal). **Anna Herzog:** conceptualization (supporting), visualization (lead),

writing – review and editing (equal). **Tobias Houska:** conceptualization (supporting), writing – original draft (supporting), writing – review and editing (equal). **Ralf Loritz:** conceptualization (supporting), writing – original draft (supporting), writing – review and editing (equal). **Diana Spieler:** conceptualization (supporting), writing – original draft (supporting), writing – review and editing (equal). **Maria Staudinger:** writing – review and editing (equal). **Larisa Tarasova:** conceptualization (supporting), writing – original draft (supporting), writing – review and editing (equal). **Stephan Thober:** conceptualization (supporting), writing – review and editing (equal). **Nicola Fohrer:** conceptualization (supporting), writing – review and editing (equal). **Doerthe Tetzlaff:** conceptualization (supporting), writing – original draft (supporting), writing – review and editing (equal). **Thorsten Wagener:** conceptualization (supporting), writing – review and editing (equal). **Björn Guse:** conceptualization (supporting), writing – original draft (supporting), writing – review and editing (equal).

Acknowledgments

Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study. Details of the reviewed studies are provided as Supporting Information.

Related WIREs Articles

Toward a theoretical framework for integrated modeling of hydrological change

Parameter estimation and uncertainty analysis in hydrological modeling

References

Abbott, M. B., J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen. 1986. "An Introduction to the European Hydrological System—Systeme Hydrologique Europeen,SHE, 1: History and Philosophy of a Physically-Based, Distributed Modelling System." *Journal of Hydrology* 87, no. 1–2: 45–59.

Abdalla, S., A. A. Kolahchi, M. Ablain, et al. 2021. "Altimetry for the Future: Building on 25 Years of Progress." *Advances in Space Research* 68, no. 2: 319–363. https://doi.org/10.1016/j.asr.2021.01.022.

Abhervé, R., C. Roques, J.-R. de Dreuzy, et al. 2024. "Improving Calibration of Groundwater Flow Models Using Headwater Streamflow Intermittence." *Hydrological Processes* 38: e15167. https://doi.org/10.1002/hyp.15167.

Ala-aho, P., C. Soulsby, H. Wang, and D. Tetzlaff. 2017. "Integrated Surface-Subsurface Model to Investigate the Role of Groundwater in Headwater Catchment Runoff Generation: A Minimalist Approach to Parameterisation." *Journal of Hydrology* 547: 664–677.

Alemayehu, T., H. V. Gupta, A. van Griensven, and W. Bauwens. 2022. "On the Calibration of Spatially Distributed Hydrologic Models for Poorly Gauged Basins: Exploiting Information From Streamflow Signatures and Remote Sensing-Based Evapotranspiration Data." *Water* 14, no. 8: 1252. https://doi.org/10.3390/w14081252.

Alsdorf, D. E., E. Rodríguez, and D. P. Lettenmaier. 2007. "Measuring Surface Water From Space." *Reviews of Geophysics* 45, no. 2: 2006RG000197. https://doi.org/10.1029/2006RG000197.

Al-Yaari, A., J. P. Wigneron, W. Dorigo, et al. 2019. "Assessment and Inter-Comparison of Recently Developed/Reprocessed Microwave Satellite Soil Moisture Products Using ISMN Ground-Based Measurements." *Remote Sensing of Environment* 224: 289–303. https://doi.org/10.1016/j.rse.2019.02.008.

Andreadis, K. M., E. A. Clark, D. P. Lettenmaier, and D. E. Alsdorf. 2007. "Prospects for River Discharge and Depth Estimation Through Assimilation of Swath-Altimetry Into a Raster-Based Hydrodynamics Model." *Geophysical Research Letters* 34, no. 10: 2007GL029721. https://doi.org/10.1029/2007GL029721.

Appel, I. 2018. "Uncertainty in Satellite Remote Sensing of Snow Fraction for Water Resources Management." *Frontiers of Earth Science* 12, no. 4: 711–727. https://doi.org/10.1007/s11707-018-0720-1.

Asadzadeh, M., and B. Tolson. 2013. "Pareto Archived Dynamically Dimensioned Search With Hypervolume-Based Selection for Multi-Objective Optimization." *Engineering Optimization* 45, no. 12: 1489–1509.

Bai, P., X. Liu, and C. Liu. 2018. "Improving Hydrological Simulations by Incorporating GRACE Data for Model Calibration." *Journal of Hydrology* 557: 291–304. https://doi.org/10.1016/j.jhydrol.2017.12.025.

Bailey, R. T., S. Park, K. Bieger, J. G. Arnold, and P. M. Allen. 2020. "Enhancing SWAT+ Simulation of Groundwater Flow and Groundwater-Surface Water Interactions Using MODFLOW Routines." *Environmental Modelling & Software* 126: 104660. https://doi.org/10.1016/j.envsoft.2020.104660.

Bailey, R. T., T. C. Wible, M. Arabi, R. M. Records, and J. Ditty. 2016. "Assessing Regional-Scale Spatio-Temporal Patterns of Groundwater–Surface Water Interactions Using a Coupled SWAT-MODFLOW Model." *Hydrological Processes* 30, no. 23: 4420–4433. https://doi.org/10.1002/hyp.10933.

Baroni, G., L. M. Scheiffele, M. Schrön, J. Ingwersen, and S. E. Oswald. 2018. "Uncertainty, Sensitivity and Improvements in Soil Moisture Estimation With Cosmic-Ray Neutron Sensing." *Journal of Hydrology* 564: 873–887. https://doi.org/10.1016/j.jhydrol.2018.07.053.

Bauer-Marschallinger, B., and C. Paulik. 2019. "Copernicus Global Land Operations Vegetation and Energy. Product User Manual (PUM). Soil Surface Moisture 1 km, Version 1, Issue l1. 30."

Bauer-Marschallinger, B., C. Paulik, S. Hochstöger, et al. 2018. "Soil Moisture From Fusion of Scatterometer and SAR: Closing the Scale Gap With Temporal Filtering." *Remote Sensing* 10, no. 7: 1030. https://doi.org/10.3390/rs10071030.

Beck, H. E., M. Pan, D. G. Miralles, et al. 2021. "Evaluation of 18 Satellite- and Model-Based Soil Moisture Products Using In Situ Measurements From 826 Sensors." *Hydrology and Earth System Sciences* 25, no. 1: 17–40. https://doi.org/10.5194/hess-25-17-2021.

Behzadpour, S. 2022. "Assessment and Mitigation of Systematic Errors in GRACE and GRACE-FO Observations. Dissertation, Graz University of Technology, Faculty of Mathematics, Physics and Geodesy."

Bennett, K. E., J. E. Cherry, B. Balk, and S. Lindsey. 2019. "Using MODIS Estimates of Fractional Snow Cover Area to Improve Streamflow Forecasts in Interior Alaska." *Hydrology and Earth System Sciences* 23, no. 5: 2439–2459. https://doi.org/10.5194/hess-23-2439-2019.

Beven, K. J. 2011. *Rainfall-Runoff Modelling: The Primer*. John Wiley & Sons.

Biancamaria, S., P. D. Bates, A. Boone, and N. M. Mognard. 2009. "Large-Scale Coupled Hydrologic and Hydraulic Modelling of the Ob River in Siberia." *Journal of Hydrology* 379, no. 1–2: 136–150. https://doi.org/10.1016/j.jhydrol.2009.09.054.

Biancamaria, S., D. P. Lettenmaier, and T. M. Pavelsky. 2016. "The SWOT Mission and Its Capabilities for Land Hydrology." *Surveys in Geophysics* 37, no. 2: 117–147. https://doi.org/10.1007/s10712-015-9346-y.

Birkel, C., and C. Soulsby. 2015. "Advancing Tracer-Aided Rainfall–Runoff Modelling: A Review of Progress, Problems and Unrealised Potential." *Hydrological Processes* 29: 5227–5240. https://doi.org/10.1002/hyp.10594.

Birkett, C. M. 1998. "Contribution of the TOPEX NASA Radar Altimeter to the Global Monitoring of Large Rivers and Wetlands." *Water Resources Research* 34, no. 5: 1223–1239. https://doi.org/10.1029/98WR00124.

Birkett, C. M., C. Reynolds, B. Beckley, and B. Doorn. 2009. "From Research to Operations: The USDA Global Reservoir and Lake Monitor." In *Coastal Altimetry*, edited by S. Vignudelli, A. G. Kostianoy, P. Cipollini, and J. Benveniste. Springer Publications.

Birkinshaw, S. J., P. Moore, C. G. Kilsby, G. M. O'donnell, A. J. Hardy, and P. A. M. Berry. 2014. "Daily Discharge Estimation at Ungauged River Sites Using Remote Sensing." *Hydrological Processes* 28, no. 3: 1043–1054. https://doi.org/10.1002/hyp.9647.

Bjerklie, D. M., S. L. Dingman, C. J. Vorosmarty, C. H. Bolster, and R. G. Congalton. 2003. "Evaluating the Potential for Measuring River Discharge From Space." *Journal of Hydrology* 278, no. 1–4: 17–38. https://doi.org/10.1016/S0022-1694(03)00129-X.

Boeing, F., O. Rakovec, R. Kumar, et al. 2022. "High-Resolution Drought Simulations and Comparison to Soil Moisture Observations in Germany." *Hydrology and Earth System Sciences* 26, no. 19: 5137–5161. https://doi.org/10.5194/hess-26-5137-2022.

Bogena, H. R., M. Schrön, J. Jakobi, et al. 2022. "COSMOS-Europe: A European Network of Cosmic-Ray Neutron Soil Moisture Sensors." *Earth System Science Data* 14: 1125–1151. https://doi.org/10.5194/essd-14-1125-2022.

Bongio, M., F. Avanzi, and C. De Michele. 2016. "Hydroelectric Power Generation in an Alpine Basin: Future Water-Energy Scenarios in a Run-Of-The-River Plant." *Advances in Water Resources* 94: 318–331. https://doi.org/10.1016/j.advwatres.2016.05.017.

Bouaziz, L., A. Weerts, J. Schellekens, et al. 2018. "Redressing the Balance: Quantifying Net Intercatchment Groundwater Flows." *Hydrology and Earth System Sciences* 22, no. 12: 6415–6434.

Bouaziz, L. J. E., F. Fenicia, G. Thirel, et al. 2021. "Behind the Scenes of Streamflow Model Performance." *Hydrology and Earth System Sciences* 25: 1069–1095. https://doi.org/10.5194/hess-25-1069-2021.

Bouaziz, L. J. E., S. C. Steele-Dunne, J. Schellekens, et al. 2020. "Improved Understanding of the Link Between Catchment-Scale Vegetation Accessible Storage and Satellite-Derived Soil Water Index." *Water Resources Research* 56, no. 3: e2019WR026365. https://doi.org/10.1029/2019WR026365.

Brauer, C. C., A. J. Teuling, P. J. J. F. Torfs, and R. Uijlenhoet. 2014. "The Wageningen Lowland Runoff Simulator (WALRUS): A Lumped Rainfall–Runoff Model for Catchments With Shallow Groundwater." *Geoscientific Model Development* 7: 2313–2332. https://doi.org/10.5194/gmd-7-2313-2014.

Bronstert, A., and E. J. Plate. 1997. "Modelling of Runoff Generation and Soil Moisture Dynamics for Hillslopes and Micro-Catchments." *Journal of Hydrology* 198, no. 1–4: 177–195. https://doi.org/10.1016/S0022-1694(96)03306.

Calmant, S., F. Seyler, and J. F. Cretaux. 2008. "Monitoring Continental Surface Waters by Satellite Altimetry." *Surveys in Geophysics* 29: 247–269. https://doi.org/10.1007/s10712-008-9051-1.

Cammalleri, C., F. Micale, and J. Vogt. 2015. "On the Value of Combining Different Modelled Soil Moisture Products for European Drought Monitoring." *Journal of Hydrology* 525: 547–558. https://doi.org/10.1016/j.jhydrol.2015.04.021.

Carroll, T. R. 2001. "Airborne Gamma Radiation Snow Survey Program: A User's Guide, Version 5.0, National Operational Hydro- Logic Remote Sensing Center (NOHRSC)." *Chanhassen* 14: 2023. https://www.nohrsc.noaa.gov/snowsurvey/historical.html.

Chen, J., A. Cazenave, C. Dahle, et al. 2022. "Applications and Challenges of GRACE and GRACE Follow-On Satellite Gravimetry." *Surveys in Geophysics* 43: 305–345. https://doi.org/10.1007/s10712-021-09685-x.

Cho, E., J. M. Jacobs, S. E. Tuttle, R. Schroeder, and C. Olheiser. 2020. "Improvement of Operational Airborne Gamma Radiation Snow Water Equivalent Estimates Using SMAP Soil Moisture." *Remote Sensing of Environment* 240: 111668. https://doi.org/10.1016/j.rse.2020.111668.

Cho, E., Y. Kwon, S. V. Kumar, and C. M. Vuyovich. 2023. "Assimilation of Airborne Gamma Observations Provides Utility for Snow Estimation in Forested Environments." *Hydrology and Earth System Sciences* 27: 4039–4056. https://doi.org/10.5194/hess-27-4039-2023.

Coe, M. T., and C. M. Birkett. 2004. "Calculation of River Discharge and Prediction of Lake Height From Satellite Radar Altimetry: Example for the Lake Chad Basin." *Water Resources Research* 40, no. 10: 2003WR002543. https://doi.org/10.1029/2003WR002543.

Coe, M. T., M. H. Costa, A. Botta, and C. Birkett. 2002. "Long-Term Simulations of Discharge and Floods in the Amazon Basin." *Journal of Geophysical Research: Atmospheres* 107, no. D20: 1–17. https://doi.org/10.1029/2001JD000740.

Corbari, C., G. Ravazzani, A. Perotto, et al. 2022. "Weekly Monitoring and Forecasting of Hydropower Production Coupling Meteo-Hydrological Modeling With Ground and Satellite Data in the Italian Alps." *Hydrology* 9, no. 2: 29. https://doi.org/10.3390/hydrology9020029.

Crétaux, J.-F., A. Arsen, S. Calmant, et al. 2011. "SOLS: A Lake Database to Monitor in the Near Real Time Water Level and Storage Variations From Remote Sensing Data." *Advances in Space Research* 47, no. 9: 1497–1507. https://doi.org/10.1016/j.asr.2011.01.004.

Creutzfeldt, B., A. Güntner, S. Vorogushyn, and B. Merz. 2010. "The Benefits of Gravimeter Observations for Modelling Water Storage Changes at the Field Scale." *Hydrology and Earth System Sciences* 14: 1715–1730. https://doi.org/10.5194/hess-14-1715-2010.

de Graaf, I. E. M., T. Gleeson, L. P. H. Van Beek, E. H. Sutanudjaja, and M. F. P. Bierkens. 2019. "Environmental Flow Limits to Global Groundwater Pumping." *Nature* 574: 90–94. https://doi.org/10.1038/s41586-019-1594-4.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002. "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II." *IEEE Transactions on Evolutionary Computation* 6, no. 2: 182–197.

Dembélé, M., N. Ceperley, S. J. Zwart, E. Salvadore, G. Mariethoz, and B. Schaefli. 2020a. "Potential of Satellite and Reanalysis Evaporation Datasets for Hydrological Modelling Under Various Model Calibration Strategies." *Advances in Water Resources* 143: 103667. https://doi.org/10.1016/j.advwatres.2020.103667.

Dembélé, M., M. Hrachowitz, H. H. G. Savenije, G. Mariéthoz, and B. Schaefli. 2020b. "Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets." *Water Resources Research* 56, no. 1: e2019WR026085. https://doi.org/10.1029/2019WR026085.

Demirel, M. C., J. Mai, G. Mendiguren, J. Koch, L. Samaniego, and S. Stisen. 2018. "Combining Satellite Data and Appropriate Objective Functions for Improved Spatial Pattern Performance of a Distributed Hydrologic Model." *Hydrology and Earth System Sciences* 22: 1299–1315. https://doi.org/10.5194/hess-22-1299-2018.

Demirel, M. C., A. Özen, S. Orta, et al. 2019. "Additional Value of Using Satellite-Based Soil Moisture and Two Sources of Groundwater Data for Hydrological Model Calibration." *Water* 11, no. 10: 2083. https://doi.org/10.3390/w11102083.

Di Marco, N., D. Avesani, M. Righetti, M. Zaramella, B. Majone, and M. Borga. 2021. "Reducing Hydrological Modelling Uncertainty by Using MODIS Snow Cover Data and a Topography-Based Distribution Function Snowmelt Model." *Journal of Hydrology* 599: 126020. https://doi.org/10.1016/j.jhydrol.2021.126020.

Dong, C. 2018. "Remote Sensing, Hydrological Modeling and In Situ Observations in Snow Cover Research: A Review." *Journal of Hydrology* 561: 573–583. https://doi.org/10.1016/j.jhydrol.2018.04.027.

Dorigo, W., I. Himmelbauer, D. Aberer, et al. 2021. "The International Soil Moisture Network: Serving Earth System Science for Over a Decade." *Hydrology and Earth System Sciences* 25: 5749–5804. https://doi.org/10.5194/hess-25-5749-2021.

Dorigo, W. A., A. Xaver, M. Vreugdenhil, et al. 2013. "Global Automated Quality Control of In Situ Soil Moisture Data From the International Soil Moisture Network." *Vadose Zone Journal* 12, no. 3: vzj2012.0097. https://doi.org/10.2136/vzj2012.0097.

Dressler, K. A., G. H. Leavesley, R. C. Bales, and S. R. Fassnacht. 2006. "Evaluation of Gridded Snow Water Equivalent and Satellite Snow Cover Products for Mountain Basins in a Hydrologic Model." *Hydrological Processes* 20, no. 4: 673–688. https://doi.org/10.1002/hyp.6130.

Duethmann, D., T. Bolch, D. Farinotti, et al. 2015. "Attribution of Streamflow Trends in Snow- and Glacier Melt Dominated Catchments of the Tarim River, Central Asia." *Water Resources Research* 51, no. 6: 24. https://doi.org/10.1002/2014WR016716.

Duethmann, D., J. Peters, T. Blume, S. Vorogushyn, and A. Güntner. 2014. "The Value of Satellite-Derived Snow Cover Images for Calibrating a Hydrological Model in Snow-Dominated Catchments in Central Asia." *Water Resources Research* 50, no. 3: 2002–2021. https://doi.org/10.1002/2013wr014382.

Duethmann, D., A. Smith, C. Soulsby, et al. 2022. "Evaluating Satellite-Derived Soil Moisture Data for Improving the Internal Consistency of Process-Based Ecohydrological Modelling." *Journal of Hydrology* 614: 128462. https://doi.org/10.1016/j.jhydrol.2022.128462.

Durand, M., C. J. Gleason, P. A. Garambois, et al. 2016. "An Intercomparison of Remote Sensing River Discharge Estimation Algorithms From Measurements of River Height, Width, and Slope." *Water Resources Research* 52, no. 6: 4527–4549. https://doi.org/10.1002/2015WR018434.

Durand, M., C. J. Gleason, T. M. Pavelsky, et al. 2023. "A Framework for Estimating Global River Discharge From the Surface Water and Ocean Topography Satellite Mission." *Water Resources Research* 59, no. 4: e2021WR031614. https://doi.org/10.1029/2021WR031614.

Ebel, B. A., K. Loague, J. E. Vanderkwaak, et al. 2007. "Near-Surface Hydrologic Response for a Steep, Unchanneled Catchment Near Coos Bay, Oregon: 2. Physics-Based Simulations." *American Journal of Science* 307, no. 4: 709–748. https://doi.org/10.2475/04.2007.03.

Efstratiadis, A., and D. Koutsoyiannis. 2010. "One Decade of Multi-Objective Calibration Approaches in Hydrological Modelling: A Review." *Hydrological Sciences Journal* 55, no. 1: 58–78.

Ershadi, A., M. F. McCabe, J. P. Evans, N. W. Chaney, and E. F. Wood. 2014. "Multi-Site Evaluation of Terrestrial Evaporation Models Using FLUXNET Data." *Agricultural and Forest Meteorology* 187: 46–61. https://doi.org/10.1016/j.agrformet.2013.11.008.

Fang, Z., H. Bogena, S. Kollet, and H. Vereecken. 2016. "Scale Dependent Parameterization of Soil Hydraulic Conductivity in 3D Simulation of Hydrological Processes in a Forested Headwater Catchment." *Journal of Hydrology* 536: 365–375. https://doi.org/10.1016/j.jhydrol.2016.03.020.

Fassnacht, S. R., J. I. López-Moreno, C. Ma, et al. 2017. "Spatio-Temporal Snowmelt Variability Across the Headwaters of the Southern Rocky Mountains." *Frontiers of Earth Science* 11: 505–514.

Fekete, B. M., U. Looser, A. Pietroniro, and R. D. Robarts. 2012. "Rationale for Monitoring Discharge on the Ground." *Journal of Hydrometeorology* 13, no. 6: 1977–1986. https://doi.org/10.1175/JHM-D-11-0126.1.

Fenicia, F., J. McDonnell, and H. Savenije. 2008. "Learning From Model Improvement: On the Contribution of Complementary Data to Process Understanding." *Water Resources Research* 44: W06419. https://doi.org/10.1029/2007WR006386.

Finger, D., M. Vis, M. Huss, and J. Seibert. 2015. "The Value of Multiple Data Set Calibration Versus Model Complexity for Improving the Performance of Hydrological Models in Mountain Catchments." *Water Resources Research* 51, no. 4: 1939–1958. https://doi.org/10.1002/2014wr015712.

Fisher, J. C. 2005. "River 3D: Watershed With Mountains." https://ian.umces.edu/media-library/river-3d-watershed-with-mountains/.

Fleming, S. W., L. Zukiewicz, M. L. Strobel, H. Hofman, and A. G. Goodbody. 2023. "SNOTEL, the Soil Climate Analysis Network, and Water Supply Forecasting at the Natural Resources Conservation Service: Past, Present, and Future." *JAWRA Journal of the American Water Resources Association* 59, no. 4: 585–599.

Fontrodona-Bach, A., B. Schaefli, R. Woods, A. J. Teuling, and J. R. Larsen. 2023. "NH-SWE: Northern Hemisphere Snow Water Equivalent Dataset Based on In Situ Snow Depth Time Series." *Earth System Science Data* 15, no. 6: 2577–2599. https://doi.org/10.5194/essd-15-2577-2023.

Frappart, F., and G. Ramillien. 2018. "Monitoring Groundwater Storage Changes Using the Gravity Recovery and Climate Experiment (GRACE) Satellite Mission: A Review." *Remote Sensing* 10, no. 6: 829. https://doi.org/10.3390/rs10060829.

Gafurov, A., and A. Bardossy. 2009. "Cloud Removal Methodology From MODIS Snow Cover Product." *Hydrology and Earth System Sciences* 13, no. 7: 1361–1373.

Gascoin, S., O. Hagolle, M. Huc, et al. 2015. "A Snow Cover Climatology for the Pyrenees From MODIS Snow Products." *Hydrology and Earth System Sciences* 19, no. 5: 2337–2351. https://doi.org/10.5194/hess-19-2337-2015.

Gaur, S., B. Singh, A. Bandyopadhyay, S. Stisen, and R. Singh. 2022. "Spatial Pattern-Based Performance Evaluation and Uncertainty Analysis of a Distributed Hydrological Model." *Hydrological Processes* 36, no. 5: e14586. https://doi.org/10.1002/hyp.14586.

Getirana, A. C. 2010. "Integrating Spatial Altimetry Data Into the Automatic Calibration of Hydrological Models." *Journal of Hydrology* 387, no. 3–4: 244–255. https://doi.org/10.1016/j.jhydrol.2010.04.013.

Getirana, A. C., A. Boone, D. Yamazaki, and N. Mognard. 2013. "Automatic Parameterization of a Flow Routing Scheme Driven by Radar Altimetry Data: Evaluation in the Amazon Basin." *Water Resources Research* 49, no. 1: 614–629. https://doi.org/10.1002/wrcr.20077.

Gleason, C. J., and M. T. Durand. 2020. "Remote Sensing of River Discharge: A Review and a Framing for the Discipline." *Remote Sensing* 12, no. 7: 1107. https://doi.org/10.3390/rs12071107.

Gleeson, T., T. Wagener, P. Döll, et al. 2021. "GMD Perspective: The Quest to Improve the Evaluation of Groundwater Representation in Continental- to Global-Scale Models." *Geoscientific Model Development* 14: 7545–7571. https://doi.org/10.5194/gmd-14-7545-2021.

Gomis-Cebolla, J., A. Garcia-Arias, M. Perpinya-Valles, and F. Frances. 2022. "Evaluation of Sentinel-1, SMAP and SMOS Surface Soil Moisture Products for Distributed Eco-Hydrological Modelling in Mediterranean Forest Basins." *Journal of Hydrology* 608: 127569. https://doi.org/10.1016/j.jhydrol.2022.127569.

Graeff, T., E. Zehe, S. Schlaeger, et al. 2010. "A Quality Assessment of Spatial TDR Soil Moisture Measurements in Homogenous and Heterogeneous Media With Laboratory Experiments." *Hydrology and Earth System Sciences* 14, no. 6: 1007–1020. https://doi.org/10.5194/hess-14-1007-2010.

Gruber, A., T. Scanlon, R. van der Schalie, W. Wagner, and W. Dorigo. 2019. "Evolution of the ESA CCI Soil Moisture Climate Data Records and Their Underlying Merging Methodology." *Earth System Science Data* 11, no. 2: 717–739. https://doi.org/10.5194/essd-11-717-2019.

Güntner, A., H. Gerdener, E. Boergens, et al. 2023. "Veränderungen der Wasserspeicherung in Deutschland seit 2002 aus Beobachtungen der Satellitengravimetrie [Water storage changes in Germany since 2002 from satellite gravity observations]." *Hydrologie und*

Wasserbewirtschaftung 67: 74–89. https://doi.org/10.5675/HyWa_2023.2_1.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez. 2009. "Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling." *Journal of Hydrology* 377, no. 1–2: 80–91.

Guse, B., J. Kiesel, M. Pfannerstill, and N. Fohrer. 2020. "Assessing Parameter Identifiability for Multiple Performance Criteria to Constrain Model Parameters." *Hydrological Sciences Journal* 65, no. 7: 1158–1172. https://doi.org/10.1080/02626667.2020.1734204.

Gyawali, D. R., and A. Bárdossy. 2022. "Development and Parameter Estimation of Snowmelt Models Using Spatial Snow-Cover Observations From MODIS." *Hydrology and Earth System Sciences* 26, no. 12: 3055–3077. https://doi.org/10.5194/hess-26-3055-2022.

Hattermann, F., V. Krysanova, F. Wechsung, and M. Wattenbach. 2004. "Integrating Groundwater Dynamics in Regional Hydrological Modelling." *Environmental Modelling and Software* 19: 1039–1051. https://doi.org/10.1016/j.envsoft.2003.11.007.

Hawdon, A., D. McJannet, and J. Wallace. 2014. "Calibration and Correction Procedures for Cosmic-Ray Neutron Soil Moisture Probes Located Across Australia." *Water Resources Research* 50, no. 6: 5029–5043. https://doi.org/10.1002/2013WR015138.

Heudorfer, B., E. Haaf, K. Stahl, and R. Barthel. 2019. "Index-Based Characterization and Quantification of Groundwater Dynamics." *Water Resources Research* 55: 5575–5592. https://doi.org/10.1029/2018WR024418.

Hill, D. F., E. A. Burakowski, R. L. Crumley, et al. 2019. "Converting Snow Depth to Snow Water Equivalent Using Climatological Variables." *Cryosphere* 13, no. 7: 1767–1784. https://doi.org/10.5194/tc-13-1767-2019.

Hrachowitz, M., M. Stockinger, M. Coenders-Gerrits, et al. 2021. "Reduction of Vegetation-Accessible Water Storage Capacity After Deforestation Affects Catchment Travel Time Distributions and Increases Young Water Fractions in a Headwater Catchment." *Hydrology and Earth System Sciences* 25, no. 9: 4887–4915. https://doi.org/10.5194/hess-25-4887-2021.

Huang, Q., D. Long, M. Du, Z. Han, and P. Han. 2020. "Daily Continuous River Discharge Estimation for Ungauged Basins Using a Hydrologic Model Calibrated by Satellite Altimetry: Implications for the SWOT Mission." *Water Resources Research* 56, no. 7: e2020WR027309. https://doi.org/10.1029/2020WR027309.

Huang, Z., Q. Tang, M. H. Lo, et al. 2019. "The Influence of Groundwater Representation on Hydrological Simulation and Its Assessment Using Satellite-Based Water Storage Variation." *Hydrological Processes* 33, no. 8: 1218–1230. https://doi.org/10.1002/hyp.13393.

Hulsman, P., M. Hrachowitz, and H. H. G. Savenije. 2021a. "Improving the Representation of Long-Term Storage Variations With Conceptual Hydrological Models in Data-Scarce Regions." *Water Resources Research* 57, no. 4: e2020WR028837. https://doi.org/10.1029/2020WR028837.

Hulsman, P., H. H. Savenije, and M. Hrachowitz. 2021b. "Learning From Satellite Observations: Increased Understanding of Catchment Processes Through Stepwise Model Improvement." *Hydrology and Earth System Sciences* 25, no. 2: 957–982.

Hulsman, P., H. C. Winsemius, C. I. Michailovsky, H. H. Savenije, and M. Hrachowitz. 2020. "Using Altimetry Observations Combined With GRACE to Select Parameter Sets of a Hydrological Model in a Data-Scarce Region." *Hydrology and Earth System Sciences* 24, no. 6: 3331–3359. https://doi.org/10.5194/hess-24-3331-2020.

Humphrey, V., M. Rodell, and A. Eicker. 2023. "Using Satellite-Based Terrestrial Water Storage Data: A Review." *Surveys in Geophysics* 44: 1489–1517. https://doi.org/10.1007/s10712-022-09754-9.

Huss, M., M. Zemp, P. C. Joerg, and N. Salzmann. 2014. "High Uncertainty in 21st Century Runoff Projections From Glacierized Basins." *Journal of Hydrology* 510: 35–48. https://doi.org/10.1016/j.jhydrol.2013.12.017.

Immerzeel, W. W., and P. Droogers. 2008. "Calibration of a Distributed Hydrological Model Based on Satellite Evapotranspiration." *Journal of Hydrology* 349: 411–424.

Iwema, J., R. Rosolem, R. Baatz, T. Wagener, and H. R. Bogena. 2015. "Investigating Temporal Field Sampling Strategies for Site-Specific Calibration of Three Soil Moisture–Neutron Intensity Parameterisation Methods." *Hydrology and Earth System Sciences* 19: 3203–3216. https://doi.org/10.5194/hess-19-3203-2015.

Jackisch, C., K. Germer, T. Graeff, et al. 2020. "Soil Moisture and Matric Potential – An Open Field Comparison of Sensor Systems." *Earth System Science Data* 12, no. 1: 683–697. https://doi.org/10.5194/essd-12-683-2020.

Jiang, D., and K. Wang. 2019. "The Role of Satellite-Based Remote Sensing in Improving Simulated Streamflow: A Review." *Water* 11, no. 8: 1615.

Jiang, L., H. Wu, J. Tao, J. S. Kimball, L. Alfieri, and X. Chen. 2020. "Satellite-Based Evapotranspiration in Hydrological Model Calibration." *Remote Sensing* 12: 428. https://doi.org/10.3390/rs12030428.

Jing, M., F. Heße, R. Kumar, et al. 2018. "Improved Regional-Scale Groundwater Representation by the Coupling of the Mesoscale Hydrologic Model (mHM v5.7) to the Groundwater Model OpenGeoSys (OGS)." *Geoscientific Model Development* 11: 1989–2007. https://doi.org/10.5194/gmd-11-1989-2018.

Johnson, J. B., and D. Marks. 2004. "The Detection and Correction of Snow Water Equivalent Pressure Sensor Errors." *Hydrological Processes* 18, no. 18: 3513–3525.

Käser, D., and D. Hunkeler. 2016. "Contribution of Alluvial Groundwater to the Outflow of Mountainous Catchments." *Water Resources Research* 52: 680–697. https://doi.org/10.1002/2014WR016730.

Kelleher, C., B. McGlynn, and T. Wagener. 2017. "Characterizing and Reducing Equifinality by Constraining a Distributed Catchment Model With Regional Signatures, Local Observations, and Process Understanding." *Hydrology and Earth System Sciences* 21: 3325–3352. https://doi.org/10.5194/hess-21-3325-2017.

Kim, H. W., K. Hwang, Q. Mu, O. L. Seung, and M. Choi. 2012. "Validation of MODIS 16 Global Terrestrial Evapotranspiration Products in Various Climates and Land Cover Types in Asia." *KSCE Journal of Civil Engineering* 16: 229–238. https://doi.org/10.1007/s12205-012-0006-1.

Kite, G. W., and P. Droogers. 2000. "Comparing Evapotranspiration Estimates From Satellites, Hydrological Models and Field Data." *Journal of Hydrology* 229, no. 1–2: 3–18. https://doi.org/10.1016/S0022-1694(99)00195-X.

Kleine, L., D. Tetzlaff, A. Smith, H. Wang, and C. Soulsby. 2020. "Using Water Stable Isotopes to Understand Evaporation, Moisture Stress, and Re-Wetting in Catchment Forest and Grassland Soils of the Summer Drought of 2018." *Hydrology and Earth System Sciences* 24: 3737–3752. https://doi.org/10.5194/hess-24-3737-2020.

Klosterhalfen, A., J. Chi, N. Kljun, et al. 2023. "Two-Level Eddy Covariance Measurements Reduce Bias in Land-Atmosphere Exchange Estimates Over a Heterogeneous Boreal Forest Landscape." *Agricultural and Forest Meteorology* 339: 109523. https://doi.org/10.1016/j.agrformet.2023.109523.

Koch, J., M. C. Demirel, and S. Stisen. 2018. "The SPAtial EFficiency Metric (SPAEF): Multiple-Component Evaluation of Spatial Patterns for Optimization of Hydrological Models." *Geoscientific Model Development* 11: 1873–1886. https://doi.org/10.5194/gmd-11-1873-2018.

Koch, J., M. C. Demirel, and S. Stisen. 2022. "Climate Normalized Spatial Patterns of Evapotranspiration Enhance the Calibration of a Hydrological Model." *Remote Sensing* 14: 315. https://doi.org/10.3390/rs14020315.

Koch, J., G. Mendiguren, G. Mariethoz, and S. Stisen. 2017. “Spatial Sensitivity Analysis of Simulated Land Surface Patterns in a Catchment Model Using a Set of Innovative Spatial Performance Metrics.” *Journal of Hydrometeorology* 18, no. 4: 1121–1142.

Köhli, M., M. Schrön, M. Zreda, U. Schmidt, P. Dietrich, and S. Zacharias. 2015. “Footprint Characteristics Revised for Field-Scale Soil Moisture Monitoring With Cosmic-Ray Neutrons.” *Water Resources Research* 51, no. 7: 5772–5790. https://doi.org/10.1002/2015WR017169.

Kollat, J. B., and P. M. Reed. 2006. “Comparing State-Of-The-Art Evolutionary Multi-Objective Algorithms for Long-Term Groundwater Monitoring Design.” *Advances in Water Resources* 29, no. 6: 792–807. https://doi.org/10.1016/j.advwatres.2005.07.010.

Kouraev, A. V., E. A. Zakharova, O. Samain, N. M. Mognard, and A. Cazenave. 2004. “Ob′ River Discharge From TOPEX/Poseidon Satellite Altimetry (1992–2002).” *Remote Sensing of Environment* 93, no. 1–2: 238–245. https://doi.org/10.1016/j.rse.2004.07.007.

Kreye, P., M. Gelleszun, and G. Meon. 2019. “Parameter Identification in Hydrological Models Using Groundwater-Level Measurements and Satellite-Based Soil Moisture.” *Hydrological Sciences Journal* 64, no. 6: 633–652.

Kuczera, G., and M. Mroczkowski. 1998. “Assessment of Hydrologic Parameter Uncertainty and the Worth of Multiresponse Data.” *Water Resources Research* 34: 1481–1489.

Kundu, D., R. W. Vervoort, and F. F. van Ogtrop. 2017. “The Value of Remotely Sensed Surface Soil Moisture for Model Calibration Using SWAT.” *Hydrological Processes* 31, no. 15: 2764–2780. https://doi.org/10.1002/hyp.11219.

Kunnath-Poovakka, A., D. Ryu, L. J. Renzullo, and B. George. 2018. “Remotely Sensed ET for Streamflow Modelling in Catchments With Contrasting Flow Characteristics: An Attempt to Improve Efficiency.” *Stochastic Environmental Research and Risk Assessment* 32: 1973–1992.

Kuppel, S., D. Tetzlaff, M. P. Maneta, and C. Soulsby. 2018. “What Can We Learn From Multi-Data Calibration of a Process-Based Ecohydrological Model?” *Environmental Modelling and Software* 101: 301–316. https://doi.org/10.1016/j.envsoft.2018.01.001.

Larnier, K., J. Monnier, P. A. Garambois, and J. Verley. 2021. “River Discharge and Bathymetry Estimation From SWOT Altimetry Measurements.” *Inverse Problems in Science and Engineering* 29, no. 6: 759–789. https://doi.org/10.1080/17415977.2020.1803858.

Le Moine, N., V. Andréassian, C. Perrin, and C. Michel. 2007. “How Can Rainfall-Runoff Models Handle Inter-Catchment Groundwater Flows? Theoretical Study Based on 1040 French Catchments.” *Water Resources Research* 43: W06428. https://doi.org/10.1029/2006WR005608.

Lettenmaier, D. P., D. Alsdorf, J. Dozier, G. J. Huffman, M. Pan, and E. F. Wood. 2015. “Inroads of Remote Sensing Into Hydrologic Science During the WRR Era.” *Water Resources Research* 51, no. 9: 7309–7342. https://doi.org/10.1002/2015WR017616.

Li, B., M. Rodell, S. Kumar, et al. 2019a. “Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges.” *Water Resources Research* 55: 7564–7586. https://doi.org/10.1029/2018wr024618.

Li, X., Y. Jing, H. Shen, and L. Zhang. 2019b. “The Recent Developments in Cloud Removal Approaches of MODIS Snow Cover Product.” *Hydrology and Earth System Sciences* 23, no. 5: 2401–2416. https://doi.org/10.5194/hess-23-2401-2019.

Li, Y., S. Grimaldi, V. R. N. Pauwels, and J. P. Walker. 2018. “Hydrologic Model Calibration Using Remotely Sensed Soil Moisture and Discharge Measurements: The Impact on Predictions at Gauged and Ungauged Locations.” *Journal of Hydrology* 557: 897–909. https://doi.org/10.1016/j.jhydrol.2018.01.013.

Lin, P., M. A. Rajib, Z.-L. Yang, et al. 2018. “Spatiotemporal Evaluation of Simulated Evapotranspiration and Streamflow Over Texas Using the WRF-Hydro-RAPID Modeling Framework.” *Journal of the American Water Resources Association* 54, no. 1: 40–54. https://doi.org/10.1111/1752-1688.12585.

Liston, G. E. 2004. “Representing Subgrid Snow Cover Heterogeneities in Regional and Global Models.” *Journal of Climate* 17, no. 6: 1381–1397. https://doi.org/10.1175/1520-0442(2004)017<1381:RSSCHI>2.0.CO;2.

Liu, G., F. W. Schwartz, K. H. Tseng, and C. K. Shum. 2015. “Discharge and Water-Depth Estimates for Ungauged Rivers: Combining Hydrologic, Hydraulic, and Inverse Modeling With Stage and Water-Area Measurements From Satellites.” *Water Resources Research* 51, no. 8: 6017–6035. https://doi.org/10.1002/2015WR016971.

Liu, Y., T. Wagener, H. E. Beck, and A. Hartmann. 2020. “What Is the Hydrologically Effective Area of a Catchment?” *Environmental Research Letters* 15: 104024. https://doi.org/10.1088/1748-9326/aba7e5.

Lo, M. H., J. S. Famiglietti, P. J.-F. Yeh, and T. H. Syed. 2010. “Improving Parameter Estimation and Water Table Depth Simulation in a Land Surface Model Using GRACE Water Storage and Estimated Base Flow Data.” *Water Resources Research* 46, no. 5: 2009WR007855. https://doi.org/10.1029/2009WR007855.

López López, P., E. H. Sutanudjaja, J. Schellekens, G. Sterk, and M. F. P. Bierkens. 2017. “Calibration of a Large-Scale Hydrological Model Using Satellite-Based Soil Moisture and Evapotranspiration Products.” *Hydrology and Earth System Sciences* 21, no. 6: 3125–3144. https://doi.org/10.5194/hess-21-3125-2017.

López-Moreno, J. I., S. R. Fassnacht, S. Beguería, and J. B. P. Latron. 2011. “Variability of Snow Depth at the Plot Scale: Implications for Mean Depth Estimation and Sampling Strategies.” *Cryosphere* 5, no. 3: 617–629. https://doi.org/10.5194/tc-5-617-2011.

Loritz, R., M. Bassiouni, A. Hildebrandt, S. K. Hassler, and E. Zehe. 2022. “Leveraging Sap Flow Data in a Catchment-Scale Hybrid Model to Improve Soil Moisture and Transpiration Estimates.” *Hydrology and Earth System Sciences* 26, no. 18: 4757–4771. https://doi.org/10.5194/hess-26-4757-2022.

Loritz, R., S. K. Hassler, C. Jackisch, et al. 2017. “Picturing and Modeling Catchments by Representative Hillslopes.” *Hydrology and Earth System Sciences* 21, no. 2: 1225–1249. https://doi.org/10.5194/hess-21-1225-2017.

Luo, S., D. Tetzlaff, A. Smith, and C. Soulsby. 2024. “Long-Term Drought Effects on Landscape Water Storage and Resilience Under Contrasting Landuses.” *Journal of Hydrology* 636: 131339. https://doi.org/10.1016/j.jhydrol.2024.131339.

Mackay, J. D., N. E. Barrand, D. M. Hannah, et al. 2018. “Glacio-Hydrological Melt and Run-Off Modelling: Application of a Limits of Acceptability Framework for Model Comparison and Selection.” *Cryosphere* 12: 2175–2210. https://doi.org/10.5194/tc-12-2175-2018.

Mai, J., J. R. Craig, and B. A. Tolson. 2020. “Simultaneously Determining Global Sensitivities of Model Parameters and Model Structure.” *Hydrology and Earth System Sciences* 24: 5835–5858. https://doi.org/10.5194/hess-24-5835-2020.

Mälicke, M., S. K. Hassler, T. Blume, M. Weiler, and E. Zehe. 2020. “Soil Moisture: Variable in Space but Redundant in Time.” *Hydrology and Earth System Sciences* 24, no. 5: 2633–2653. https://doi.org/10.5194/hess-24-2633-2020.

McMillan, H., G. Coxon, R. Araki, et al. 2023. “When Good Signatures Go Bad: Applying Hydrologic Signatures in Large Sample Studies.” *Hydrological Processes* 37, no. 9: e14987. https://doi.org/10.1002/hyp.14987.

McNamara, J. P., D. Tetzlaff, K. Bishop, et al. 2011. “Storage as a Metric of Catchment Comparison.” *Hydrological Processes* 25, no. 21: 3364–3371. https://doi.org/10.1002/hyp.8113.

Mei, Y., J. Mai, H. X. Do, et al. 2023. “Can Hydrological Models Benefit From Using Global Soil Moisture, Evapotranspiration, and Runoff

Products as Calibration Targets?" *Water Resources Research* 59, no. 2: e2022WR032064. https://doi.org/10.1029/2022WR032064.

Merz, R., L. Tarasova, and S. Basso. 2020. "Parameter's Controls of Distributed Catchment Models—How Much Information Is in Conventional Catchment Descriptors?" *Water Resources Research* 56: e2019WR026008. https://doi.org/10.1029/2019WR026008.

Meyer Oliveira, A., A. S. Fleischmann, and R. C. D. Paiva. 2021. "On the Contribution of Remote Sensing-Based Calibration to Model Hydrological and Hydraulic Processes in Tropical Regions." *Journal of Hydrology* 597: 126184. https://doi.org/10.1016/j.jhydrol.2021.126184.

Miller, Z. S., E. H. Peitzsch, E. A. Sproles, K. W. Birkeland, and R. T. Palomaki. 2022. "Assessing the Seasonal Evolution of Snow Depth Spatial Variability and Scaling in Complex Mountain Terrain." *Cryosphere* 16, no. 12: 4907–4930. https://doi.org/10.5194/tc-16-4907-2022.

Milzow, C., P. E. Krogh, and P. Bauer-Gottwein. 2011. "Combining Satellite Radar Altimetry, SAR Surface Soil Moisture and GRACE Total Storage Changes for Hydrological Model Calibration in a Large Poorly Gauged Catchment." *Hydrology and Earth System Sciences* 15, no. 6: 1729–1743. https://doi.org/10.5194/hess-15-1729-2011.

Miralles, D. G., R. A. M. De Jeu, J. H. Gash, T. R. H. Holmes, and A. J. Dolman. 2011a. "Magnitude and Variability of Land Evaporation and Its Components at the Global Scale." *Hydrology and Earth System Sciences* 15: 967–981. https://doi.org/10.5194/hess-15-967-2011.

Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman. 2011b. "Global Land-Surface Evaporation Estimated From Satellite-Based Observations." *Hydrology and Earth System Sciences* 15: 453–469. https://doi.org/10.5194/hess-15-453-2011.

Monteith, J. L. 1965. *The State and Movement of Water in Living Organisms: 19th Symposia of the Society for Experimental Biology*, edited by G. E. Fogg, 205–234. Cambridge University Press.

Mu, Q., M. Zhao, and S. W. Running. 2011. "Improvements to a MODIS Global Terrestrial Evapotranspiration Algorithm." *Remote Sensing of Environment* 115, no. 8: 1781–1800.

Nash, J. E., and J. V. Sutcliffe. 1970. "River Flow Forecasting Through Conceptual Models Part I—A Discussion of Principles." *Journal of Hydrology* 10, no. 3: 282–290.

Neill, A. J., C. Birkel, M. P. Maneta, D. Tetzlaff, and C. Soulsby. 2021. "Structural Changes to Forests During Regeneration Affect Water Flux Partitioning, Water Ages and Hydrological Connectivity: Insights From Tracer-Aided Ecohydrological Modeling." *Hydrology and Earth System Sciences* 25: 4861–4886. https://hess.copernicus.org/articles/25/4861/2021/.

Nesru, M., A. Shetty, and M. K. Nagaraj. 2020. "Multi-Variable Calibration of Hydrological Model in the Upper Omo-Gibe Basin, Ethiopia." *Acta Geophysica* 68, no. 2: 537–551.

Newman, A. J., M. P. Clark, K. Sampson, et al. 2015. "Development of a Large-Sample Watershed-Scale Hydrometeorological Dataset for the Contiguous USA: Dataset Characteristics and Assessment of Regional Variability in Hydrologic Model Performance." *Hydrology and Earth System Sciences* 19: 209–223. https://doi.org/10.5194/hess-19-209-2015.

Nguyen, T. V., B. Uniyal, D. A. Tran, and T. B. T. Pham. 2022. "On the Evaluation of Both Spatial and Temporal Performance of Distributed Hydrological Models Using Remote Sensing Products." *Remote Sensing* 14, no. 9: 1959. https://doi.org/10.3390/rs14091959.

Nielsen, K., E. Zakharova, A. Tarpanelli, O. B. Andersen, and J. Benveniste. 2022. "River Levels From Multi Mission Altimetry, a Statistical Approach." *Remote Sensing of Environment* 270: 112876. https://doi.org/10.1016/j.rse.2021.112876.

Nijzink, R. C., S. Almeida, I. G. Pechlivanidis, et al. 2018. "Constraining Conceptual Hydrological Models With Multiple Information Sources." *Water Resources Research* 54, no. 10: 8332–8362. https://doi.org/10.1029/2017wr021895.

Odusanya, A. E., B. Mehdi, C. Schürz, et al. 2019. "Multi-Site Calibration and Validation of SWAT With Satellite-Based Evapotranspiration in a Data-Sparse Catchment in Southwestern Nigeria." *Hydrology and Earth System Sciences* 23, no. 2: 1113–1144. https://doi.org/10.5194/hess-23-1113-2019.

Oki, T., and S. Kanae. 2006. "Global Hydrological Cycles and World Water Resources." *Science* 313: 1068–1072. https://doi.org/10.1126/science.1128845.

Orth, R. 2021. "Global Soil Moisture Data Derived Through Machine Learning Trained With In-Situ Measurements." *Scientific Data* 8, no. 1: 1–14. https://doi.org/10.1038/s41597-021-00964-1.

Pan, S. L., Y. P. Xu, H. T. Gu, B. Yu, and W. D. Xuan. 2022. "Evaluation of Remote Sensing-Based Evapotranspiration Datasets for Improving Hydrological Model Simulation in Humid Region of East China." *Remote Sensing* 14, no. 18: 4546. https://doi.org/10.3390/rs14184546.

Parajka, J., and G. Blöschl. 2008a. "Spatio-Temporal Combination of MODIS Images – Potential for Snow Cover Mapping." *Water Resources Research* 44, no. 3: 2007WR006204. https://doi.org/10.1029/2007wr006204.

Parajka, J., and G. Blöschl. 2008b. "The Value of MODIS Snow Cover Data in Validating and Calibrating Conceptual Hydrologic Models." *Journal of Hydrology* 358, no. 3–4: 240–258. https://doi.org/10.1016/j.jhydrol.2008.06.006.

Parajka, J., L. Holko, Z. Kostka, and G. Blöschl. 2012. "MODIS Snow Cover Mapping Accuracy in a Small Mountain Catchment – Comparison Between Open and Forest Sites." *Hydrology and Earth System Sciences* 16: 2365–2377. https://doi.org/10.5194/hess-16-2365-2012.

Parajka, J., R. Merz, and G. Blöschl. 2007. "Uncertainty and Multiple Objective Calibration in Regional Water Balance Modelling: Case Study in 320 Austrian Catchments." *Hydrological Processes* 21, no. 4: 435–446. https://doi.org/10.1002/hyp.6253.

Parajka, J., V. Naeimi, G. Blöschl, and J. Komma. 2009. "Matching ERS Scatterometer Based Soil Moisture Patterns With Simulations of a Conceptual Dual Layer Hydrologic Model Over Austria." *Hydrology and Earth System Sciences* 13, no. 2: 259–271. https://doi.org/10.5194/hess-13-259-2009.

Parajka, J., V. Naeimi, G. Blöschl, W. Wagner, R. Merz, and K. Scipal. 2006. "Assimilating Scatterometer Soil Moisture Data Into Conceptual Hydrologic Models at the Regional Scale." *Hydrology and Earth System Sciences* 10, no. 3: 353–368. https://doi.org/10.5194/hess-10-353-2006.

Pastorello, G., C. Trotta, E. Canfora, et al. 2020. "The FLUXNET2015 Dataset and the ONEFlux Processing Pipeline for Eddy Covariance Data." *Scientific Data* 7, no. 1: 225.

Paulik, C., W. Dorigo, W. Wagner, and R. Kidd. 2014. "Validation of the ASCAT Soil Water Index Using In Situ Data From the International Soil Moisture Network." *International Journal of Applied Earth Observation and Geoinformation* 30: 1–8. https://doi.org/10.1016/j.jag.2014.01.007.

Peker, I., and A. Sorman. 2021. "Application of SWAT Using Snow Data and Detecting Climate Change Impacts in the Mountainous Eastern Regions of Turkey." *Water* 13, no. 14: 1982. https://doi.org/10.3390/w13141982.

Peng, J., A. Loew, O. Merlin, and N. E. C. Verhoest. 2017. "A Review of Spatial Downscaling of Satellite Remotely Sensed Soil Moisture." *Reviews of Geophysics* 55, no. 2: 341–366. https://doi.org/10.1002/2016rg000543.

Pfannerstill, M., B. Guse, and N. Fohrer. 2014. "A Multi-Storage Groundwater Concept for the SWAT Model to Emphasize Nonlinear Groundwater Dynamics in Lowland Catchments." *Hydrological Processes* 28: 5599–5612.

Pianosi, F., K. Beven, J. Freer, et al. 2016. "Sensitivity Analysis of Environmental Models: A Systematic Review With Practical Workflow."

Environmental Modelling & Software 79: 214–232. https://doi.org/10.1016/j.envsoft.2016.02.008.

Pirazzini, R., L. Leppänen, G. Picard, et al. 2018. "European In-Situ Snow Measurements: Practices and Purposes." *Sensors* 18, no. 7: 2016. https://doi.org/10.3390/s18072016.

Poméon, T., B. Diekkrüger, A. Springer, J. Kusche, and A. Eicker. 2018. "Multi-Objective Validation of SWAT for Sparsely-Gauged West African River Basins – A Remote Sensing Approach." *Water* 10: 451.

Pool, D. R., and J. H. Eychaner. 1995. "Measurements of Aquifer-Storage Change and Specific Yield Using Gravity Surveys." *Groundwater* 33: 425–432. https://doi.org/10.1111/j.1745-6584.1995.tb00299.x.

Pool, S., K. Fowler, and M. Peel. 2024. "Benefit of Multivariate Model Calibration for Different Climatic Regions." *Water Resources Research* 60: e2023WR036364. https://doi.org/10.1029/2023WR036364.

Poyatos, R., V. Granda, V. Flo, et al. 2021. "Global Transpiration Data From Sap Flow Measurements: The SAPFLUXNET Database." *Earth System Science Data* 13, no. 6: 2607–2649. https://doi.org/10.5194/essd-13-2607-2021.

Prigent, C., D. P. Lettenmaier, F. Aires, and F. Papa. 2016. "Toward a High-Resolution Monitoring of Continental Surface Water Extent and Dynamics, at Global Scale: From GIEMS (Global Inundation Extent From Multi-Satellites) to SWOT (Surface Water Ocean Topography)." *Surveys in Geophysics* 37: 339–355. https://doi.org/10.1007/s10712-015-9339-x.

Ragettli, S., F. Pellicciotti, W. W. Immerzeel, et al. 2015. "Unraveling the Hydrology of a Himalayan Catchment Through Integration of High Resolution In Situ Data and Remote Sensing With an Advanced Simulation Model." *Advances in Water Resources* 78: 94–111. https://doi.org/10.1016/j.advwatres.2015.01.013.

Rajib, A., G. R. Evenson, H. E. Golden, and C. R. Lane. 2018. "Hydrologic Model Predictability Improves With Spatially Explicit Calibration Using Remotely Sensed Evapotranspiration and Biophysical Parameters." *Journal of Hydrology* 567: 668–683. https://doi.org/10.1016/j.jhydrol.2018.10.024.

Rajib, M. A., V. Merwade, and Z. Yu. 2016. "Multi-Objective Calibration of a Hydrologic Model Using Spatially Distributed Remotely Sensed/In-Situ Soil Moisture." *Journal of Hydrology* 536: 192–207. https://doi.org/10.1016/j.jhydrol.2016.02.037.

Rakovec, O., R. Kumar, J. Mai, et al. 2016. "Multiscale and Multivariate Evaluation of Water Fluxes and States Over European River Basins." *Journal of Hydrometeorology* 17, no. 1: 287–307. https://doi.org/10.1175/jhm-d-15-0054.1.

Reinecke, R., S. Gnann, L. Stein, et al. 2024. "Uncertainty in Model Estimates of Global Groundwater Depth." *Environmental Research Letters* 19, no. 11: 114066.

Rientjes, T. H. M., L. P. Muthuwatta, M. G. Bos, M. J. Booij, and H. A. Bhatti. 2013. "Multi-Variable Calibration of a Semi-Distributed Hydrological Model Using Streamflow Data and Satellite-Based Evapotranspiration." *Journal of Hydrology* 505: 276–290. https://doi.org/10.1016/j.jhydrol.2013.10.006.

Robinson, D. A., C. S. Campbell, J. W. Hopmans, et al. 2008. "Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review." *Vadose Zone Journal* 7, no. 1: 358–389. https://doi.org/10.2136/vzj2007.0143.

Rodhe, A., L. Nyberg, and K. Bishop. 1996. "Transit Times for Water in a Small Till Catchment From a Step Shift in the Oxygen 18 Content of the Water Input." *Water Resources Research* 32, no. 12: 3497–3511.

Rosolem, R., H. V. Gupta, W. J. Shuttleworth, X. Zeng, and L. G. G. de Gonçalves. 2012. "A Fully Multiple-Criteria Implementation of the Sobol' Method for Parameter Sensitivity Analysis." *Journal of Geophysical Research* 117: D07103. https://doi.org/10.1029/2011JD016355.

Roux, E., M. Cauhope, M. P. Bonnet, S. Calmant, P. Vauchel, and F. Seyler. 2008. "Daily water stage estimated from satellite altimetric data for large river basin monitoring/Estimation de hauteurs d'eau journalières a partir de données d'altimétrie radar pour la surveillance des grands basins fluviaux." *Hydrological Sciences Journal* 53, no. 1: 81–99. https://doi.org/10.1623/hysj.53.1.81.

Roy, A., A. Royer, and R. Turcotte. 2010. "Improvement of Springtime Streamflow Simulations in a Boreal Environment by Incorporating Snow-Covered Area Derived From Remote Sensing Data." *Journal of Hydrology* 390, no. 1–2: 35–44. https://doi.org/10.1016/j.jhydrol.2010.06.027.

Ruhoff, A. L., A. R. Paz, L. E. O. C. Aragao, et al. 2013. "Assessment of the MODIS Global Evapotranspiration Algorithm Using Eddy Covariance Measurements and Hydrological Modelling in the Rio Grande Basin." *Hydrological Sciences Journal* 58, no. 8: 1658–1676.

Samaniego, L., R. Kumar, and S. Attinger. 2010. "Multiscale Parameter Regionalization of a Grid-Based Hydrologic Model at the Mesoscale." *Water Resources Research* 46: W05523. https://doi.org/10.1029/2008WR007327.

Samaniego, L., R. Kumar, S. Thober, et al. 2017. "Toward Seamless Hydrologic Predictions Across Spatial Scales." *Hydrology and Earth System Sciences* 21, no. 9: 4323–4346.

Schneider, R., A. Tarpanelli, K. Nielsen, H. Madsen, and P. Bauer-Gottwein. 2018. "Evaluation of Multi-Mode CryoSat-2 Altimetry Data Over the Po River Against In Situ Data and a Hydrodynamic Model." *Advances in Water Resources* 112: 17–26. https://doi.org/10.1016/j.advwatres.2017.11.027.

Schöber, J., S. Achleitner, R. Kirnbauer, F. Schöberl, and H. Schönlaub. 2010. "Hydrological Modelling of Glacierized Catchments Focussing on the Validation of Simulated Snow Patterns – Applications Within the Flood Forecasting System of the Tyrolean River Inn." *Advances in Geosciences* 27: 99–109. https://doi.org/10.5194/adgeo-27-99-2010.

Schumacher, M., E. Forootan, A. I. van Dijk, et al. 2018. "Improving Drought Simulations Within the Murray-Darling Basin by Combined Calibration/Assimilation of GRACE Data Into the WaterGAP Global Hydrology Model." *Remote Sensing of Environment* 204: 212–228.

Schwatke, C., D. Dettmering, W. Bosch, and F. Seitz. 2015. "DAHITI—An Innovative Approach for Estimating Water Level Time Series Over Inland Waters Using Multi-Mission Satellite Altimetry." *Hydrology and Earth System Sciences* 19: 4345–4364. https://doi.org/10.5194/hess-19-4345-2015.

Seibert, J. 2000. "Multi-Criteria Calibration of a Conceptual Runoff Model Using a Genetic Algorithm." *Hydrology and Earth System Sciences* 4, no. 2: 215–224.

Seibert, J., and S. Bergström. 2022. "A Retrospective on Hydrological Catchment Modelling Based on Half a Century With the HBV Model." *Hydrology and Earth System Sciences* 26: 1371–1388. https://doi.org/10.5194/hess-26-1371-2022.

Serreze, M. C., M. P. Clark, and A. Frei. 2001. "Characteristics of Large Snowfall Events in the Montane Western United States as Examined Using Snowpack Telemetry (SNOTEL) Data." *Water Resources Research* 37, no. 3: 675–688.

Seyler, F., S. Calmant, J. S. Da Silva, D. M. Moreira, F. Mercier, and C. K. Shum. 2013. "From TOPEX/Poseidon to Jason-2/OSTM in the Amazon Basin." *Advances in Space Research* 51, no. 8: 1542–1550. https://doi.org/10.1016/j.asr.2012.11.002.

Sheffield, J., E. F. Wood, M. Pan, et al. 2018. "Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions." *Water Resources Research* 54, no. 12: 9724–9758.

Sikder, M. S., M. Bonnema, C. M. Emery, et al. 2021. "A Synthetic Data Set Inspired by Satellite Altimetry and Impacts of Sampling on Global

Spaceborne Discharge Characterization." *Water Resources Research* 57, no. 2: e2020WR029035. https://doi.org/10.1029/2020WR029035.

Silvestro, F., S. Gabellani, R. Rudari, F. Delogu, P. Laiolo, and G. Boni. 2015. "Uncertainty Reduction and Parameter Estimation of a Distributed Hydrological Model With Ground and Remote-Sensing Data." *Hydrology and Earth System Sciences* 19, no. 4: 1727–1751. https://doi.org/10.5194/hess-19-1727-2015.

Sirisena, T. J. G., S. Maskey, and R. Ranasinghe. 2020. "Hydrological Model Calibration With Streamflow and Remote Sensing Based Evapotranspiration Data in a Data Poor Basin." *Remote Sensing* 12, no. 22: 3768. https://doi.org/10.3390/rs12223768.

Sleziak, P., J. Szolgay, K. Hlavčová, M. Danko, and J. Parajka. 2020. "The Effect of the Snow Weighting on the Temporal Stability of Hydrologic Model Efficiency and Parameters." *Journal of Hydrology* 583: 124639. https://doi.org/10.1016/j.jhydrol.2020.124639.

Smith, A., D. Tetzlaff, L. Kleine, M. Maneta, and C. Soulsby. 2021. "Quantifying the Effects of Land Use and Model Scale on Water Partitioning and Water Ages Using Tracer-Aided Ecohydrological Models." *Hydrology and Earth System Sciences* 25, no. 4: 2239–2259. https://doi.org/10.5194/hess-25-2239-2021.

Sobol, I. M. 2001. "Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates." *Mathematics and Computers in Simulation* 55, no. 1–3: 271–280. https://doi.org/10.1016/S0378-4754(00)00270-6.

Soltani, M., E. Bjerre, J. Koch, and S. Stisen. 2021. "Integrating Remote Sensing Data in Optimization of a National Water Resources Model to Improve the Spatial Pattern Performance of Evapotranspiration." *Journal of Hydrology* 603: 127026.

Soulsby, C., K. Piegat, J. Seibert, and D. Tetzlaff. 2011. "Catchment-Scale Estimates of Flow Path Partitioning and Water Storage Based on Transit Time and Runoff Modelling." *Hydrological Processes* 25, no. 25: 3960–3976.

Soulsby, C., D. Tetzlaff, and M. Hrachowitz. 2009. "Tracers and Transit Times: Windows for Viewing Catchment Scale Storage?" *Hydrological Processes* 23: 3503–3507.

Sprenger, M., R. W. H. Carroll, J. Dennedy-Frank, et al. 2022. "Variability of Snow and Rainfall Partitioning Into Evapotranspiration and Summer Runoff Across Nine Mountainous Catchments." *Geophysical Research Letters* 49, no. 13: e2022GL099324. https://doi.org/10.1029/2022GL099324.

Stadnyk, T. A., and T. L. Holmes. 2023. "Large Scale Hydrologic and Tracer Aided Modelling: A Review." *Journal of Hydrology* 618: 129177.

Stampoulis, D., J. T. Reager, C. H. David, et al. 2019. "Model-Data Fusion of Hydrologic Simulations and GRACE Terrestrial Water Storage Observations to Estimate Changes in Water Table Depth." *Advances in Water Resources* 128: 13–27. https://doi.org/10.1016/j.advwatres.2019.04.004.

Staudinger, M., J. Seibert, and H. J. van Meerveld. 2021. "Representation of bi-Directional Fluxes Between Groundwater and Surface Water in a Bucket-Type Hydrological Model." *Water Resources Research* 57, no. 9: e2020WR028835.

Staudinger, M., M. Stoelzle, F. Cochand, J. Seibert, M. Weiler, and D. Hunkeler. 2019. "Your Work Is My Boundary Condition! Challenges and Approaches for a Closer Collaboration Between Hydrologists and Hydrogeologists." *Journal of Hydrology* 571: 235–243. https://doi.org/10.1016/j.jhydrol.2019.01.058.

Stisen, S., J. Koch, T. O. Sonnenborg, et al. 2018. "Moving Beyond Run-Off Calibration-Multivariable Optimization of a Surface-Subsurface-Atmosphere Model." *Hydrological Processes* 32, no. 17: 2654–2668. https://doi.org/10.1002/hyp.13177.

Stoelzle, M., M. Weiler, K. Stahl, A. Morhard, and T. Schuetz. 2015. "Is There a Superior Conceptual Groundwater Model Structure for Baseflow Simulation?" *Hydrological Processes* 29: 1301–1313. https://doi.org/10.1002/hyp.10251.

Sulistioadi, Y. B., K. H. Tseng, C. K. Shum, et al. 2015. "Satellite Radar Altimetry for Monitoring Small Rivers and Lakes in Indonesia." *Hydrology and Earth System Sciences* 19, no. 1: 341–359. https://doi.org/10.5194/hess-19-341-2015.

Sun, W., H. Ishidaira, and S. Bastola. 2012. "Calibration of Hydrological Models in Ungauged Basins Based on Satellite Radar Altimetry Observations of River Water Level." *Hydrological Processes* 26, no. 23: 3524–3537. https://doi.org/10.1002/hyp.8429.

Sutanudjaja, E. H., L. P. H. van Beek, S. M. de Jong, F. C. van Geer, and M. F. P. Bierkens. 2014. "Calibrating a Large-Extent High-Resolution Coupled Groundwater-Land Surface Model Using Soil Moisture and Discharge Data." *Water Resources Research* 50, no. 1: 687–705. https://doi.org/10.1002/2013wr013807.

Széles, B., J. Parajka, P. Hogan, et al. 2020. "The Added Value of Different Data Types for Calibrating and Testing a Hydrologic Model in a Small Catchment." *Water Resources Research* 56, no. 10: e2019WR026153. https://doi.org/10.1029/2019WR026153.

Széles, B., J. Parajka, P. Hogan, et al. 2021. "Stepwise Prediction of Runoff Using Proxy Data in a Small Agricultural Catchment." *Journal of Hydrology and Hydromechanics* 69, no. 1: 65–75. https://doi.org/10.2478/johh-2020-0029.

Tarasova, L., M. Knoche, J. Dietrich, and R. Merz. 2016. "Effects of Input Discretization, Model Complexity, and Calibration Strategy on Model Performance in a Data-Scarce Glacierized Catchment in Central Asia." *Water Resources Research* 52, no. 6: 4674–4699. https://doi.org/10.1002/2015wr018551.

Tarpanelli, A., G. Amarnath, L. Brocca, C. Massari, and T. Moramarco. 2017. "Discharge Estimation and Forecasting by MODIS and Altimetry Data in Niger-Benue River." *Remote Sensing of Environment* 195: 96–106. https://doi.org/10.1016/j.rse.2017.04.015.

Tarpanelli, A., S. Barbetta, L. Brocca, and T. Moramarco. 2013. "River Discharge Estimation by Using Altimetry Data and Simplified Flood Routing Modeling." *Remote Sensing* 5, no. 9: 4145–4162. https://doi.org/10.3390/rs5094145.

Tarpanelli, A., S. Camici, K. Nielsen, L. Brocca, T. Moramarco, and J. Benveniste. 2021. "Potentials and Limitations of Sentinel-3 for River Discharge Assessment." *Advances in Space Research* 68, no. 2: 593–606. https://doi.org/10.1016/j.asr.2019.08.005.

Tetzlaff, D., C. Birkel, J. Dick, J. Geris, and C. Soulsby. 2014. "Storage Dynamics in Hydropedological Units Control Hillslope Connectivity, Runoff Generation, and the Evolution of Catchment Transit Time Distributions." *Water Resources Research* 50, no. 2: 969–985.

Tetzlaff, D., J. Buttle, S. K. Carey, K. McGuire, H. Laudon, and C. Soulsby. 2015. "Tracer-Based Assessment of Flow Paths, Storage and Runoff Generation in Northern Catchments: A Review." *Hydrological Processes* 29, no. 16: 3475–3490.

Tetzlaff, D., H. Laudon, S. Luo, and C. Soulsby. 2024. "Ecohydrological Resilience and the Landscape Water Storage Continuum in Droughts." *Nature Water* 2: 915–918. https://doi.org/10.1038/s44221-024-00300-y.

Tigabu, T. B., P. D. Wagner, G. Hörmann, and N. Fohrer. 2020. "Modelling the Spatio-Temporal Flow Dynamics of Groundwater-Surface Water Interactions of the Lake Tana Basin, Upper Blue Nile, Ethiopia." *Hydrology Research* 51, no. 6: 1537–1559. https://doi.org/10.2166/nh.2020.046.

Tong, R., J. Parajka, B. Széles, et al. 2022. "The Value of Satellite Soil Moisture and Snow Cover Data for the Transfer of Hydrological Model Parameters to Ungauged Sites." *Hydrology and Earth System Sciences* 26, no. 7: 1779–1799. https://doi.org/10.5194/hess-26-1779-2022.

Tourian, M. J., C. Schwatke, and N. Sneeuw. 2017. "River Discharge Estimation at Daily Resolution From Satellite Altimetry Over an Entire

River Basin." *Journal of Hydrology* 546: 230–247. https://doi.org/10.1016/j.jhydrol.2017.01.009.

Tran, B. N., J. van der Kwast, S. Seyoum, R. Uijlenhoet, G. Jewitt, and M. Mul. 2023. "Uncertainty Assessment of Satellite Remote-Sensing-Based Evapotranspiration Estimates: A Systematic Review of Methods and Gaps." *Hydrology and Earth System Sciences* 27: 4505–4528. https://doi.org/10.5194/hess-27-4505-2023.

Tripathi, V., A. Groh, M. Horwath, and R. Ramsankaran. 2022. "Scaling Methods of Leakage Correction in GRACE Mass Change Estimates Revisited for the Complex Hydro-Climatic Setting of the Indus Basin." *Hydrology and Earth System Sciences* 26, no. 17: 4515–4535. https://doi.org/10.5194/hess-26-4515-2022.

Tsai, Y.-L. S., A. Dietz, N. Oppelt, and C. Kuenzer. 2019. "Remote Sensing of Snow Cover Using Spaceborne SAR: A Review." *Remote Sensing* 11, no. 12: 1456. https://doi.org/10.3390/rs11121456.

Tuo, Y., G. Marcolini, M. Disse, and G. Chiogna. 2018. "A Multi-Objective Approach to Improve SWAT Model Calibration in Alpine Catchments." *Journal of Hydrology* 559: 347–360. https://doi.org/10.1016/j.jhydrol.2018.02.055.

Tuttle, S. E., J. M. Jacobs, C. M. Vuyovich, C. Olheiser, and E. Cho. 2018. "Intercomparison of Snow Water Equivalent Observations in the Northern Great Plains." *Hydrological Processes* 32: 817–829. https://doi.org/10.1002/hyp.11459.

Vervoort, R. W., S. F. Miechels, F. F. van Ogtrop, and J. H. A. Guillaume. 2014. "Remotely Sensed Evapotranspiration to Calibrate a Lumped Conceptual Model: Pitfalls and Opportunities." *Journal of Hydrology* 519: 3223–3236. https://doi.org/10.1016/j.jhydrol.2014.10.034.

Wagner, W., G. Lemoine, and H. Rott. 1999. "A Method for Estimating Soil Moisture From ERS Scatterometer and Soil Data." *Remote Sensing of Environment* 70, no. 2: 191–207.

Wang, Y., J. Hu, R. Li, et al. 2023. "Increasing Cloud Coverage Deteriorates Evapotranspiration Estimating Accuracy From Satellite, Reanalysis and Land Surface Models Over East Asia." *Geophysical Research Letters* 50: e2022GL102706. https://doi.org/10.1029/2022GL102706.

Werth, S., and A. Güntner. 2010. "Calibration Analysis for Water Storage Variability of the Global Hydrological Model WGHM." *Hydrology and Earth System Sciences* 14: 59–78. https://doi.org/10.5194/hess-14-59-2010.

Werth, S., A. Güntner, S. Petrovic, and R. Schmidt. 2009. "Integration of GRACE Mass Variations Into a Global Hydrological Model." *Earth and Planetary Science Letters* 277, no. 1–2: 166–173.

Wiederhold, H., A. J. Kallesøe, R. Kirsch, R. Mecking, R. Pechnig, and F. Skowronek. 2021. "Geophysical Methods Help to Assess Potential Groundwater Extraction Sites." *Grundwasser* 26, no. 4: 367–378. https://doi.org/10.1007/s00767-021-00495-x.

Winkler, M., H. Schellander, and S. Gruber. 2021. "Snow Water Equivalents Exclusively From Snow Depths and Their Temporal Changes: The Δsnow Model." *Hydrology and Earth System Sciences* 25, no. 3: 1165–1187. https://doi.org/10.5194/hess-25-1165-2021.

Wu, S., D. Tetzlaff, X. Yang, A. Smith, and C. Soulsby. 2023. "Integrating Tracers and Soft Data Into Multi-Criteria Calibration: Implications From Distributed Modeling in a Riparian Wetland." *Water Resources Research* 59, no. 11: e2023WR035509.

Wu, S., D. Tetzlaff, X. Yang, and C. Soulsby. 2022. "Disentangling the Influence of Landscape Characteristics, Hydroclimatic Variability and Land Management on Surface Water NO3-N Dynamics: Spatially Distributed Modelling Over 30 Yr in a Lowland Mixed Land Use Catchment." *Water Resources Research* 58, no. 2: e2021WR030566. https://doi.org/10.1029/2021WR030566.

Xiao, L., T. Che, and L. Dai. 2020. "Evaluation of Remote Sensing and Reanalysis Snow Depth Datasets Over the Northern Hemisphere During 1980–2016." *Remote Sensing* 12, no. 19: 3253. https://www.mdpi.com/2072-4292/12/19/3253.

Xie, J., X. Liu, S. Jasechko, et al. 2024. "Majority of Global River Flow Sustained by Groundwater." *Nature Geoscience* 17: 770–777. https://doi.org/10.1038/s41561-024-01483-5.

Xu, X., J. Li, and B. A. Tolson. 2014. "Progress in Integrating Remote Sensing Data and Hydrologic Modeling." *Progress in Physical Geography* 38, no. 4: 464–498.

Yang, X., D. Tetzlaff, C. Mueller, K. Knoeller, D. Borchardt, and C. Soulsby. 2023. "Upscaling Tracer-Aided Ecohydrological Modelling to Larger Catchments: Implications for Process Representation and Heterogeneity in Landscape Organization." *Water Resources Research* 59: e2022WR033033. https://doi.org/10.1029/2022WR033033.

Yang, X., S. Tian, W. Feng, et al. 2020. "Spatio-Temporal Evaluation of Water Storage Trends From Hydrological Models Over Australia Using GRACE Mascon Solutions." *Remote Sensing* 12: 3578. https://doi.org/10.3390/rs12213578.

Yassin, F., S. Razavi, H. Wheater, G. Sapriza-Azuri, B. Davison, and A. Pietroniro. 2017. "Enhanced Identification of a Hydrologic Model Using Streamflow and Satellite Water Storage Data: A Multicriteria Sensitivity Analysis and Optimization Approach." *Hydrological Processes* 31: 3320–3333. https://doi.org/10.1002/hyp.11267.

Yoon, Y., P. A. Garambois, R. C. Paiva, M. Durand, H. Roux, and E. Beighley. 2016. "Improved Error Estimates of a Discharge Algorithm for Remotely Sensed River Measurements: Test Cases on Sacramento and Garonne Rivers." *Water Resources Research* 52, no. 1: 278–294. https://doi.org/10.1002/2015WR017319.

Yu, X., G. Bhatt, C. J. Duffy, et al. 2015. "A Coupled Surface-Subsurface Modeling Framework to Assess the Impact of Climate Change on Freshwater Wetlands." *Climate Research* 66: 211–228. https://doi.org/10.3354/cr01348.

Zakharova, E., K. Nielsen, G. Kamenev, and A. Kouraev. 2020. "River Discharge Estimation From Radar Altimetry: Assessment of Satellite Performance, River Scales and Methods." *Journal of Hydrology* 583: 124561. https://doi.org/10.1016/j.jhydrol.2020.124561.

Zhang, X., L. Jiang, Z. Liu, et al. 2023. "Flow Regime Changes in the Lancang River, Revealed by Integrated Modeling With Multiple Earth Observation Datasets." *Science of the Total Environment* 862: 160656. https://doi.org/10.1016/j.scitotenv.2022.160656.

Zhang, Y., F. H. S. Chiew, L. Zhang, and H. Li. 2009. "Use of Remotely Sensed Actual Evapotranspiration to Improve Rainfall-Runoff Modeling in Southeast Australia." *Journal of Hydrometeorology* 10: 969–980.

Zink, M., L. Samaniego, R. Kumar, et al. 2016. "The German Drought Monitor." *Environmental Research Letters* 11, no. 7: 074002. https://doi.org/10.1088/1748-9326/11/7/074002.

Zipper, S., K. M. Befus, R. Reinecke, et al. 2023. "GroMoPo: A Groundwater Model Portal for Findable, Accessible, Interoperable, and Reusable (FAIR) Modeling." *Groundwater* 61, no. 6: 764–767. https://doi.org/10.1111/gwat.13343.

Zitzler, E., M. Laumanns, and L. Thiele. 2001. "SPEA 2: Improving the Strength Pareto Evolutionary Algorithm. Technical Report 103 Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH) Zurich."

Zreda, M., W. J. Shuttleworth, X. Zeng, et al. 2012. "COSMOS: The COsmic-Ray Soil Moisture Observing System." *Hydrology and Earth System Sciences* 16, no. 11: 4079–4099. https://doi.org/10.5194/hess-16-4079-2012.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Appendix A

TABLE A1 | List of data sets on hydrological variables, including main features.

Variable	Data set/Product	Spatial sampling	Temporal range	Temporal resolution	Obser-vation technique	Download link
Snow	MODIS Snow Cover (also provides raw NDSI values for calculation for fractional SCE; National Snow and Ice Data Center)	500 m	2002–present	Daily/8 days	Satellite	https://nsidc.org/data/myd10a1/versions/6#anchor-1; https://nsidc.org/data/mod10a1/versions/6
	Landsat Fractional Snow-Covered Area (for Landsat 4–5 TM, Landsat 7 ETM+, and Landsat 8 OLI/TIRS; USGS)	30 m	1984–present	16 days	Satellite	https://www.usgs.gov/landsat-missions/landsat-fractional-snow-covered-area-science-products
	GlobSnow v3.0 Northern Hemisphere snow water equivalent dataset (based on SMMR, SSM/I, and SSMIS sensors combined with ground-based weather station data; European Space Agency)	25 km	1979–2018	Daily or monthly	Satellite	https://www.globsnow.info/swe/archive_v3.0/
	Copernicus Global Land Service—SWE	5 km	2006–present	Daily	Satellite	https://land.copernicus.eu/global/products/swe
Soil mosture	SMAPL3E (based on SMAP radiometer)	9 km	2015–present	1–3 days	Satellite	https://smap.jpl.nasa.gov/data/
	ASCAT	25 km	2006–present	1-2d	Satellite	https://hsaf.meteoam.it/Products/ProductsList?type=soil_moisture
	ERS-SCAT	~50 km	1992–2011	3-4d	Satellite	https://earth.esa.int/eogateway/catalog/ers-2-scatterometer-surface-soil-moisture-time-series-and-orbit-product-in-high-and-nominal-resolution-ssm-h-n-ts-ssm-h-n
	AMSR-E	25 km	2002–2011	Daily	Satellite	https://nsidc.org/data/ae_land3
	AMSR-2	25 km	2012–present	Daily	Satellite	http://nsidc.org/data/au_land
	SMOS	35 km	2009–present	Daily	Satellite	https://earth.esa.int/web/guest/-/ level-2-soil-moisture-6900
	ESA-CCI Soil Moisture	0.25°	1978–2018	daily	satellite	https://www.esa-soilmoisture-cci.org/node/145
	SCATSAR (SWI product based on ASCAT and Sentinel 1 CSAR) (available for Europe)	1 km	2015–present	Daily	Satellite	https://land.copernicus.eu/global/products/swi
	SWI product based on ASCAT	0.1°	2007–present	Daily	Satellite	https://land.copernicus.eu/global/products/SWI
	International soil moisture monitoring network (ISMN)	n.a.	1952–present		In situ	https://ismn.earth/
	SoMo.ml	0.25°	2000–2019	Daily	In situ combined with machine learning	https://springernature.figshare.com/collections/Global_soil_moisture_from_in_situ_measurements_using_machine_learning_-_SoMo_ml/5142185

(Continues)

TABLE A1 | (Continued)

Variable	Data set/Product	Spatial sampling	Temporal range	Temporal resolution	Obser-vation technique	Download link
Groundwater	Global patterns of groundwater table depth	1,603,781 sites			In situ	https://www.science.org/doi/10.1126/science.1229881
Terrestrial water storage	GRACE/GRACE-FO (Terrestrial water storage anomalies)	~100 km	2002–present	Monthly	Satellite	GFZ: https://isdc.gfz-potsdam.de/grace-isdc/JPL: https://grace.jpl.nasa.gov/data/get-data/CSR: https://www2.csr.utexas.edu/grace/
	Precipitation isotope data: The Global Network of Isotopes in Precipitation (GNIP) data	n.a.	Initiated in 1960	Usually monthly	In situ	https://www.iaea.org/services/networks/gnip
	Streamwater isotope data: The Global Network of Isotopes in Rivers (GNIR) data	n.a.	Initiated in 2002	Usually monthly	In situ	https://www.iaea.org/services/networks/gnir
Evapotranspiration	MODIS ET/MOD16A2	0.5 km	2000–present	8-day composite	Satellite	https://lpdaac.usgs.gov/products/mod16a2v006/ or https://lpdaac.usgs.gov/products/mod16a2v061/
	MODIS ET/MOD16A2	0.5 km	2000–present	Annual	Satellite	https://lpdaac.usgs.gov/products/mod16a3v006/ or https://lpdaac.usgs.gov/products/mod16a3gfv061/
	GLEAM/GLEAM v3.6a	0.25°	1980–2021	Daily	Satellite	https://www.gleam.eu/
	GLEAM/GLEAM v3.6b	0.25°	2003–2021	Daily	Satellite	https://www.gleam.eu/
	SSEBop	1 km	Depends on region, 2000-present	Depends on region; daily, monthly, annual	Satellite	https://earlywarning.usgs.gov/ssebop/modis/daily https://earlywarning.usgs.gov/fews
	Fluxnet/FLUXNET2015	212 sites globally	Depends on site, up to 2014	Half-hourly to yearly	In situ	https://fluxnet.org/data/fluxnet2015-dataset/
	FluxCom	RS: 0.0833°; RS+ METEO: 0.5°	RS: 2001–2015; RS + METEO depending on climate forcing	RS: 8-daily; RS + METEO: daily	In situ, satellite; combined with machine learning	https://fluxcom.org/
Altimetric water level	Hydroweb	n.a.	1992–2011	35 days	Satellite	https://hydroweb.theia-land.fr/
	Global Reservoirs and Lakes Monitor	n.a.	1991–present	10–35 days	Satellite	https://ipad.fas.usda.gov/cropexplorer/global_reservoir/
	DAHITI	n.a.	1992–present	10–35 days	Satellite	https://dahiti.dgfi.tum.de/

TABLE A2 | List of past and current satellite missions providing estimates of water surface elevation, including main features.

Mission	Space Agency	Time period	Repeat period (d)	Inter-track distance (km)	Along-track distance (m)
Geosat	US Navy	1985–1990	17	165	
ERS-1	ESA	1991–2000	35	80	
TOPEX/Poseidon	NASA/CNES	1992–2005	10	315	620
ERS-2	ESA	1995–2011	35	80	374
Jason-1	CNES/NASA	2001–2013	10	315	294
Envisat	ESA	2002–2012	35	80	374
Jason-2	CNES/NASA/ EUMETSAT/NOAA	2008–2019	10	315	294
SARAL/AltiKa	CNES	2013–present	35	80	173
Jason-3	CNES/NASA/ EUMETSAT/NOAA	2016–present	10	315	
Sentinel-3A/B	ESA	2016–present	27	52	
Jason-CS/Sentinel-6	EUMETSAT/NASA	2020–present	10	315	