https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

Open Access Earth System Science Data Discussions

State-of-the-art hydrological datasets exhibit low water balance consistency globally

Hao Huang[1], Junguo Liu[1,2], Aifang Chen[3], Melissa Ruiz-Vásquez[4,5], René Orth[5]

[1]School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China

[2]Henan Provincial Key Laboratory of Hydrosphere and Watershed Water Security, North China University of Water Resources and Electric Power, Zhengzhou, China

[3]School of Environment and Civil Engineering, Dongguan University of Technology, Dongguan, China

[4]Max Planck Institute for Biogeochemistry, Jena, Germany

[5]Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany

Correspondence to: Junguo Liu (junguo.liu@gmail.com)

Abstract

The proliferation and diversification of hydrological datasets have significantly advanced hydrological research. However, the coherence across these datasets remains poorly understood, hindering the comparability of findings derived from different data sources and variables. Here, we demonstrate that state-of-the-art hydrological datasets exhibit overall low consistency when evaluated through the lens of water balance – specifically, the relationship between variations in soil moisture and the difference between precipitation, evapotranspiration, and runoff. Our analysis reveals that satellite-based precipitation datasets generally show the highest consistency, while gauge-based datasets perform better in densely monitored regions of the Northern Hemisphere. For evapotranspiration, runoff, and soil moisture, reanalysis datasets demonstrate broader areas of higher consistency compared to gauge- or satellite-based products. Spatial patterns of consistency are strongly influenced by aridity and temperature, which affect measurement and modelling accuracy, while vegetation cover further modulates the performance of soil moisture datasets. Notably, dataset consistency has improved significantly in northern mid-latitudes over recent decades, likely reflecting advancements in observational technologies and the effects of climate warming. These findings underscore the importance of continued efforts to enhance dataset coherence and reliability for robust hydrological assessments.

1 Introduction

Over the past decades, the advancement of hydrological science and interconnected water-related research fields was accompanied by the emergence of datasets that depict the spatiotemporal changes of variables in the water cycle (Tang et al., 2024; Zarei and Destouni, 2024; Gebrechorkos et al., 2024; Douville et al., 2021; Oki and Kanae, 2006; Wang-Erlandsson et al., 2022; Mehta et al., 2024; Markonis et al., 2024). At the same time, understanding the consistency across the increasing suite of datasets is crucial not only for research on the responses and interactions within hydrology, but also for practitioners and management in terms of regional water scarcity (Mekonnen and Hoekstra, 2016; Mehta et al., 2024), ecosystem function and water availability (Denissen et al., 2022), and the Earth system resilience (Wang-Erlandsson et al., 2022; Jaramillo and Destouni, 2015). Nevertheless, the water balance consistency among different suites remains largely unknown, while current studies mostly detail the dataset performance in terms of accuracies against observations and/or reference data, modeling behaviors, or water and energy balance closure (Tang et al., 2024; Gebrechorkos et al., 2024; Pan et al., 2020; Zarei and Destouni, 2024; Abolafia-Rosenzweig et al., 2021).

Gridded hydrological datasets are derived based on different types of observations and methods, such as (i) spatial interpolation based on gauge/station/*in-situ* measurements (Dorigo et al., 2011; Pastorello et al., 2020; Do et al., 2018; Harris et al., 2020), (ii) radiative transfer modelling based on satellite measurements (Cooley et al., 2022; Mccabe et al., 2017), (iii) land surface modelling with integrated data assimilation of hydrological and other variables (Muñoz-Sabater et al., 2021). Meanwhile, datasets also could be developed based on a combination of these approaches and observations (Beck et al., 2019; Yao et al., 2014). In this context, each of these approaches is characterized by inherent advantages and disadvantages. For example, in the case of precipitation (*P*), gauge-based datasets are based on ground truth but at the same time they are influenced by errors related to wind and air flow anomalies around the gauges, by the spatial distribution of gauges which potentially misses some of the spatial heterogeneity of precipitation patterns, and by uncertainties in spatial interpolation (Lanza et al., 2022; Mishra and Coulibaly, 2009; La et al., 2002). By contrast, satellite-based *P* datasets can capture spatial patterns more consistently (Tang et al., 2022; Ashouri et al., 2015; Funk et al., 2015), but have difficulties in estimating *P* amounts arriving to the surface. Further, reanalysis datasets based on land surface models show strengths in addressing temporal gaps caused by missing records and incomplete observation periods (Hersbach et al., 2020; Gelaro et al., 2017), but suffer from inaccurate or incomplete consideration of land surface processes affecting hydrological dynamics. From this aspect, machine-learning algorithms present an alternative opportunity to produce seamless data, firstly happening for estimating *P* (Ashouri et al., 2015) and recently also being applied to evapotranspiration (*ET*), runoff (*R*), and soil moisture (*SM*) datasets (Nelson et al., 2024; Ghiggi et al., 2019; O and Orth, 2021).

As a result of the different derivation approaches and the influence of environmental factors, disagreements between hydrological datasets remain (Hirschi et al., 2025; Markonis et al., 2024; Sun et al., 2018). These uncertainties limit the fundamental understanding of patterns, changes, and variabilities of water variables (Markonis et al., 2024; Wang et al., 2024; Han et al., 2024; Douville et al., 2021; Greve et al., 2014; Zhang et al., 2024a; Denissen et al., 2022). The scarcity of observations across time, space, and hydrological variables hinders a comprehensive analysis of datasets' performance and reliability. However, observations are not our only source of knowledge about Nature, but known physical laws also provide information. This way, for example the water balance equation can be used to evaluate the consistency across combinations of hydrological datasets, a question which has remained largely unclear because assessments are usually specific to individual datasets (Zarei and Destouni, 2024; Abolafia-Rosenzweig et al., 2021). Such a combinatorial and factorial analysis requires (i) gridded datasets of all involved variables and (ii) independence between them in the sense that they are not derived with, e.g., the same model or approach which inherently enforces water balance closure. Thanks to the recent emergence of many hydrological datasets (Muñoz-Sabater et al., 2021; Ghiggi et al., 2019; Miralles et al., 2025), these requirements are now met, opening a novel opportunity for hydrological dataset evaluation.

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

In this study, we evaluate the water balance consistency across a comprehensive set of *P*, *ET*, *R* and *SM* datasets. This encompasses gauge/station-based, satellite-based and reanalysis-based datasets, and offers 8,294 combinations of water balance-variables from independently derived datasets (Fig. 1a). For each combination, we evaluate adjusted R^2 as the performance of linear regression of temporal changes in $P-ET-R$ against changes in *SM* (ΔSM) to determine its water balance consistency. Then, combining an individual dataset with all possible combinations of datasets for the remaining water balance-variables we can assess its performance through the average of the R^2 scores obtained for all considered combinations. This way, the common limitations and strengths of different derivation-based datasets for each variable (i.e., *P*, *ET*, *R*, and *SM*) are distinguished across space and time. In addition to determining the performance of a large set of considered hydrological datasets across the globe, we also evaluate the resulting spatial patterns for possible causes in order to provide guidance for further dataset development.

2 Materials and Methods

2.1 Data and Independent combinations

We utilized 20 *P* datasets, 11 *ET* datasets, 7 *R* datasets, and 9 *SM* datasets to obtain respective monthly values across the global land area, where the *P*, *ET*, and *R* values are monthly amounts and ΔSM values are the soil moisture differences between the last day and the first day of each month. According to their sources, these datasets were summarized into three categories:

- Gauge/station-based products: CPC (Xie et al., 2010), CRU TS v4.06 (Harris et al., 2020), UDel v5.01 (Legates and Willmott, 1990), EM-EARTH (Tang et al., 2022), GPCC v2022 (Schneider et al., 2022), and PREC/L (Chen et al., 2002) for *P*, X-BASE (Nelson et al., 2024) for *ET*, GRUN (Ghiggi et al., 2019) for *R*, as well as SoMo.ml (O and Orth, 2021) for ΔSM.
- Satellite-based products: CHIRPS v2.0 (Funk et al., 2015), CMAP (Xie and Arkin, 1997), CMORPH v1 (Xie et al., 2017), GPCP(M) v2.3 (Adler et al., 2018), GPCP(D) v1.3 (Huffman et al., 2001), GPM IMERG v07 (Huffman et al., 2023), PERSIANN-CDR (Ashouri et al., 2015), MSWEP v2.8 (Beck et al., 2019) for *P*, MODIS (Running et al., 2021), PT-JPL (Fisher et al., 2008), PML-v2 (Zhang et al., 2019), GLASS (Yao et al., 2014) for *ET*, GLEAM v4.1 (Miralles et al., 2025) for both *ET* and ΔSM, SMAP L4 v7 (Reichle et al., 2019) for both *R* and ΔSM, as well as ESA CCI v08.1 (Gruber et al., 2019) for ΔSM.
- Reanalysis products: 20CR v3 (Slivinski et al., 2021), JRA-55 (Japan Meteorological Agency, 2013), ERA5 (Hersbach et al., 2020), NCEP-NCAR R1 (Kistler et al., 2001), and NCEP-DOE R2 (Kanamitsu et al., 2002) for *P*, MERRA-2 (Gelaro et al., 2017) for *P*, *ET*, *R*, and ΔSM, as well as GLDAS-2.0 (Rodell et al., 2004), GLDAS-2.1 (Rodell et al., 2004), GLDAS-2.2 (Li et al., 2019), ERA5-land (Muñoz-Sabater et al., 2021) for *ET*, *R*, and ΔSM.

All the datasets were either provided in or resampled in 0.25-degree resolution by linear interpolation, and their temporal coverages are within the period of Jan-2000 to Dec-2022. Various components of GLDAS (i.e., -2.0, -2.1, and -2.2) were used here because they are based on different forcings, models, and data assimilation strategies (see more details in Tables S1−S4). The ΔSM from different datasets were the depth-weighted averages of their available soil layers (Li et al., 2023a). SoMo.ml ΔSM covers 0−50 cm related to the commonly observed depths in *in-*

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

situ measurements; ESA CCI ΔSM represents the top surface layer (of < 2 cm thickness) captured by satellite observations; GLEAM ΔSM, SMAP ΔSM, and MERRA-2 ΔSM represent a root zone layer of 0−100 cm; and GLDAS-2.0/2.1/2.2 and ERA5-land ΔSM cover deeper depths (> 100 cm). Despite different depths, the ΔSM was assumed to record the variability of $P - ET - R$ in the water balance, on the basis that its variability constitutes a large portion of the variability in terrestrial water storage (Freedman et al., 2014). In this context, a suite of P, ET, R, and SM datasets forms a considered combination, such as

$$P_{\mathrm{CPC}},\ ET_{\mathrm{X\text{-}BASE}},\ R_{\mathrm{GUN}},\ \text{and}\ \Delta SM_{\mathrm{SoMo.ml}}$$

Among the considered datasets as listed above, 13,860 combinations (that is 20×11×7×9 for P, ET, R, and S) were initially available. However, considering the temporal availability and the dependence between dataset sources of different water balance components, parts of combinations were excluded by three rules:

1) The combinations with short overlapping time periods cannot be considered. In particular,

- SMAP L4 products have only one year overlap (i.e., 2015) with 20CR v3, so the combinations with P from 20CR v3 and R and/or ΔSM from SMAP were not considered;
- The combinations with GRUN R (covering until 2014) and R and/or ΔSM from SMAP (starting from 2015) were not available;
- The combinations with water balance components from GLDAS-2.0 (also covering until 2014) and SMAP were not available;
- The combinations with SMAP L4 products and either PT-JPL or UDel v5.01 (covering until 2017) were not considered.

2) The combinations with water balance components from the same dataset source were not considered, which include the combinations with GLEAM ET and ΔSM, the combinations with SMAP ET and ΔSM, and the combinations with any two or more variables from MERRA-2/GLDAS/ERA5-land. In this perspective, since the difference between ERA5 and ERA5-land was mainly because of the non-linear dynamical downscaling technique (Muñoz-Sabater et al., 2021), the combinations with ERA5 P and ERA5-land $ET/R/\Delta S$ were also not considered.

3) If a dataset was driven by another dataset, the water balance components from these two datasets were also not considered in combination. In particular:

- GRUN was driven by GSWP3, a dynamically downscaled and bias-corrected version of the 20CR, so the combinations with 20CR P and GRUN R were excluded;
- SoMo.ml was driven by meteorological data from ERA5, so the combinations with ERA5 P and SoMo.ml ΔSM were excluded;
- PML-v2 used the GLDAS-2.1 meteorological forcings, which includes GPCP(D) v1.3, so the combinations with GPCP(D) P and PML ET, as well as those with GPCP(D) P and GLDAS-2.1 ET, were excluded;
- GLEAM v4.1 used P from MSWEP v2.8 as one of the inputs, so the combinations with MSWEP P and GLEAM $ET/\Delta SM$ were excluded;
- The inputted P for SMAP L4 was from CPC and GPCP(M), and therefore, the combinations with CPC/GPCP(M) P and SMAP $R/\Delta SM$ were excluded;
- Since the land surface component of MERRA-2 bias adjusted P by using CPC, CMAP, and GPCP(M), the combinations with CPC/CMAP/GPCP(M) P and MERRA-2 $ET/R/\Delta SM$ were excluded;

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

- The GLDAS-2.2 was forced with the meteorological analysis fields from the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS)(Rui et al., 2022), which includes ERA5, so the combinations with ERA5 P and GLDAS-2.2 $ET/R/\Delta SM$ were excluded.

At the same time, there are different levels of (in)dependence such that the decision on whether or not to consider certain datasets as independent is not always straightforward. The following cases are not fully independent but considered sufficiently independent for the context of our study:

- The datasets driven by similar forcings, such as SoMo.ml ΔSM and ERA5-land ET and R, are considered to form independent combinations;
- MSWEP generated based on a group of P datasets including ERA5 is considered sufficiently independent from the ERA5-land ET, R, and ΔSM;
- ESA CCI ΔSM which was assimilated into GLEAM ΔSM is considered independent from GLEAM ET.

After applying these exclusion rules, there remained 8,294 independent combinations.

2.2 Performance assessment in terms of water balance consistency

For each considered combination of hydrological datasets, we assess adjusted R^2 scores in water balance in each grid cell through a linear regression model considering all available months:

$$(\mathbf{P} - \mathbf{ET} - \mathbf{R})_s = k \cdot \mathbf{\Delta SM}_s \tag{1}$$

where s is the spatial index (grid cell) and k is the proportionality factor. Note that this is not supposed to equal to 1 in our context because of the differences in units between the left side of the equation (mm for **P**, **ET**, and **R**) and the right side ($m^3 \cdot m^{-3}$ for $\mathbf{\Delta SM}$). The linear regression model lets us avoid the conversion of $\mathbf{\Delta SM}$ unit from $m^3 \cdot m^{-3}$ to mm, reducing uncertainties from considering soil moisture datasets with different soil depths. **P**, **ET**, **R**, and $\mathbf{\Delta SM}$ are $M \times 1$ vectors, where M is the number of months. We removed the models with M smaller than 36 to ensure enough input data. The adjusted R^2 score for each model was used to represent the ability of each combination of datasets to describe the variability in water balance at each grid point. Since the water balance is a physical law that should be obeyed according to mass balance, the ability of describing variability here is attributed to the performance of each combination for each grid cell.

Since different independent combinations have different temporal coverages (i.e., different M), we analyzed whether the varying M would affect the accuracy results. For this purpose, a fixed study period of Feb-2003 to Dec-2014 (where M is fixed to be 143) was selected. We calculated the degree of water balance closure, evaluated based on the adjusted R^2 score, of all available independent combinations for this fixed M. We compared the R^2 values between varying M and fixed M for each considered combination by calculating their linearly regressed R^2 scores and slopes. Most of the regressed R^2 and slopes are distributed between 0.9 and 1 (Fig. S1). This indicates that the considered time period has no significant influence on the resulting degree of water balance closure. Therefore, we assessed the performance across different time periods for different combinations of datasets, depending on their temporal coverage and overlap (ensuring a minimum overlap of 3 years). This allows us to involve a larger number of combinations compared to a fixed M, while we also provided the results calculated based on the combinations used in the following temporal changes analysis, which have only large and less varying M (Fig.

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

S2).

In a next step, the overall performance in terms of water balance consistency for each individual dataset in each grid cell was inferred from the averaged R^2 across all combinations of datasets containing the respective dataset. In other words, the performance of each individual dataset is assessed through the R^2 scores in water balance for describing variability when combining it with all suitable combinations of state-of-the-art datasets for the other water balance components.

Since the performance inferred from water balance consistency is based on the ability to describe variability in water balance, different temporal resolutions directly affected magnitudes and frequencies of the variability (Maurer and Hidalgo, 2008). Accordingly, we repeated the above calculations for the datasets available at daily and yearly resolutions, where 3,647 and 8,294 (the same as monthly) independent combinations were considered, respectively (Text S1−S2).

2.3 Potential influence factors on dataset performance

To further understand how the global spatial patterns of each dataset's performance (Figs. S3−S6) were influenced, we considered a set of potential influence factors of the spatial patterns: soil texture, aridity index, tree cover fraction, area equipped for irrigation, monthly mean temperature, observation density, and topography. For the first five factors, we calculated them for each independent combination because the factors are changing from dataset to dataset, and then obtained the averages for each dataset through all the considered combinations that include this dataset. In detail,

- Soil clay content was used to indicate soil texture influence, since small particles and large surface areas can create small pore sizes to hold water tightly, affecting *SM* conditions and through local water cycles to influence other water variables (Cleophas et al., 2022). The clay contents were provided by the Harmonized World Soil Database version 2.0 (HWSD v2.0) (Nachtergaele et al., 2023) for seven soil layers, and the layers used for each independent combination were selected according to the depth of *SM* dataset in that combination and depth-weighted for a whole layer.
- Regarding the aridity index, we used the multi-year averages of **ET** and divided them by those of **P** to obtain an aridity index map for each independent combination (O and Orth, 2021; Li et al., 2022).
- The tree cover fraction from NASA Vegetation Continuous Fields Version 1 data product (Hansen and Song, 2018).
- Area equipped for irrigation from Mehta et al. (2024) were averaged among the available periods for each independent combination.
- The monthly mean 2m air temperature was averaged based on the daily average temperature from ERA5 and calculated for each month in each considered combination.
- Unlike the upper factors, the observation density is different from variable to variable, not from combination to combination. We counted the number of stations/sites for different observation networks of the water variables: CPC global stations for *P*, eddy covariance sites in FLUXNET2015 (Pastorello et al., 2020) and AmeriFlux for *ET*, streamflow stations in the Global Streamflow Indices and Metadata Archive (GSIM) (Do et al., 2018) for *R*, and sites of *in-situ* measurement in the International Soil Moisture Network (ISMN) (Dorigo et al., 2011) and the National Center for Monitoring and Early Warning of Natural Disasters of Brazil (CEMADEN) (Zeri et al., 2020) for *SM*. Here, we referred to Ruiz-Vásquez et al. (2022) to sum up the stations/sites located in each grid cell and its

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

eight neighboring grid cells (Fig. S7).

- The topography information is represented by the standard deviation of 15 arc-second elevation (Noaa National Centers for Environmental Information, 2022) within each 0.25º grid cell.

An explainable machine learning method was applied for quantitative attribution (Li et al., 2022) in order to determine the relative roles of the considered factors for the resulting global spatial patterns of each dataset's performance. For each dataset, we trained one random forest model, where the global performance map was the target variable, the seven maps of the above-described factors were the predictors, and a common hyperparameter setting (numbers of estimators: 100; maximum features: 30%) was used (Li et al., 2022). Before training, the correlation matrix of the seven predictors for each random forest model was calculated to evaluate the potential collinearity between predictors. Since the correlations are within a range of -0.5−0.6 (Figs. S8−S11 for *P*, *ET*, *R*, and *SM* datasets, respectively), collinearity is not a major issue to affect our model predictions (Dormann et al., 2012). The performance of random forest models was determined by the cross-validation out-of-bag R^2, which mainly distributes around 0.8 for all the trained models and therefore indicates the usefulness of these models for the following attribution (Fig. S12). Then, SHapley Additive exPlanations (SHAP) feature importance was calculated to quantify the marginal contributions of predictors to each dataset's overall accuracy (Li et al., 2023a), and we identified the relative importance among predictors by ranking their global averaged absolute SHAP values (Li et al., 2023b).

2.4 Temporal changes in dataset performance

Since the temporal coverages of independent combinations are inconsistent, the independent combinations with less than two thirds of available monthly data for either the first period of Jan-2000 to Dec-2010 or the second period of Jan-2011 to Dec-2022 were removed in the temporal changes analysis. The remaining independent combinations (n = 2,589) were used to calculate the water balance consistency for the first period and the second period separately. The overall performance in terms of water balance consistency was calculated for the first or the second periods of each dataset by averaging the respective period's adjusted R^2 scores among all the independent combinations of the datasets considered in this study. In this way, the temporal change in performance for each dataset was obtained by subtracting the overall performance of the first period from that of the second period.

To account for the uncertainties of these temporal changes, bootstrap confidence intervals (Kulesa et al., 2015) were calculated for the performance in both the first and the second periods of the 2,589 independent combinations. For each of these independent combinations, whose number of available monthly data for the first and the second periods are denoted as M_1 and M_2, respectively, we obtained 100 random samples for the first/second period with replacement. The amount of data in one sample is M_1 for the first period and M_2 for the second period, and 100 samples indicate that 100 adjusted R^2 scores were calculated for the first/second period based on equation (1). Accordingly, a bootstrap distribution for the first/second period with 100 samples was obtained, and its confidence interval was evaluated by the 5th and 95th percentiles. When the 5th percentile of the second period is higher than the 95th percentile of the first period, or the 95th percentile of the second period is lower than the 5th percentile of the first period, the change in performance of this independent combination from the first to the second period is significant. Finally, grids in the map of temporal changes in performance for each dataset were masked by n/a (i.e., not available) if they did not have over 50% independent combinations showing

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

significant changes.

3 Results

3.1 Water balance consistency of considered datasets

Fig. 1b-e summarizes the performance of considered datasets in terms of their water balance consistency, based on monthly calculations (see Methods). Colors distinguish gauge-based, satellite-based and reanalysis datasets. Overall, the R^2 scores are fairly low, indicating prevailing inconsistencies across considered datasets in terms of the water balance. From the combinations with top ten performance, it is likely that the *P* from PERSIANN-CDR, *ET* from PT-JPL, *R* from GRUN, and *SM* from GLDAS-2.1 would contribute to high water balance consistency (Fig. S13).

For *P* datasets, the overall performance of satellite-based datasets is generally higher than gauge-based and reanalysis datasets, where the CHIPRS v2.0 and PERSIANN-CDR show the highest global medians (Fig. 1b). This is related to their limited spatial coverage omitting high-latitude regions with typically low water balance consistency (Fig. S3), while for 50ºS−50ºN, PERSIANN-CDR, GPM IMERG v07 and MSWEP v2.8 show the highest medians (Fig. 1b). Besides, GPM IMERG v07 and MSWEP v2.8 exhibit the largest areas with the best performance across datasets. Fig. 2 maps the types of datasets with the highest water balance consistency for each considered variable. It shows that given the comparatively good performance of GPM IMERG v07 and MSWEP v2.8, satellite-based precipitation dataset types perform best across most of the globe, particularly in the tropics and subtropics (Fig. S14 and Fig. 2a). Gauge-based *P* datasets perform best in high-latitude regions which in the Northern Hemisphere are characterized by abundant *in-situ* observations (Fig. S7).

ET and *R* datasets show similar global patterns and medians of overall performance among the different dataset types (Figs. S4−S5 and Fig. 1cd). However, for the spatial patterns, PT-JPL and GLDAS-2.2 have distinctly larger areas with the best performance compared to other *ET* datasets (Fig. S14b), leading to comparable best-performance areas between satellite-based and reanalysis *ET* datasets (Fig. 2b). Similarly, gauge-based and reanalysis *R* datasets show the largest areas with the best performance (Fig. 2c), where GRUN and ERA5-land datasets are the respective main contributors (Fig. S14c).

Among *SM* datasets, SoMo.ml and ESA CCI v08.1 have the lowest global medians of overall performance. This is because they only represent the surface layers instead of the entire soil column (Fig. 1e). Meanwhile, the *SM* datasets with simulations of deep soil layer generally performed better in most global regions, such as the reanalysis and GLDAS-2 products (Fig. 2d and Fig. S14d).

Additionally, we calculated our analysis at daily and annual time scales. Results indicate substantially less water balance consistency with the lowest R^2 scores at the annual scale (Fig. 1b−e and Fig. S15). However, different temporal resolutions did not alter the relative ranking patterns among the datasets (Fig. 1b−d), except for *SM* whose memory is likely to be more sensitive to the varying resolutions (Fig. 1e).

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

Open Access Earth System Science Data Discussions

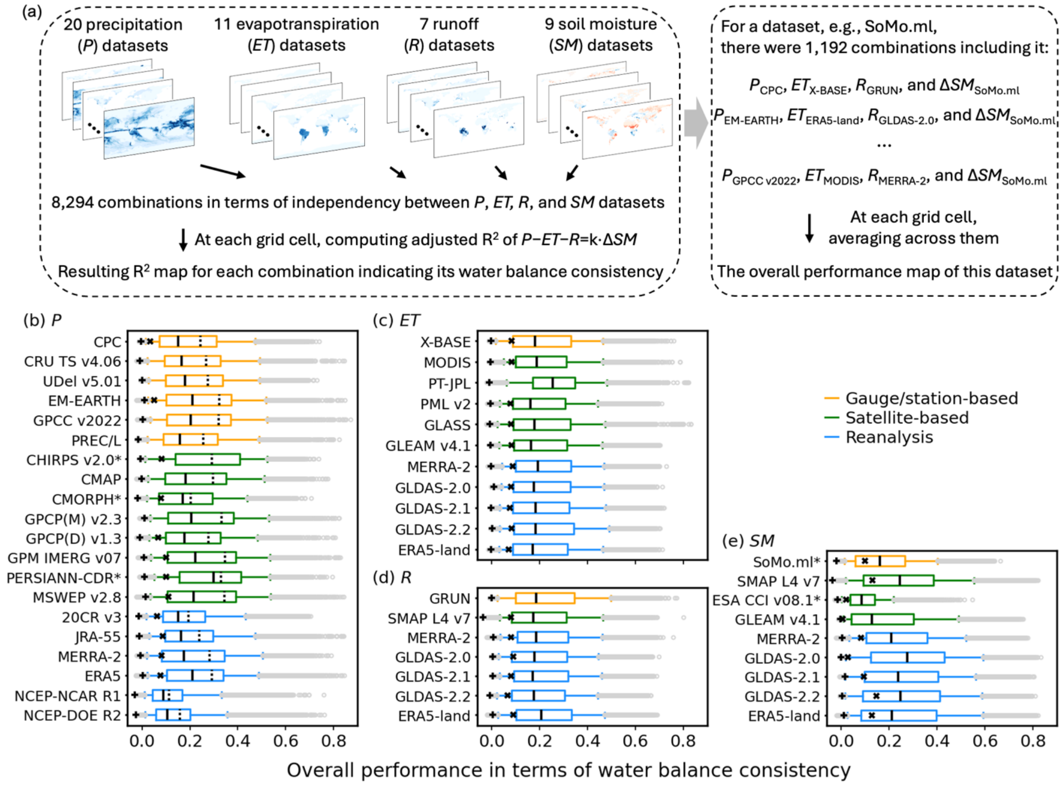

Fig. 1. Illustration of water-balance approach and calculated performance of considered datasets. (a) Performance is determined based on R^2 scores measuring consistency of each dataset when combined with all independent datasets in terms of the water balance (Methods). (b–e) The boxplots summarize the performance of considered datasets. Colors indicate the type of each dataset. Each box shows the median value, as well as the 5^{th}, 25^{th}, 75^{th}, and 95^{th} percentiles of the global pattern of water balance consistency derived from monthly data. Median results for performing the analysis with daily and annual data are indicated through crosses and pluses, respectively (Text S1−S2). Asterisks (*) following the name of *P* dataset indicate its limited spatial coverage omitting high-latitude regions with typically low performance, and dashed line in each box indicates median of only 50ºS−50ºN. * of *SM* dataset indicates that the dataset does not consider the entire soil column.

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

Open Access Earth System Science Data Discussions

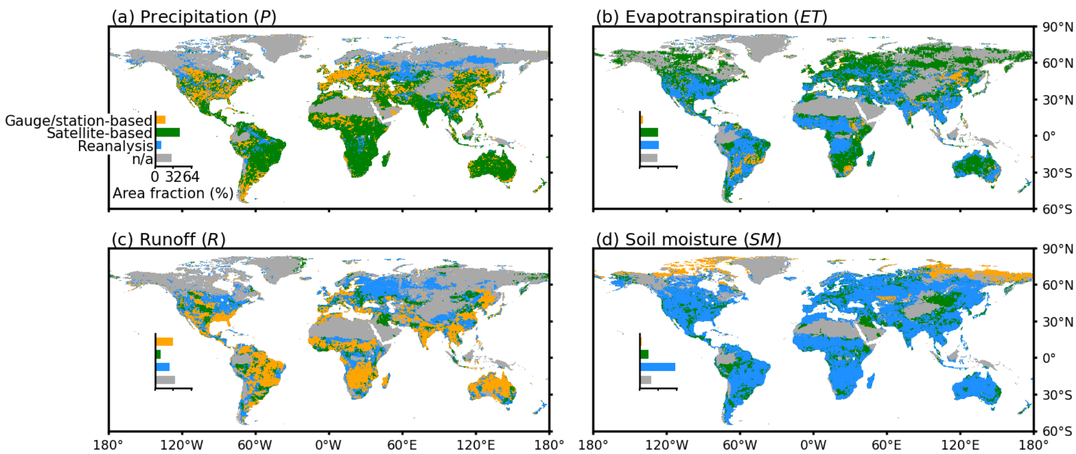

Fig. 2. Types of best-performing datasets across hydrological variables. Colors indicate type of dataset with the highest water balance consistency. Gray color indicates that multiple datasets show similar water balance consistency (with R^2 scores varied within 5%) or low water balance consistency (with all R^2 scores lower than 0.2).

3.2 Potential reasons influencing water balance consistency

Next, we aim to diagnose possible reasons for regional discrepancies of dataset performance in terms of water balance consistency. For this purpose, we consider a large set of variables that may affect the water balance consistency of a given dataset, including soil and vegetation characteristics, climate, and gauge density (Methods). By applying an explainable machine learning method (i.e., SHAP), temperature and aridity (i.e., *ET*/*P*) were diagnosed as the key factors to influence the spatial performance patterns of *P*, *ET*, and *R* datasets, while for *SM* datasets temperature and tree cover are critical (Fig. 3 and Figs. S16−S19). Our results demonstrate that the performance of *P* datasets is higher in the sub-humid and sub-arid regions (where the aridity index is 0.6−1.0) with monthly mean temperatures between 10ºC and 15ºC (Fig. 3a and Fig. S20). The results for *ET* and *R* datasets are largely similar to those of *P* datasets (Fig. 3b−c and Figs. S21−S22), while comparatively good performance of *SM* datasets is found in regions with a moderate tree cover fraction (5−50 %) and warm temperature (10−15 ºC) (Fig. 3d and Fig. S23). These influence patterns were summarized according to medians across dataset performance, while using maximum does not alter the results (Fig. S24).

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

Open Access Earth System Science Data Discussions

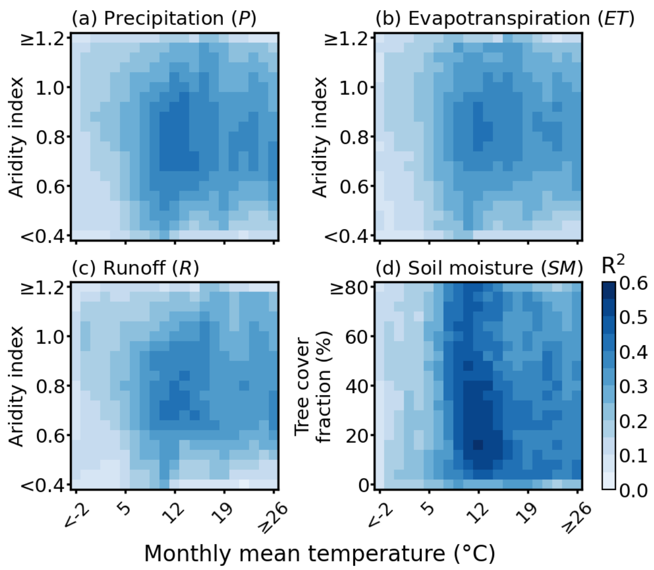

Fig. 3. Influence of temperature and aridity (or tree cover for *SM*) on water balance consistency of datasets. The consistency through water balance is quantified by R^2 scores (Fig. 1), and median R^2 scores across *P*/*ET*/*R*/*SM* datasets for each climate/vegetation class (Figs. S20−S23) are shown.

3.3 Temporal changes in water balance consistency of dataset

We furthermore assess changes in the diagnosed dataset performance inferred from water balance consistency over time. This is done by splitting our study period and repeating the analysis for the sub-periods 2000−2010 and 2011−2022, and includes an assessment of significance (Methods). For the *P* datasets, the majority of global grid cells show no temporal change in water balance consistency, and among the grid cells with temporal changes, we found mostly increases (Fig. 4a). These increasing changes were mainly observed in middle- and high-latitude regions of the Northern Hemisphere, while the *P* dataset from ERA5 shows the highest median level of performance improvement (Fig. S25). At the same time, we find similar spatial patterns of changes in water balance consistency for *ET*, *R*, and *S* datasets, with most grid cells showing no change (Fig. 4b−d). Among the grid cells with significant changes, performance in terms of water balance consistency increases prevail and are mostly located in high-latitude regions and in regions with scarce observations in the Northern Hemisphere (Fig. 4b−d, Fig. S14 and S26−S28).

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

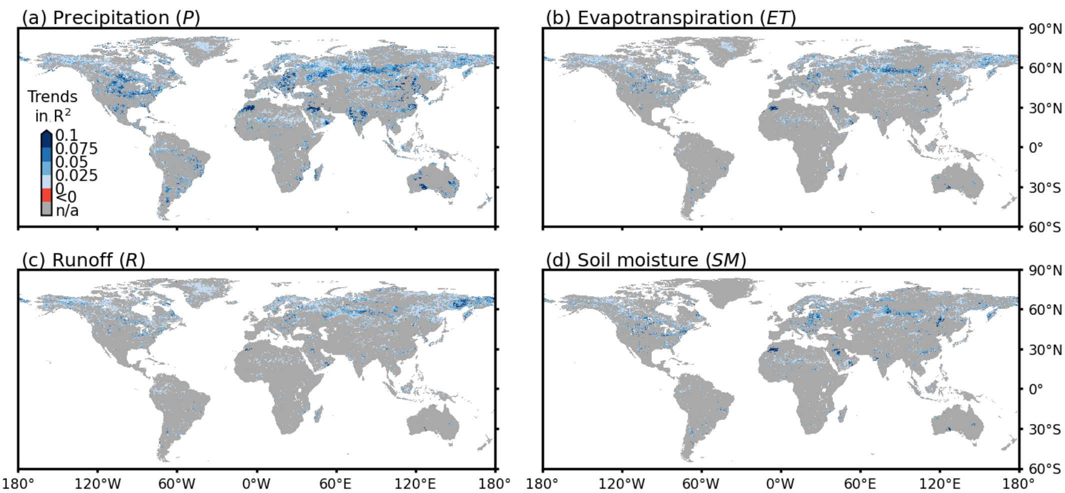

Fig. 4. Temporal changes in water balance consistency of *P*, *ET*, *R*, and *SM* datasets from 2000−2010 to 2011−2022. Based on the changes in R scores for each dataset (Figs. S25−S28), median values are shown in each grid cell where at least half the considered datasets showed significant changes (Methods), representing common temporally changing patterns.

4 Discussion

The spatial performance patterns derived from our water balance consistency approach reveal high similarity among *P* datasets (Fig. S3), consistent with findings from recent studies on *P* dataset agreement (Markonis et al., 2024; Dosio et al., 2021). Beyond these similarities, our grid cell-level comparisons suggest that satellite-based *P* datasets outperform others in large regions of southern America, Africa, south and Southeast Asia, and inner Australia, while gauge-based *P* datasets excel in many grid cells across the United States, Europe, and East Asia (Fig. 2). This suggests that the satellite-based *P* datasets are superior in regions with sparse or no gauging stations (Fig. S7), compared to gauge-based and reanalysis datasets. However, all *P* datasets exhibit higher water balance consistency in moderately humid or dry regions, with long-term mean temperature also influencing the performance (Fig. 3 and Fig. S20). Lower consistency of gauge-based datasets in humid and dry regions may stem from challenges in mapping spatial variability of extreme rainfall (Mishra and Coulibaly, 2009) and accurately recording light precipitation events (Lanza et al., 2022), as consistency is based on seasonal variabilities in water balance. Additionally, *P* datasets show lower consistency in cold regions because of difficulties in measuring solid precipitation (La et al., 2002). Similarly, satellite-derived precipitation is relatively insensitive to light rainfall (Laviola et al., 2013), struggles with extreme rainfall estimates (likely due to retrieval algorithms and infrequent temporal sampling of polar orbits) (Barlow et al., 2019), and often fails to detect snowfall or perform well over snow- and ice-covered surfaces (Alijanian et al., 2017). In contrast, reanalysis datasets perform better in cold regions, benefiting from assimilated meteorological observations and atmospheric states (Barlow et al., 2019; Dosio et al., 2021; Sun et al., 2018).

The *ET*, *R*, and *SM* datasets generally show global spatial performance patterns similar to those of *P* datasets (Figs. S4−S6). This is partly because uncertainties in *P* datasets propagate through the water cycle (Fallah et al., 2020), affecting the water balance consistency of *ET*, *R*, and *SM* datasets. Nevertheless, our approach identifies distinct relative performances across hydrological variables and dataset types (Fig. 2 and Fig. S14), as it considers independent combinations of datasets. For *ET*, the satellite-based PT-JPL dataset performs comparatively well, likely due to its advanced consideration of plant physiological limitations and water stress. The reanalysis dataset GLDAS-2.2 also performs comparatively well, probably due to its assimilation of terrestrial water storage (Table S2 and Fig. S14). For *R*, the machine-learning model-driven GRUN, constrained by *P* and temperature in large basins, and ERA5-land dataset, perform best in most regions (Tables S3 and Fig. S14). For *SM*, reanalysis datasets perform best, likely because they are constrained by physical laws and considers deeper soil moisture variability (Table S4 and Fig. S14). Overall, our results highlight the importance of physical constraints and of data assimilation in enhancing water balance consistency of hydrological variables (Pan et al., 2020; Tang et al., 2024; Yang et al., 2023; Ruiz-Vásquez et al., 2023).

Dataset performance varied significantly across time scales, with the highest correspondence at the monthly scale, where seasonal variability is well-captured and synoptic weather variability is mitigated. This explains the extremely lower water balance consistency observed at the annual scale for all datasets. At a daily time scale, the variability of the involved variables is high, including more extreme values, and apparently under-constrained by available observations (Maurer and Hidalgo, 2008; Fisher et al., 2008). Furthermore, we find widespread increases in water balance consistency across hydrological variables during our study period in mid-to-high latitude regions of the Northern Hemisphere (Fig. 4). These regions have experienced reduced snow-cover durations (Bormann et al., 2018) and extents (Mudryk et al., 2020), as well as less snowfall (O'gorman, 2014), which has weakened *R* seasonality (Wang et al., 2024) and enhanced the influence of *P* variability on *R* seasonality (Han et al., 2024). Given the influence patterns in Fig. 3, higher temperatures and reduced solid precipitation likely enhance *P* dataset performance. Also, the absence of strong increases in extreme precipitation events in these regions (Asadieh and Krakauer, 2015) may contribute to improved consistency. Previous studies have shown that models incorporating updated vegetation information, such as leaf area index (LAI) seasonality, perform better in these regions (Ruiz-Vásquez et al., 2023; Nogueira et al., 2021), aligning with our observed improvements over time (Fig. 4). This underscores the importance of accurately representing the coupling between *ET* and *SM* for dataset performance, as inferred from our approach.

Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the online repository. Additionally, their access links are provided in the following. CPC is available at https://www.psl.noaa.gov/data/gridded/data.cpc.globalprecip.html; CRU TS v4.06 is available at https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.06/; UDel v5.01 is available at https://climate.geog.udel.edu/; EM-EARTH is available at https://www.frdr-dfdr.ca/repo/dataset/8d30ab02-f2bd-4d05-ae43-11f4a387e5ad; GPCC v2022 is available at https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-monthly_v2022_doi_download.html; PREC/L is available at https://psl.noaa.gov/data/gridded/data.precl.html; CHIRPS v2.0 is available at https://www.chc.ucsb.edu/data/chirps; CMAP is available at https://psl.noaa.gov/data/gridded/data.cmap.html; CMORPH v1 is available at https://www.ncei.noaa.gov/products/climate-data-records/precipitation-cmorph; GPCP(M) v2.3 is available at https://psl.noaa.gov/data/gridded/data.gpcp.html; GPCP(D) v1.3 is available at https://rda.ucar.edu/datasets/d728007/; GPM IMERG v07 is available at https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_07/summary?keywords=%22IMERG%20final%22; PERSIANN-CDR is available at https://www.ncei.noaa.gov/products/climate-data-records/precipitation-persiann; MSWEP v2.8 is available at https://www.gloh2o.org/mswep/; 20CR v3 is available at https://psl.noaa.gov/data/gridded/data.20thC_ReanV3.html; JRA-55 is available at https://rda.ucar.edu/datasets/d628000/; ERA5 is available at https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview; NCEP-NCAR R1 is available at https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html; NCEP-DOE R2 is available at https://psl.noaa.gov/data/gridded/data.ncep.reanalysis2.html; MERRA-2 is available at https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/data_access/;X-BASE is available at https://meta.icos-cp.eu/collections/_l85vWiIV81AifoxCkty50YI; MODIS is available at https://lpdaac.usgs.gov/products/mod16a2gfv061/; PT-JPL is available at http://josh.yosh.org/; PML-v2 is available at https://doi.org/10.5281/zenodo.10647618 (Zhang et al., 2024b); GLASS is available at http://www.glass.umd.edu/Download.html; GLEAM v4.1 is available at https://www.gleam.eu/; GLDAS-2.0/2.1/2.2 are available at https://disc.gsfc.nasa.gov/datasets?keywords=GLDAS; ERA5-land is available at https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview; GRUN is available at https://figshare.com/articles/dataset/GRUN_Global_Runoff_Reconstruction/9228176; SMAP L4 v7 is available at https://nsidc.org/data/spl4smgp/versions/7; SoMo.ml is available at https://www.bgc-jena.mpg.de/geodb/projects/Data.php; ESA CCI v08.1 is available at https://climate.esa.int/en/projects/soil-moisture/.

Code availability

The core codes for calculating the water balance consistency of each combination and each dataset, as well as assessing the potential influence based on explainable machine learning and uncertainties of the temporal changes based on bootstrap confidence intervals, are available at https://github.com/HowHuang/WaterBalanceConsistency.

Author contributions

R.O. conceived the original idea, which was further developed in collaboration with H.H. and J.L. H.H. aggregated the datasets used in this study, did the analysis, and prepared the original paper. H.H., J.L., A.C., M.R.V., and R.O. contributed to interpreting the results and discussion and improving the paper.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

We thank the respective dataset development groups for their continuous efforts in advancing the

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

state-of-the-art datasets and ensuring easy access. We furthermore thank Sophia Walther (Max Planck Institute for Biogeochemistry, Jena), as well as the Biogeochemical System Modeling, Biometry and Environmental System Analysis (led by Carsten F. Dormann), and Sensor-based Geoinformatics (led by Teja Kattenborn) groups at the University of Freiburg, for fruitful discussions.

Competing interests

The authors declare no competing interests.

Financial support

J.L. acknowledges support from the National Natural Science Foundation of China (Grant No. 42361144001), the 111 Project (Grant No. D25014), the National Foreign Experts Program (Category S) (Grant No. S20240116), and the Henan Province Foreign Scientist Studio for Synergistic Management of Water, Food, Energy, and Carbon (Grant No. GZS2024013). A.C. acknowledges support from Guangdong Provincial Key Laboratory of Intelligent Disaster Prevention and Emergency Technologies for Urban Lifeline Engineering (2022) (Grant No. 2022B1212010016). H.H. acknowledges a scholarship from the China Scholarship Council for visiting the University of Freiburg to complete this study and work with many cool scientists. R.O. acknowledges support from the German Research Foundation (Emmy Noether Grant No. 391059971).

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

References

Abolafia-Rosenzweig, R., Pan, M., Zeng, J. L., and Livneh, B.: Remotely sensed ensembles of the terrestrial water budget over major global river basins: An assessment of three closure techniques, Remote Sensing of Environment, 252, 10.1016/j.rse.2020.112191, 2021.

Adler, R. F., Sapiano, M., Huffman, G. J., Wang, J., Gu, G., Bolvin, D., Chiu, L., Schneider, U., Becker, A., Nelkin, E., Xie, P., Ferraro, R., and Shin, D. B.: The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation, Atmosphere (Basel), 9, 10.3390/atmos9040138, 2018.

Alijanian, M., Rakhshandehroo, G. R., Mishra, A. K., and Dehghani, M.: Evaluation of satellite rainfall climatology using CMORPH, PERSIANN-CDR, PERSIANN, TRMM, MSWEP over Iran, International Journal of Climatology, 37, 4896-4914, 10.1002/joc.5131, 2017.

Asadieh, B. and Krakauer, N. Y.: Global trends in extreme precipitation: climate models versus observations, Hydrology and Earth System Sciences, 19, 877-891, 10.5194/hess-19-877-2015, 2015.

Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P.: PERSIANN-CDR Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies, Bulletin of the American Meteorological Society, 96, 69-+, 10.1175/Bams-D-13-00068.1, 2015.

Barlow, M., Gutowski, W. J., Gyakum, J. R., Katz, R. W., Lim, Y.-K., Schumacher, R. S., Wehner, M. F., Agel, L., Bosilovich, M., Collow, A., Gershunov, A., Grotjahn, R., Leung, R., Milrad, S., and Min, S.-K.: North American extreme precipitation events and related large-scale meteorological patterns: a review of statistical methods, dynamics, modeling, and trends, Climate Dynamics, 53, 6835-6875, 10.1007/s00382-019-04958-z, 2019.

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., McVicar, T. R., and Adler, R. F.: MSWEP V2 Global 3-Hourly 0.1 degrees Precipitation: Methodology and Quantitative Assessment, Bulletin of the American Meteorological Society, 100, 473-502, 10.1175/Bams-D-17-0138.1, 2019.

Bormann, K. J., Brown, R. D., Derksen, C., and Painter, T. H.: Estimating snow-cover trends from space, Nature Climate Change, 8, 923-927, 10.1038/s41558-018-0318-3, 2018.

Chen, M. Y., Xie, P. P., Janowiak, J. E., and Arkin, P. A.: Global land precipitation: A 50-yr monthly analysis based on gauge observations, Journal of Hydrometeorology, 3, 249-266, Doi 10.1175/1525-7541(2002)003<0249:Glpaym>2.0.Co;2, 2002.

Cleophas, F., Isidore, F., Musta, B., Mohd Ali, B. N., Mahali, M., Zahari, N. Z., and Bidin, K.: Effect of soil physical properties on soil infiltration rates, Journal of Physics: Conference Series, 2314, 10.1088/1742-6596/2314/1/012020, 2022.

Cooley, S. S., Fisher, J. B., and Goldsmith, G. R.: Convergence in water use efficiency within plant functional types across contrasting climates, Nat Plants, 8, 341-345, 10.1038/s41477-022-01131-z, 2022.

Denissen, J. M. C., Teuling, A. J., Pitman, A. J., Koirala, S., Migliavacca, M., Li, W. T., Reichstein, M., Winkler, A. J., Zhan, C. H., and Orth, R.: Widespread shift from ecosystem energy to water limitation with climate change, Nature Climate Change, 12, 677-684, 10.1038/s41558-022-01403-8, 2022.

Do, H. X., Gudmundsson, L., Leonard, M., and Westra, S.: The Global Streamflow Indices and Metadata Archive (GSIM) - Part 1: The production of a daily streamflow archive and metadata, Earth System Science Data, 10, 765-785, 10.5194/essd-10-765-2018, 2018.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch,

M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrology and Earth System Sciences, 15, 1675-1698, 10.5194/hess-15-1675-2011, 2011.
Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., and Lautenbach, S.: Collinearity: a review of methods to deal with it and a simulation study evaluating their performance, Ecography, 36, 27-46, 10.1111/j.1600-0587.2012.07348.x, 2012.
Dosio, A., Pinto, I., Lennard, C., Sylla, M. B., Jack, C., and Nikulin, G.: What can we know about recent past precipitation over Africa? Daily characteristics of African precipitation from a large ensemble of observational products for model evaluation, Earth and Space Science, 8, 10.1029/2020EA001466, 2021.
Douville, H., Raghavan, K., Renwick, J., Allan, R. P., Arias, P. A., Barlow, M., Cerezo-Mota, R., Cherchi, A., Gan, T., and Gergis, J.: Water cycle changes, 2021.
Fallah, A., O, S., and Orth, R.: Climate-dependent propagation of precipitation uncertainty into the water cycle, Hydrology and Earth System Sciences, 24, 3725-3735, 10.5194/hess-24-3725-2020, 2020.
Fisher, J. B., Tu, K. P., and Baldocchi, D. D.: Global estimates of the land-atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites, Remote Sensing of Environment, 112, 901-919, 10.1016/j.rse.2007.06.025, 2008.
Freedman, F. R., Pitts, K. L., and Bridger, A. F. C.: Evaluation of CMIP climate model hydrological output for the Mississippi River Basin using GRACE satellite observations, Journal of Hydrology, 519, 3566-3577, 10.1016/j.jhydrol.2014.10.036, 2014.
Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J.: The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes, Scientific Data, 2, 10.1038/sdata.2015.66, 2015.
Gebrechorkos, S. H., Leyland, J., Dadson, S. J., Cohen, S., Slater, L., Wortmann, M., Ashworth, P. J., Bennett, G. L., Boothroyd, R., Cloke, H., Delorme, P., Griffith, H., Hardy, R., Hawker, L., McLelland, S., Neal, J., Nicholas, A., Tatem, A. J., Vahidi, E., Liu, Y. X., Sheffield, J., Parsons, D. R., and Darby, S. E.: Global-scale evaluation of precipitation datasets for hydrological modelling, Hydrology and Earth System Sciences, 28, 3099-3118, 10.5194/hess-28-3099-2024, 2024.
Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Journal of Climate, Volume 30, 5419-5454, 10.1175/JCLI-D-16-0758.1, 2017.
Ghiggi, G., Humphrey, V., Seneviratne, S. I., and Gudmundsson, L.: GRUN: an observation-based global gridded runoff dataset from 1902 to 2014, Earth System Science Data, 11, 1655-1674, 10.5194/essd-11-1655-2019, 2019.
Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., and Seneviratne, S. I.: Global assessment of trends in wetting and drying over land, Nature Geoscience, 7, 716-721, 10.1038/ngeo2247, 2014.

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., and Dorigo, W.: Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology, Earth System Science Data, 11, 717-739, 10.5194/essd-11-717-2019, 2019.

Han, J., Liu, Z., Woods, R., McVicar, T. R., Yang, D., Wang, T., Hou, Y., Guo, Y., Li, C., and Yang, Y.: Streamflow seasonality in a snow-dwindling world, Nature, 629, 1075-1081, 10.1038/s41586-024-07299-y, 2024.

Hansen, M. and Song, X.: Vegetation Continuous Fields (VCF) Yearly Global 0.05 Deg, https://doi.org/10.5067/MEaSUREs/VCF/VCF5KYR.001, 2018.

Harris, I., Osborn, T. J., Jones, P., and Lister, D.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, Scientific Data, 7, 109, 10.1038/s41597-020-0453-3, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Q J Roy Meteor Soc, 146, 1999-2049, 10.1002/qj.3803, 2020.

Hirschi, M., Stradiotti, P., Crezee, B., Dorigo, W., and Seneviratne, S. I.: Potential of long-term satellite observations and reanalysis products for characterising soil drying: trends and drought events, Hydrology and Earth System Sciences, 29, 397-425, 10.5194/hess-29-397-2025, 2025.

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., and Tan, J.: GPM IMERG Final Precipitation L3 1 day 0.1 degree x 0.1 degree V07, 10.5067/GPM/IMERGDF/DAY/07, 2023.

Huffman, G. J., Adler, R. F., Morrissey, M. M., Bolvin, D. T., Curtis, S., Joyce, R., McGavock, B., and Susskind, J.: Global precipitation at one-degree daily resolution from multisatellite observations, Journal of Hydrometeorology, 2, 36-50, Doi 10.1175/1525-7541(2001)002<0036:Gpaodd>2.0.Co;2, 2001.

Japan Meteorological Agency, J.: JRA-55: Japanese 55-year Reanalysis, Monthly Means and Variances, https://doi.org/10.5065/D60G3H5B, 2013.

Jaramillo, F. and Destouni, G.: Local flow regulation and irrigation raise global human water consumption and footprint, Science, 350, 1248-1251, 10.1126/science.aad1010, 2015.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S. K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP-DOE AMIP-II reanalysis (R-2), Bulletin of the American Meteorological Society, 83, 1631-1643, 10.1175/Bams-83-11-1631(2002)083<1631:Nar>2.3.Co;2, 2002.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, Bulletin of the American Meteorological Society, 82, 247-267, 10.1175/1520-0477(2001)082<0247:Tnnyrm>2.3.Co;2, 2001.

Kulesa, A., Krzywinski, M., Blainey, P., and Altman, N.: Sampling distributions and the bootstrap, Nat Methods, 12, 477-478, 10.1038/nmeth.3414, 2015.

La, B. P., Lanza, L. G., and Stagi, L.: Tipping bucket mechanical errors and their influence on rainfall statistics and extremes, Water Sci Technol, 45, 1-10, 2002.

Lanza, L. G., Cauteruccio, A., and Stagnaro, M.: Rain gauge measurements, in: Rainfall,

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

Elsevier, 77-108, 2022.
Laviola, S., Levizzani, V., Cattani, E., and Kidd, C.: The 183-WSL fast rain rate retrieval algorithm. Part II: Validation using ground radar measurements, Atmos Res, 134, 77-86, 10.1016/j.atmosres.2013.07.013, 2013.
Legates, D. R. and Willmott, C. J.: Mean seasonal and spatial variability in gauge-corrected, global precipitation, International Journal of Climatology, 10, 111-127, 10.1002/joc.3370100202, 1990.
Li, B., Rodell, M., Sheffield, J., Wood, E., and Sutanudjaja, E.: Long-term, non-anthropogenic groundwater storage changes simulated by three global-scale hydrological models, Sci Rep, 9, 10746, 10.1038/s41598-019-47219-z, 2019.
Li, W., Migliavacca, M., Forkel, M., Denissen, J. M. C., Reichstein, M., Yang, H., Duveiller, G., Weber, U., and Orth, R.: Widespread increasing vegetation sensitivity to soil moisture, Nature Communications, 13, 3959, 10.1038/s41467-022-31667-9, 2022.
Li, W., Reichstein, M., O, S., May, C., Destouni, G., Migliavacca, M., Kraft, B., Weber, U., and Orth, R.: Contrasting Drought Propagation Into the Terrestrial Water Cycle Between Dry and Wet Regions, Earth's Future, 11, e2022EF003441, 10.1029/2022ef003441, 2023a.
Li, W., Pacheco-Labrador, J., Migliavacca, M., Miralles, D., Hoek van Dijke, A., Reichstein, M., Forkel, M., Zhang, W., Frankenberg, C., Panwar, A., Zhang, Q., Weber, U., Gentine, P., and Orth, R.: Widespread and complex drought effects on vegetation physiology inferred from space, Nat Commun, 14, 4640, 10.1038/s41467-023-40226-9, 2023b.
Markonis, Y., Godoy, M. R. V., Pradhan, R. K., Pratap, S., Thomson, J. R., Hanel, M., Paschalis, A., Nikolopoulos, E., and Papalexiou, S. M.: Spatial partitioning of terrestrial precipitation reveals varying dataset agreement across different environments, Communications Earth & Environment, 5, 10.1038/s43247-024-01377-9, 2024.
Maurer, E. P. and Hidalgo, H. G.: Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods, Hydrology and Earth System Sciences, 12, 551-563, 2008.
McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N. E. C., Franz, T. E., Shi, J., Gao, H., and Wood, E. F.: The Future of Earth Observation in Hydrology, Hydrol Earth Syst Sci, 21, 3879-3914, 10.5194/hess-21-3879-2017, 2017.
Mehta, P., Siebert, S., Kummu, M., Deng, Q., Ali, T., Marston, L., Xie, W., and Davis, K. F.: Half of twenty-first century global irrigation expansion has been in water-stressed regions, Nature Water, 2, 254-261, 10.1038/s44221-024-00206-9, 2024.
Mekonnen, M. M. and Hoekstra, A. Y.: Four billion people facing severe water scarcity, Science Advance, 2, e1500323, 10.1126/sciadv.1500323, 2016.
Miralles, D. G., Bonte, O., Koppa, A., Baez-Villanueva, O. M., Tronquo, E., Zhong, F., Beck, H. E., Hulsman, P., Dorigo, W., Verhoest, N. E. C., and Haghdoost, S.: GLEAM4: global land evaporation and soil moisture dataset at 0.1 resolution from 1980 to near present, Sci Data, 12, 416, 10.1038/s41597-025-04610-y, 2025.
Mishra, A. K. and Coulibaly, P.: Developments in hydrometric network design: A review, Reviews of Geophysics, 47, 10.1029/2007rg000243, 2009.
Mudryk, L., Santolaria-Otín, M., Krinner, G., Ménégoz, M., Derksen, C., Brutel-Vuilmet, C., Brady, M., and Essery, R.: Historical Northern Hemisphere snow cover trends and projected changes in the CMIP6 multi-model ensemble, Cryosphere, 14, 2495-2514, 10.5194/tc-14-2495-2020, 2020.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 4349-4383, 10.5194/essd-13-4349-2021, 2021.
Nachtergaele, F., van Velthuizen, H., Verelst, L., Wiberg, D., Henry, M., Chiozza, F., Yigini, Y., Aksoy, E., Batjes, N., and Boateng, E.: Harmonized world soil database version 2.0, FAO2023.
Nelson, J. A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N., Migliavacca, M., Wohlfahrt, G., Šigut, L., Ibrom, A., Papale, D., Göckede, M., Duveiller, G., Knohl, A., Hörtnagl, L., Scott, R. L., Zhang, W., Hamdi, Z. M., Reichstein, M., Aranda-Barranco, S., Ardö, J., Op de Beeck, M., Billesbach, D., Bowling, D., Bracho, R., Brümmer, C., Camps-Valls, G., Chen, S., Cleverly, J. R., Desai, A., Dong, G., El-Madany, T. S., Euskirchen, E. S., Feigenwinter, I., Galvagno, M., Gerosa, G. A., Gielen, B., Goded, I., Goslee, S., Gough, C. M., Heinesch, B., Ichii, K., Jackowicz-Korczynski, M. A., Klosterhalfen, A., Knox, S., Kobayashi, H., Kohonen, K.-M., Korkiakoski, M., Mammarella, I., Gharun, M., Marzuoli, R., Matamala, R., Metzger, S., Montagnani, L., Nicolini, G., O'Halloran, T., Ourcival, J.-M., Peichl, M., Pendall, E., Ruiz Reverter, B., Roland, M., Sabbatini, S., Sachs, T., Schmidt, M., Schwalm, C. R., Shekhar, A., Silberstein, R., Silveira, M. L., Spano, D., Tagesson, T., Tramontana, G., Trotta, C., Turco, F., Vesala, T., Vincke, C., Vitale, D., Vivoni, E. R., Wang, Y., Woodgate, W., Yepez, E. A., Zhang, J., Zona, D., and Jung, M.: X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, FLUXCOM-X, Biogeosciences, 21, 5079-5115, 10.5194/bg-21-5079-2024, 2024.
NOAA National Centers for Environmental Information: ETOPO 2022 15 Arc-Second Global Relief Model, 10.25921/fd45-gt74, 2022.
Nogueira, M., Boussetta, S., Balsamo, G., Albergel, C., Trigo, I. F., Johannsen, F., Miralles, D. G., and Dutra, E.: Upgrading Land-Cover and Vegetation Seasonality in the ECMWF Coupled System: Verification With FLUXNET Sites, METEOSAT Satellite Land Surface Temperatures, and ERA5 Atmospheric Reanalysis, J Geophys Res Atmos, 126, e2020JD034163, 10.1029/2020JD034163, 2021.
O, S. and Orth, R.: Global soil moisture data derived through machine learning trained with in-situ measurements, Scientific Data, 8, 170, 10.1038/s41597-021-00964-1, 2021.
O'Gorman, P. A.: Contrasting responses of mean and extreme snowfall to climate change, Nature, 512, 416-418, 10.1038/nature13625, 2014.
Oki, T. and Kanae, S.: Global hydrological cycles and world water resources, Science, 313, 1068-1072, 10.1126/science.1128845, 2006.
Pan, S. F., Pan, N. Q., Tian, H. Q., Friedlingstein, P., Sitch, S., Shi, H., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Lienert, S., Lombardozzi, D., Nabel, J. E. M. S., Otté, C., Poulter, B., Zaehle, S., and Running, S. W.: Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling, Hydrology and Earth System Sciences, 24, 1485-1509, 10.5194/hess-24-1485-2020, 2020.
Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardo, J., Arkebauer, T., Arndt, S. K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A.,

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

Beamesderfer, E., Marchesini, L. B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Boike, J., Bolstad, P. V., Bonal, D., Bonnefond, J. M., Bowling, D. R., Bracho, R., Brodeur, J., Brummer, C., Buchmann, N., Burban, B., Burns, S. P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T. R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B. D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P. S., D'Andrea, E., da Rocha, H., Dai, X., Davis, K. J., Cinti, B., Grandcourt, A., Ligne, A., De Oliveira, R. C., Delpierre, N., Desai, A. R., Di Bella, C. M., Tommasi, P. D., Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrene, E., Dunn, A., Dusek, J., Eamus, D., Eichelmann, U., ElKhidir, H. A. M., Eugster, W., Ewenz, C. M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., Gharun, M., Gianelle, D., Gielen, B., Gioli, B., Gitelson, A., Goded, I., Goeckede, M., Goldstein, A. H., Gough, C. M., Goulden, M. L., Graf, A., Griebel, A., Gruening, C., Grunwald, T., Hammerle, A., Han, S., Han, X., Hansen, B. U., Hanson, C., Hatakka, J., He, Y., Hehn, M., Heinesch, B., Hinko-Najera, N., Hortnagl, L., Hutley, L., Ibrom, A., Ikawa, H., Jackowicz-Korczynski, M., Janous, D., Jans, W., Jassal, R., Jiang, S., Kato, T., Khomik, M., Klatt, J., Knohl, A., Knox, S., Kobayashi, H., Koerber, G., Kolle, O., Kosugi, Y., Kotani, A., Kowalski, A., Kruijt, B., Kurbatova, J., Kutsch, W. L., Kwon, H., Launiainen, S., Laurila, T., Law, B., Leuning, R., Li, Y., Liddell, M., Limousin, J. M., Lion, M., Liska, A. J., Lohila, A., Lopez-Ballesteros, A., Lopez-Blanco, E., Loubet, B., Loustau, D., Lucas-Moffat, A., Luers, J., Ma, S., Macfarlane, C., Magliulo, V., Maier, R., Mammarella, I., Manca, G., Marcolla, B., Margolis, H. A., Marras, S., Massman, W., Mastepanov, M., Matamala, R., Matthes, J. H., Mazzenga, F., McCaughey, H., McHugh, I., McMillan, A. M. S., Merbold, L., Meyer, W., Meyers, T., Miller, S. D., Minerbi, S., Moderow, U., Monson, R. K., Montagnani, L., Moore, C. E., Moors, E., Moreaux, V., Moureaux, C., Munger, J. W., Nakai, T., Neirynck, J., Nesic, Z., Nicolini, G., Noormets, A., Northwood, M., Nosetto, M., Nouvellon, Y., Novick, K., Oechel, W., Olesen, J. E., Ourcival, J. M., Papuga, S. A., Parmentier, F. J., Paul-Limoges, E., Pavelka, M., Peichl, M., Pendall, E., Phillips, R. P., Pilegaard, K., Pirk, N., Posse, G., Powell, T., Prasse, H., Prober, S. M., Rambal, S., Rannik, U., Raz-Yaseef, N., Rebmann, C., Reed, D., Dios, V. R., Restrepo-Coupe, N., Reverter, B. R., Roland, M., Sabbatini, S., Sachs, T., Saleska, S. R., Sanchez-Canete, E. P., Sanchez-Mejia, Z. M., Schmid, H. P., Schmidt, M., Schneider, K., Schrader, F., Schroder, I., Scott, R. L., Sedlak, P., Serrano-Ortiz, P., Shao, C., Shi, P., Shironya, I., Siebicke, L., Sigut, L., Silberstein, R., Sirca, C., Spano, D., Steinbrecher, R., Stevens, R. M., Sturtevant, C., Suyker, A., Tagesson, T., Takanashi, S., Tang, Y., Tapper, N., Thom, J., Tomassucci, M., Tuovinen, J. P., Urbanski, S., Valentini, R., van der Molen, M., van Gorsel, E., van Huissteden, K., Varlagin, A., Verfaillie, J., Vesala, T., Vincke, C., Vitale, D., Vygodskaya, N., Walker, J. P., Walter-Shea, E., Wang, H., Weber, R., Westermann, S., Wille, C., Wofsy, S., Wohlfahrt, G., Wolf, S., Woodgate, W., Li, Y., Zampedri, R., Zhang, J., Zhou, G., Zona, D., Agarwal, D., Biraud, S., Torn, M., and Papale, D.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, Sci Data, 7, 225, 10.1038/s41597-020-0534-3, 2020.

Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J. M., Kimball, J. S., Ardizzone, J. V., Bosch, D., Colliander, A., Cosh, M., Kolassa, J., Mahanama, S. P., Prueger, J., Starks, P., and Walker, J. P.: Version 4 of the SMAP Level-4 Soil Moisture Algorithm and Data Product, J Adv Model Earth Sy, 11, 3106-3130, 10.1029/2019ms001729, 2019.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and

https://doi.org/10.5194/essd-2025-376
Preprint. Discussion started: 24 July 2025

Toll, D.: The Global Land Data Assimilation System, Bulletin of the American Meteorological Society, 85, 381-394, 10.1175/bams-85-3-381, 2004.
Rui, H., Beaudoing, H., and Loeser, C.: README Document for NASA GLDAS Version 2 Data Products, 2022.
Ruiz-Vásquez, M., O, S., Arduini, G., Boussetta, S., Brenning, A., Bastos, A., Koirala, S., Balsamo, G., Reichstein, M., and Orth, R.: Impact of updating vegetation information on land surface model performance, 10.22541/essoar.168182273.38487150/v1, 2023.
Ruiz-Vásquez, M., O, S., Brenning, A., Koster, R. D., Balsamo, G., Weber, U., Arduini, G., Bastos, A., Reichstein, M., and Orth, R.: Exploring the relationship between temperature forecast errors and Earth system variables, Earth System Dynamics, 13, 1451-1471, 10.5194/esd-13-1451-2022, 2022.
Running, S., Mu, Q., Zhao, M., and Moreno, A.: MODIS/Terra Net Evapotranspiration Gap-Filled 8-Day L4 Global 500m SIN Grid V061, https://doi.org/10.5067/MODIS/MOD16A2GF.061, 2021.
Schneider, U., Hänsel, S., Finger, P., Rustemeier, E., and Ziese, M.: GPCC Full Data Monthly Product Version 2022 at 0.25°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historical Data, 10.5676/DWD_GPCC/FD_M_V2022_025, 2022.
Slivinski, L. C., Compo, G. P., Sardeshmukh, P. D., Whitaker, J. S., McColl, C., Allan, R. J., Brohan, P., Yin, X., Smith, C. A., Spencer, L. J., Vose, R. S., Rohrer, M., Conroy, R. P., Schuster, D. C., Kennedy, J. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Kubota, H., Lee, T. C., Lorrey, A. M., Luterbacher, J., Mock, C. J., Przybylak, R. K., Pudmenzky, C., Slonosky, V. C., Tinz, B., Trewin, B., Wang, X. L., Wilkinson, C., and Wood, K.: An Evaluation of the Performance of the Twentieth Century Reanalysis Version 3, Journal of Climate, 34, 1417-1438, 10.1175/Jcli-D-20-0505.1, 2021.
Sun, Q. H., Miao, C. Y., Duan, Q. Y., Ashouri, H., Sorooshian, S., and Hsu, K. L.: A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons, Reviews of Geophysics, 56, 79-107, 10.1002/2017rg000574, 2018.
Tang, G. Q., Clark, M. P., and Papalexiou, S. M.: EM-Earth The Ensemble Meteorological Dataset for Planet Earth, Bulletin of the American Meteorological Society, 103, E996-E1018, 10.1175/Bams-D-21-0106.1, 2022.
Tang, R. L., Peng, Z., Liu, M., Li, Z. L., Jiang, Y. Z., Hu, Y. X., Huang, L. X., Wang, Y. Z., Wang, J. R., Jia, L., Zheng, C. L., Zhang, Y. Q., Zhang, K., Yao, Y. J., Chen, X. L., Xiong, Y. J., Zeng, Z. Z., and Fisher, J. B.: Spatial-temporal patterns of land surface evapotranspiration from global products, Remote Sensing of Environment, 304, 10.1016/j.rse.2024.114066, 2024.
Wang, H., Liu, J., Klaar, M., Chen, A., Gudmundsson, L., and Holden, J.: Anthropogenic climate change has influenced global river flow seasonality, Science, 383, 1009-1014, 10.1126/science.adi9501, 2024.
Wang-Erlandsson, L., Tobian, A., van der Ent, R. J., Fetzer, I., te Wierik, S., Porkka, M., Staal, A., Jaramillo, F., Dahlmann, H., Singh, C., Greve, P., Gerten, D., Keys, P. W., Gleeson, T., Cornell, S. E., Steffen, W., Bai, X. M., and Rockstrom, J.: A planetary boundary for green water, Nature Reviews Earth & Environment, 3, 380-392, 10.1038/s43017-022-00287-8, 2022.
Xie, P. P. and Arkin, P. A.: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, Bulletin of the American

Meteorological Society, 78, 2539-2558, 10.1175/1520-0477(1997)078<2539:Gpayma>2.0.Co;2, 1997.

Xie, P. P., Chen, M., and Shi, W.: CPC unified gauge-based analysis of global daily precipitation, 24th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc,

Xie, P. P., Joyce, R., Wu, S. R., Yoo, S. H., Yarosh, Y., Sun, F. Y., and Lin, R.: Reprocessed, Bias-Corrected CMORPH Global High-Resolution Precipitation Estimates from 1998, Journal of Hydrometeorology, 18, 1617-1641, 10.1175/Jhm-D-16-0168.1, 2017.

Yang, Y. T., Roderick, M. L., Guo, H., Miralles, D. G., Zhang, L., Fatichi, S., Luo, X. Z., Zhang, Y. Q., McVicar, T. R., Tu, Z. Y., Keenan, T. F., Fisher, J. B., Gan, R., Zhang, X. Z., Piao, S. L., Zhang, B. Q., and Yang, D. W.: Evapotranspiration on a greening Earth, Nature Reviews Earth & Environment, 4, 626-641, 10.1038/s43017-023-00464-3, 2023.

Yao, Y. J., Liang, S. L., Li, X. L., Hong, Y., Fisher, J. B., Zhang, N. N., Chen, J. Q., Cheng, J., Zhao, S. H., Zhang, X. T., Jiang, B., Sun, L., Jia, K., Wang, K. C., Chen, Y., Mu, Q. Z., and Feng, F.: Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations, J Geophys Res-Atmos, 119, 4521-4545, 10.1002/2013jd020864, 2014.

Zarei, M. and Destouni, G.: A global multi catchment and multi dataset synthesis for water fluxes and storage changes on land, Sci Data, 11, 1333, 10.1038/s41597-024-04203-1, 2024.

Zeri, M., Costa, J., Urbano, D., Cuartas, L., Ivo, A., Marengo, J., and Alvala, R.: A soil moisture dataset over the Brazilian semiarid region, Mendeley Data Version, 2, 2020.

Zhang, W. X., Zhou, T. J., and Wu, P. L.: Anthropogenic amplification of precipitation variability over the past century, Science, 385, 427-432, 10.1126/science.adp0212, 2024a.

Zhang, Y. Q., Kong, D. D., and Xu, Z. W.: PML_V2 global evapotranspiration and gross primary production (2000.02-2023.12), Zenodo, 10.5281/zenodo.10647617, 2024b.

Zhang, Y. Q., Kong, D. D., Gan, R., Chiew, F. H. S., McVicar, T. R., Zhang, Q., and Yang, Y. T.: Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002-2017, Remote Sensing of Environment, 222, 165-182, 10.1016/j.rse.2018.12.031, 2019.