Data-driven worldwide quantification of large-scale hydroclimatic co-variation patterns and comparison with reanalysis and Earth System modeling

Navid Ghajarnia[1], Zahra Kalantari[2], and Georgia Destouni[1]

[1]Stockholm University
[2]Department of Physical Geography, Stockholm University

November 23, 2022

Abstract

Large-scale co-variations of freshwater fluxes and storages on land can critically regulate green (vegetation) and blue (hydrosphere) water balances, land-atmosphere interactions, and hydroclimatic hazards. Such essential co-variation patterns still remain largely unknown over large scales and in different climates around the world. To contribute to bridging this large-scale knowledge gap, we synthesize and decipher different data time series over the period 1980-2010 for 6405 hydrological catchments around the world. From observation-based data, we identify dominant large-scale co-variation patterns between main freshwater fluxes and soil moisture (SM) for different world parts and climates. These co-variation patterns are also compared with those obtained from reanalysis products and Earth System Models (ESMs). The observation-based datasets robustly show the strongest large-scale hydrological co-variation relationship to be that between SM and runoff (R), consistently across the study catchments and their different climate characteristics. The predominantly strongest large-scale SM-R co-variation relationship, however, is also the most misrepresented by ESMs and reanalysis products, followed by that between precipitation and R. Comparison between corresponding observation-based and ESM results also shows that an ESM may perform well for individual hydrological variables, but still fail in representing the patterns of large-scale co-variations between variables.

Data-driven worldwide quantification of large-scale hydroclimatic co-variation patterns and comparison with reanalysis and Earth System modeling

Navid Ghajarnia[1]*, Zahra Kalantari[1&2], and Georgia Destouni[1]

[1] Department of Physical Geography, Bolin Center for Climate Research, Stockholm University, SE-10691 Stockholm, Sweden

[2] Department of Sustainable Development, Environmental Science and Engineering, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

*Corresponding author: Navid Ghajarnia (navid.ghajarnia@natgeo.su.se)

Key points

- The closest large-scale co-variations of studied freshwater fluxes and storages in different climates are between soil moisture and runoff
- Blue-water precipitation-runoff and green-water precipitation-evapotranspiration co-variations are mostly weaker
- Soil moisture-runoff and precipitation-runoff co-variations are the most misrepresented by studied reanalysis products and Earth System models

Abstract

Large-scale co-variations of freshwater fluxes and storages on land can critically regulate green (vegetation) and blue (hydrosphere) water balances, land-atmosphere interactions, and hydroclimatic hazards. Such essential co-variation patterns still remain largely unknown over large scales and in different climates around the world. To contribute to bridging this large-scale knowledge gap, we synthesize and decipher different data time series over the period 1980-2010 for 6405 hydrological catchments around the world. From observation-based data, we identify dominant large-scale co-variation patterns between main freshwater fluxes and soil moisture (SM) for different world parts and climates. These co-variation patterns are also compared with those obtained from reanalysis products and Earth System Models (ESMs). The observation-based datasets robustly show the strongest large-scale hydrological co-variation relationship to be that between SM and runoff (R), consistently across the study catchments and their different climate characteristics. The predominantly strongest large-scale SM-R co-variation relationship, however, is also the most misrepresented by ESMs and reanalysis products, followed by that between precipitation and R. Comparison between corresponding observation-based and ESM results also shows that an ESM may perform well for individual hydrological variables, but still fail in representing the patterns of large-scale co-variations between variables.

Keywords:

Large-scale hydro-climate, Hydrological correlation patterns, Observational data, Earth System Models, Reanalysis products, Hydrological catchments, Multi-catchment study

1. Introduction

Different hydrological variables may be expected to co-vary (in similar or various ways) over both small and large spatiotemporal scales on land, since the world's freshwater fluxes and storages, even though heterogeneous, are still coupled hydraulically by local continuity of water pressure, momentum and mass, as well as hydrologically by overarching water balance (large-scale continuity of mass) (Destouni et al., 2010, 2013). These multi-level linkages imply connectivity across all local interfaces of soil water with groundwater and vegetation water (Maxwell & Condon, 2016), and of subsurface with surface water (Bosson et al., 2012), as well as between blue (hydrosphere) and green (vegetation) water (Falkenmark & Rockström, 2006). Resulting large-scale co-variations of the latter regulate the balance of green and blue water resource availability and security (Mastrotheodoros et al., 2020), essential for both the green-water uses by terrestrial vegetation, forestry and agriculture (Gudmundsson et al., 2014; Orth et al., 2020), and the various blue-water uses by freshwater ecosystems (Albert et al., 2020) and human societies (for households, energy supply, agricultural irrigation, and other sectors) (Medellín-Azuara et al., 2007; Koutsouris et al., 2010; Madani & Lund, 2010; Naumann et al., 2015). These large-scale co-variations also regulate land-atmosphere interactions (Maxwell & Kollet, 2008; Seneviratne et al., 2010) and hydro-climatic hazards (Raymond et al., 2020), e.g., through co-occurrences of high precipitation (P), soil moisture (SM) and runoff (R) anomalies for floods (Kalantari et al., 2019), and low P, SM, R and actual evapotranspiration (ET) anomalies for droughts (Orth & Destouni, 2018) and concurrent heat-drought events (Raymond et al., 2020). Figure 1 illustrates schematically main connections of P, ET and R fluxes, and contributions to these from SM, groundwater and surface water storage variations, which may to some degree also regulate resulting large-scale flux co-variation patterns on land.

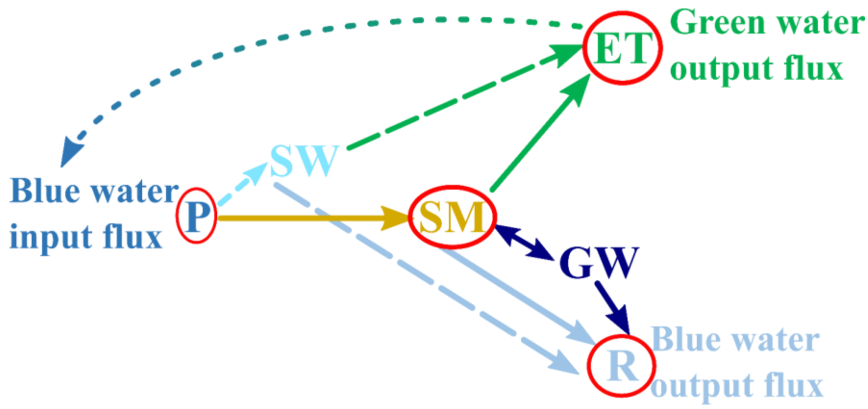

Figure 1. Simple schematic illustration of main freshwater fluxes and storage connections on land. Main freshwater fluxes of precipitation (P), total actual evapotranspiration (ET), and total runoff (R) interact with soil moisture (SM), surface water (SW), and groundwater (GW) storages and their respective flux contributions to total ET and R.

Local mechanistic relationships and co-variations of water flux and storage variables have been amply investigated in local hydrological studies. However, large-scale hydrological co-variation patterns remain largely unknown in and across different parts and climates of the world (Van Loon, 2015; Thorslund et al., 2017). This large-scale knowledge gap leads to relatively poor hydrological and Earth System model links and performances in the changing climate across the world (Khatami et al., 2019; Saft, M., M. C. Peel, A. W. Western, 2016; Thirel et al., 2015; Vaze et al., 2010; Asokan et al., 2016; Bring et al., 2015, 2019; Törnqvist et al., 2014). One reason for such poor model performances and cross-model inconsistencies may be that large-scale models have mostly focused on just near-surface interactions of main water fluxes with SM (Clark et al., 2015; Brocca et al., 2010). However, Earth's major freshwater storage is subsurface, including both groundwater and SM all the way down to the depth of the groundwater table; through the subsurface water pathways that continuously feed into the surface water networks of whole catchments (Cvetkovic et al., 2012), these deeper subsurface water systems may play important regulatory roles for resulting large-scale ET and R fluxes (Moshir Panahi et al., 2020; Destouni & Verrot, 2014). Furthermore, the re-circulating part of ET that feeds into total P in the same hydrological catchment (Figure 1) may also regulate emergent large-scale hydrological co-variation patterns, and differently so in various parts and climates of the world.

Recent community identification of key hydrological knowledge gaps has emphasized needs for hydrological research to go beyond fragmented local understanding (Blöschl et al., 2019) to multi-catchment studies that can identify large-scale hydrological patterns (Ghajarnia et al., 2020; Orth et al., 2020; Berghuijs et al., 2019; Orth & Destouni, 2018; Destouni et al., 2013; Blöschl, 2006). In their recent multi-catchment study over Europe, Ghajarnia et al. (2020) found that, among the main flux and storage variables depicted in Figure 1, sufficiently long-term time-series for consistent data-driven determination of large-scale co-variations based on different data sources are primarily available for P, R, ET, and SM. In the present study, we investigate emergent co-variation patterns of these key hydro-climatic variables worldwide based on multi-catchment data from 6405 hydrological catchments of different scales and climates around the world (Figure 2a). These study catchments are selected based on consistent data availability from different comparative sources, with the aim to answer the following main research questions: (1) What large-scale hydrological co-variation patterns emerge from different observation-based datasets for P, ET, R, and SM in different parts and climates of the world? (2) How well do reanalysis products and Earth System model (ESM) results represent the observation-based emergent hydrological co-variation patterns?

To this end, we considered root-zone (and not just near-surface) SM along with P, R, and ET over each of the 6405 catchments, comparatively based on consistent monthly data over the time period 1980-2010 from observations, reanalysis products, and ESM results. Observational gauge-based and gridded data were then available for quantification of catchment-scale P and R (and surface temperature (T), also included in the study for climate consideration), while only reanalysis data were available to complement these with, in different comparative combinations, for corresponding catchment-scale ET and SM.

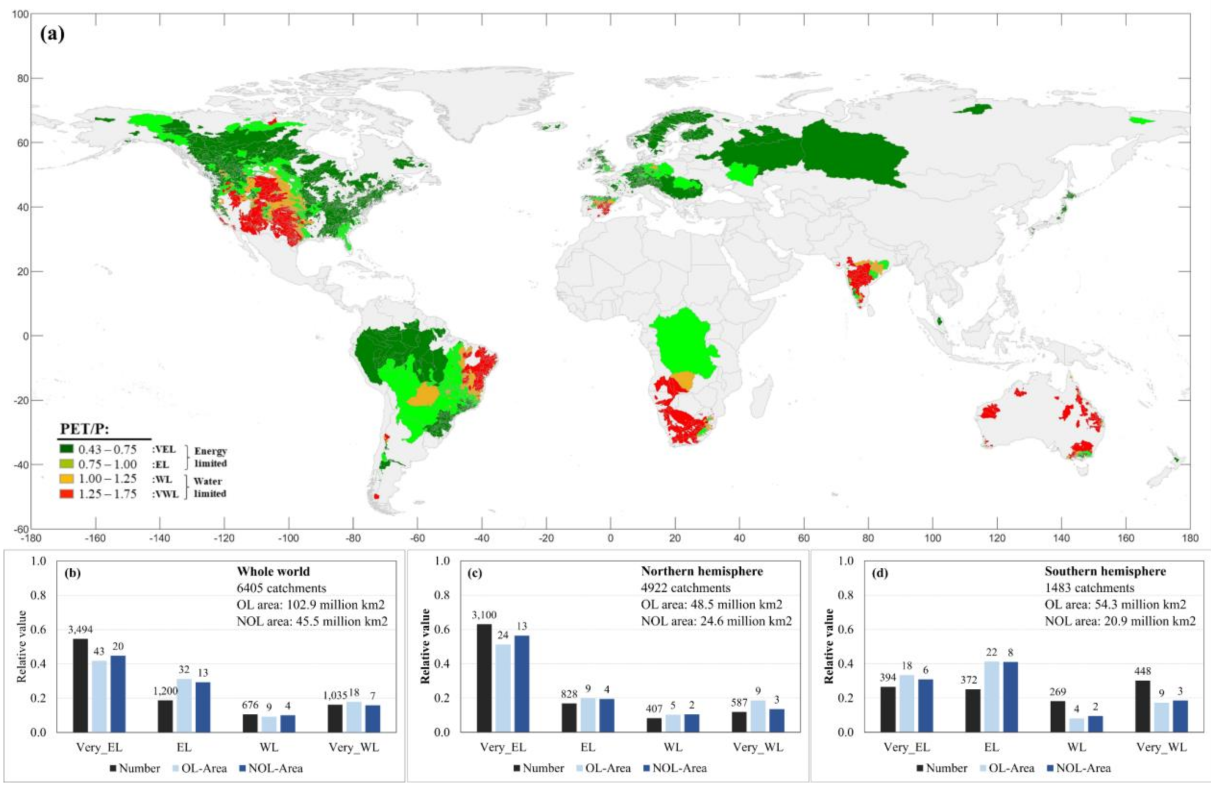

Figure 2. Map of the 6405 catchments included in this study and their corresponding area coverage. Panel (a) shows the location of catchments, classified into four categories of very energy-limited (VEL), energy-limited (EL), water-limited (WL), and very water-limited (VWL) conditions, based on their aridity index (PET/P, where PET is potential evapotranspiration estimated based on ref. (Langbein, 1949) and P is precipitation). Panels (b-d) show the relative numbers and areas of catchments in the different categories; area values include both overlapping (OL) and non-overlapping (NOL) catchment area over the: (b) whole world, (c) Northern hemisphere, and (d) Southern hemisphere. Absolute number and area values are shown above each bar and in total inside the (b-d) panels. Relative values are obtained by dividing the absolute values (numbers above each bar) with the corresponding total values (numbers given top-right in each panel)

2. Materials and Methods

2.1 Datasets

Table 1 lists the different sources of data used for each variable and synthesized in 10 different datasets: three semi-observational (Semi-Obs1-3); two pure reanalysis datasets from either ERA5 or GLDAS; and outputs from five ESMs in the Coupled Model Intercomparison

Project Phase 6 (CMIP6; Eyring et al., 2016). The purely observational data for R, P, and T in the three semi-observational datasets are retrieved from station-based GSIM (Global Streamflow Indices and Metadata (Do et al., 2018; Gudmundsson et al., 2018)), and gridded GPCC-V7 (Global Precipitation Climatology Centre-Version7 (Schneider et al., 2016)) and GHCN-CAMS (Global Historical Climatology Network-Climate Anomaly Monitoring System (Fan & van den Dool, 2008)), respectively. With no globally consistent observational database available for SM or ET, we used in the semi-observational datasets alternative data for these variables from different reanalysis products, which many researchers regard as the closest available alternative to direct observations (Dee et al., 2011).

The Semi-Obs datasets include SM and ET data from different (or independent) reanalysis products (Table 1) in order to facilitate checking of the possible dependence of associated observation-based co-variation results on specific relationships built into different reanalysis products. The three independent Semi-Obs datasets thus include observational data for P, R and T in different combinations with SM and ET data from the reanalysis products ERA5 (European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis 5th Generation (Copernicicus Climate Change Service (C3S), 2017)), GLDAS (Global Land Data Assimilation System; NOAH025 M2.0 (Beaudoing, H. and M. Rodell, 2019; Rodell et al., 2004)), and GLEAM-3.3a (Global Land Evaporation Amsterdam Model (Martens et al., 2017; Miralles et al., 2011)). The three Semi-Obs datasets are independent by including data for SM and ET from these different reanalysis products. This comparative approach was used to test the robustness of and quantify the uncertainty in resulting co-variation patterns from the different Semi-Obs datasets.

Table 1. Data used in 10 different combinations of datasets

Variable	Semi-observational			Reanalysis		Earth System Models (ESMs)				
	1	2	3	1	2	1	2	3	4	5
Precipitation	GPCC-V7[a]	GPCC-V7[a]	GPCC-V7[a]	ERA5[d]	GLDAS NOAH025 M2.0[b]	BCC-CSM2-MR[g]	CanESM5[h]	EC-Earth3-Veg[i]	MIROC6[j]	MPI-ESM[k]
Soil moisture	GLDAS NOAH025 M2.0[b]	ERA5[d]	GLEAM v3.3a[f]	ERA5[d]	GLDAS NOAH025 M2.0[b]	BCC-CSM2-MR[g]	CanESM5[h]	EC-Earth3-Veg[i]	MIROC6[j]	MPI-ESM[k]
Runoff	GSIM[c]	GSIM[c]	GSIM[c]	ERA5[d]	GLDAS NOAH025 M2.0[b]	BCC-CSM2-MR[g]	CanESM5[h]	EC-Earth3-Veg[i]	MIROC6[j]	MPI-ESM[k]
Actual Evapo-transpiration	ERA5[d]	GLEAM v3.3a[f]	GLDAS NOAH025 M2.0[b]	ERA5[d]	GLDAS NOAH025 M2.0[b]	BCC-CSM2-MR[g]	CanESM5[h]	EC-Earth3-Veg[i]	MIROC6[j]	MPI-ESM[k]
Temperature	GHCN-CAMS[e]	GHCN-CAMS[e]	GHCN-CAMS[e]	ERA5[d]	GLDAS NOAH025 M2.0[b]	BCC-CSM2-MR[g]	CanESM5[h]	EC-Earth3-Veg[i]	MIROC6[j]	MPI-ESM[k]

[a] Global Precipitation Climatology Centre-Version7 (Schneider et al., 2016)
[b] Global Land Data Assimilation System (GLDAS) NOAH025 M2.0 (Beaudoing and Rodell, 2019; Rodell et al., 2004)
[c] Global Streamflow Indices and Metadata (Do et al., 2018; Gudmundsson et al., 2018)
[d] European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis 5th Generation (ERA5) (Copernicicus Climate Change Service (C3S), 2017)
[e] Global Historical Climatology Network-Climate Anomaly Monitoring System (Fan & van den Dool, 2008)
[f] Global Land Evaporation Amsterdam Model (Martens et al., 2017; Miralles et al., 2011)
[g] Beijing Climate Center Climate System Model (Wu et al., 2018)
[h] Canadian Earth System Model version 5 (Swart et al., 2019)
[i] European Community Earth-Vegetation model version 3 (EC-Earth Consortium, 2019)
[j] Model for Interdisciplinary Research on Climate (Tatebe & Watanabe, 2018)
[k] Max Planck Institute for Meteorology-Earth System Model (Neubauer et al., 2019)

After analyzing and obtaining large-scale hydrological co-variation patterns from the Semi-Obs datasets, we further compared these with corresponding co-variation patterns obtained from the pure reanalysis datasets (ERA5 and GLDAS), or from the CMIP6 results of each of the following ESMs: BCC-CSM2-MR (Beijing Climate Center Climate System Model (Wu et al., 2018)), CanESM5 (Canadian Earth System Model version 5 (Swart et al., 2019)), EC-Earth3-Veg (European Community Earth-Vegetation model version 3 (EC-Earth Consortium, 2019)), MIROC6 (Model for Interdisciplinary Research on Climate (Tatebe & Watanabe, 2018)), and MPI-ESM (Max Planck Institute for Meteorology-Earth System Model (Neubauer et al., 2019)). The selection of ERA5 and GLDAS for this comparison was based on these being the most widely used reanalysis products (Koohi et al., 2019). The five comparative ESMs were chosen because they were found to be relatively well performing in previous catchment-based hydrological ESM evaluations (Bring et al., 2019; Bring et al., 2015). For further exploration of results on large-scale hydrological co-variation patterns obtained with the root-zone SM mainly used in this study, we also compared these with corresponding results obtained by instead using in the Semi-Obs 2 dataset remotely-sensed surface-layer SM data from the ESA-CCI SM-v04.5 combined product (European Space Agency-Climate Change Initiative soil moisture product version 4.5 (Dorigo et al., 2017; Gruber et al., 2017, 2019)).

2.2 Selection of the hydrological catchments

Observation-based R data time series were needed and used for overall water-balance closure in each hydrological study catchment. The GSIM database with independent outlet discharge data for 30 958 hydrological catchments around the world was used for this purpose. However, in terms of R data (discharge divided by contributing catchment area), the GSIM database only included 8217 catchments that met the condition of having at least 300 non-missing monthly R values (corresponding to 25 years) within the study period 1980-2010. Furthermore, based on quality controls of the R data, 6405 catchments were finally included in the analysis. The quality controls included comparisons of reported catchment area and area of the corresponding catchment polygon in the GSIM database. Based on these, catchments with wrongly delineated catchment polygons (e.g. triangle polygons introduced as the catchment), and catchments with clear outlier discharge values (too high or too low in the outlier range) were removed from further analysis.

The final set of 6405 catchments in different parts of the world, with various climatic, geographic, and hydrological conditions, was further classified into four categories (see Figure 2a) based on their aridity index (calculated as PET/P ratio, where PET is potential evapotranspiration estimated by Langbein (1949) and P is precipitation). Hydrological co-variations in many of the included study catchments may also be influenced by various anthropogenic activities, such as agriculture and its irrigation, dams and associated artificial reservoirs, and enhanced ET by these (Destouni et al., 2013). However, we did not in this study seek to investigate and resolve all drivers and catchment-internal processes that can influence hydrological co-variations. The study aim was instead to identify if and what dominant large-scale patterns may emerge for such co-variations across a global set of catchments with widely varying drivers and internal processes, and further evaluate the performance of reanalysis and ESMs in capturing these large-scale patterns.

2.3 Spatial aggregations

To capture different large-scale co-variation patterns that may emerge under widely different hydroclimatic conditions, the studied catchments are divided into four main such classes (Figure 2a): very energy-limited (VEL; dark green), energy-limited (EL; light green), water-limited (WL; orange), and very water-limited (VWL; red). This classification is based on catchment-average aridity index (PET/P), where PET is potential evapotranspiration estimated by Langbein (1949). The R data is station-based and directly representative of integrating discharge from each catchment, while the data for each gridded study variable (P, ET, SM, T) is aggregated through area-weighted averaging over each catchment to get corresponding catchment-wise variable time series. In this area-weighted averaging over each catchment, overlying areas of each grid with the catchment polygons are used as the averaging weights. Moreover, all catchment-wise data are further aggregated over all catchments in each climate category to get large-scale variable time series representative of each climate category. In this cross-catchment aggregation, the total surface area of each catchment is the weight in the area-weighted averaging. This procedure was followed for all observational, reanalysis and ESM data, to get catchment–wise and climate-category aggregated monthly time series for all variables.

Since some of the study catchments are overlapping, we have also calculated and compared the total overlapping (OL) and non-overlapping (NOL) areas of each climate category globally (Figure 2b), and over the Northern (Figure 2c) and Southern (Figure 2d) hemispheres. This shows that the OL and NOL catchment areas sample similar relative area shares of the different climate categories (relative OL and NOL area values are similar for and across the categories). As such, the OL catchment statistics can be used as a relatively unbiased sample representation of the NOL global or hemispheric catchment area conditions and statistics. By considering non-weighted OL catchment statistics we get a larger sample of catchments from which we can also assess how individual catchments of any scale behave and compare statistically with the large-scale area-weighted behavior and statistics obtained for each climate-zone aggregation of catchments.

2.3 Statistical co-variation and model error quantifications

Co-variation patterns were quantified and compared in terms of resulting coefficients of determination (r^2) for linear regression between pairs of variable anomalies in each catchment, and various r^2 statistics across the catchments in each climate category, for all comparative datasets (Table 1). The absolute monthly value in the time series of each variable were transformed to normalized monthly anomalies as:

$$NorX_{m,y} = \frac{X_{m,y} - \overline{X_m}}{\sigma_{X_m}} \quad (1)$$

where $NorX_{m,y}$ is normalized anomaly of variable $X_{m,y}$ for month m of each year y, and $\overline{X_m}$ and σ_{X_m} are the long-term average value and standard deviation of monthly X, respectively, for each month m over all years in the study period 1980-2010.

To quantify large-scale co-variation patterns from the r^2 results for each variable pair in and across all catchments of each climate category, globally and in each hemisphere, area-weighted average r^2 values were calculated over the entire variable pair time series, as well as for each specific month across all years in the series. A corresponding area-weighted standard deviation (σ^*) among catchment results was also calculated as (Hawley et al., 1988):

$$\sigma^* = \sqrt{var(r^2)\sum_{j=1}^{N}\left(\frac{A_j}{\sum_{j=1}^{N} A_j}\right)^2} \tag{2}$$

where A_j is the area of catchment j and N is the number of catchments. In this way, larger catchments have greater influence than smaller ones on resulting area-weighted co-variation statistics. In addition, non-weighted r^2 statistics were also calculated to represent hydrological co-variation behaviors and statistics irrespective of catchment scale, for comparison with the above-described area-weighted result statistics.

Moreover, to evaluate and rank ESM and reanalysis product performance, statistical error measures were calculated for individual hydroclimatic variables in terms of mean absolute error (MAE), standard deviation (SD), correlation coefficient (CC), and root mean squared difference (RMSD). These error measures were calculated for absolute variable values in each catchment and further averaged across the catchments. In these calculations for individual hydrological variable performance, ESM and reanalysis product results were compared with purely observational data from GPCC for P, GSIM for R, GHCN-CAMS for T, and reanalysis data from GLEAM3.3-a for SM and ET. The GLEAM3.3-a dataset was chosen for the latter so that ERA5 and GLDAS result performance could also be assessed against some alternative way to obtain the SM and ET variables.

For error measure and ranking based on co-variation results for different pairs of variables, the r^2 results from the ESM and reanalysis products were compared with corresponding reference results from the Semi-Obs1 dataset. The MAE for r^2 values obtained from the former was then calculated by comparison with the reference r^2 values obtained from the Semi-Obs1 dataset, and further divided by the SD of the r^2 time series in the Semi-Obs1 dataset.

3. Results and Discussion

Figure 3 shows resulting co-variation patterns between different variable pairs for the catchments in each climate category and for all studied datasets (Table 1). The results show consistent, robust large-scale co-variation patterns obtained from the three Semi-Obs datasets (Figure 3a-c), but considerable differences from these in results obtained from pure reanalysis data (Figure 3d,e) and ESM outputs (Figure 3f-j).

The Semi-Obs dataset results (Figure 3a-c) are consistent with those from reanalysis (Figure 3d-e) in showing the overall strongest co-variation relationship (highest average r^2) between SM and R (even stronger from reanalysis than the Semi-Obs datasets). The Semi-Obs datasets further consistently - and compared with the other results apparently surprisingly - exhibit the weakest co-variation relationship (lowest average r^2) between P and R. The reanalysis and most ESM results exhibit considerably stronger P-R relationships, but with different changes towards either stronger or weaker average r^2 for drier, more water-limited climate conditions. In contrast, the Semi-Obs datasets show consistently stable and small r^2 between P and R across all hydroclimatic conditions.

Results are overall consistent across all studied datasets in showing stronger (higher average r^2 for the) P-ET and SM-ET relationships in drier and more water-limited climate. However, the ESM results exhibit widely varying r^2 values for these relationships and mostly too high r^2 for the P-ET relationship in the driest conditions, compared with both the Semi-Obs and the reanalysis results. The P-SM co-variation strength is that differing the most between the different Semi-Obs datasets, but these differences are similar to those exhibited by the reanalysis and ESM results, as are also the average, relatively small r^2 levels resulting from the different datasets.

Overall, the hydrological co-variation patterns emerging from the Semi-Obs datasets (Figure a-c) are more consistent with those obtained from reanalysis products (Figure 3d-e) than the ESM results (Figure 3f-j). The latter show diverse biases and inconsistencies among the different ESMs, as well as between these and the Semi-Obs and reanalysis results, for all studied variable co-variations. Furthermore, even though the reanalysis data capture the overall strongest SM-R co-variation, they tend to overestimate this as well as the P-R co-variation strength. Most importantly, the ERA5 (Figure 3d) and GLDAS (Figure 3e) results diverge in how the SM-R and P-R co-variation strengths change with drier and more water-limited climate. For the weaker P-ET, P-SM, and SM-ET co-variations, the reanalysis results are more consistent with those for the Semi-Obs datasets.

The ESM results are particularly inconsistent with other datasets for BCC-CSM2-MR (Figure 3f), which totally misrepresents the SM-R, P-R and P-ET co-variations, and greatly underestimates the strongest first two. CanESM5 (Figure 3g) overestimates co-variations of P-SM, SM-ET, P-R (in the VEL and EL categories), and P-ET (in the VWL category), while MPI-ESM (Figure 3h) and EC-Earth3-Veg (Figure 3i) overestimate the P-R co-variations and also imply unrealistic increases in these for drier conditions, and MIROC6 (Figure 2j) underestimates SM-R and overestimates P-R co-variation.

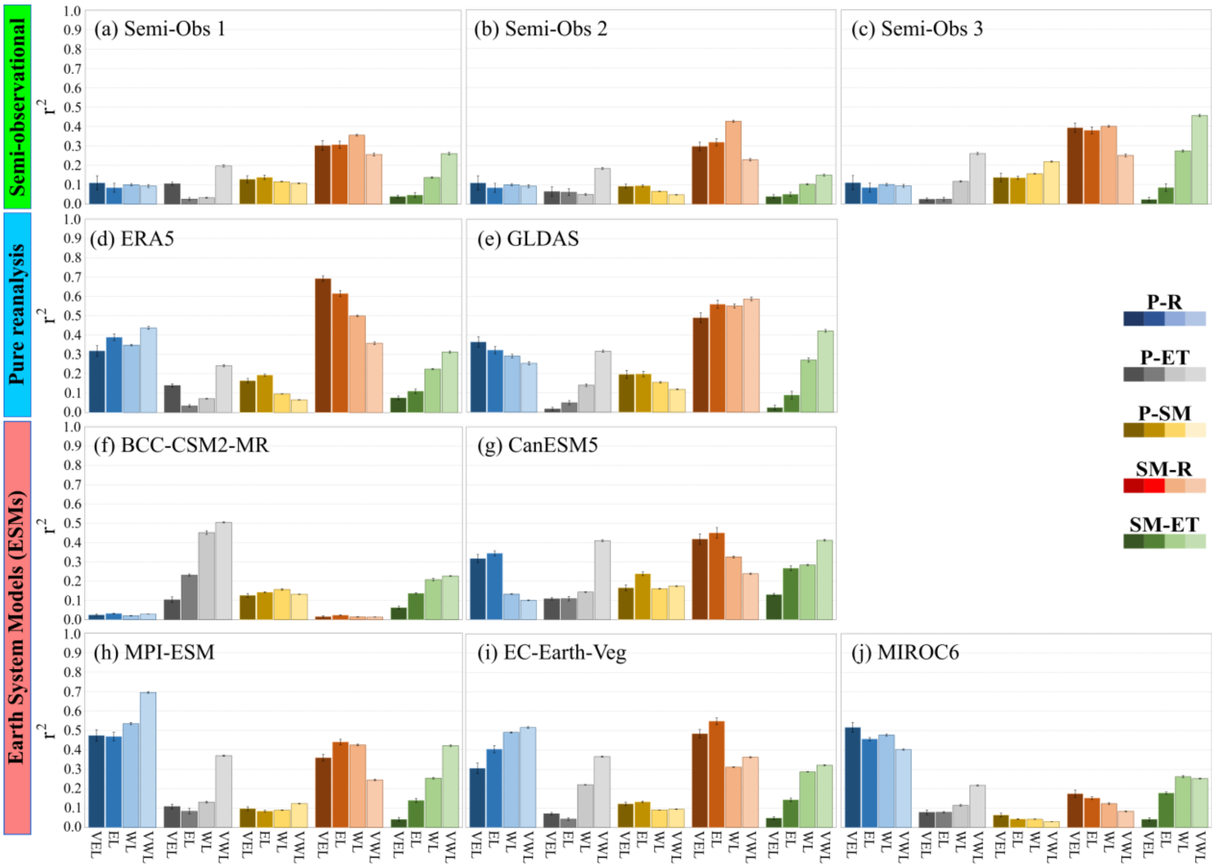

Figure 3 – Area-weighted average coefficient of determination (r^2) values for linear regression of co-variations in precipitation-runoff (P-R), precipitation-evapotranspiration (P-ET), precipitation-soil moisture (P-SM), soil moisture-runoff (SM-R), and soil moisture-evapotranspiration (SM-ET). Results are shown for data from the (a-c) semi-observational datasets, (d, e) reanalysis datasets, and (f-j) Earth System Model (ESM) outputs. Catchments are classified as in Figure 2 into four categories of very energy-limited (VEL), energy-limited (EL), water-limited (WL), very water-limited (VWL) conditions. Area-weighted standard deviation of the r^2 values among catchments (calculated according to Eq. (2)) is also shown by error bars. See Table 1 for reanalysis product and ESM acronyms and source references.

Supporting Figure S1 shows r^2 box plots for each variable pair with non-weighted statistics, corresponding to the area-weighted statistics in Figure 3. Result comparison shows much greater variability of r^2 values among catchments for non-weighted than area-weighted statistics. This may be due to smaller catchments varying more between them than larger one, and the effects of small-catchment variability being dampened by area-weighting. Furthermore, catchment scale weighting (or not) affects more the P-R than other co-variation results, which show smaller differences between corresponding area-weighted and non-weighted results. The average level of r^2 for the P-R relationship is greater for non-weighted than area-weighted statistics, with the latter being more influenced by large-catchment behavior and the difference being most notable for wet, energy limited catchments (VEL and EL categories). This may be due to that larger catchments encompass greater spatial heterogeneity within them (rather than between them), e.g., in human activities, groundwater flows into stream networks, prevalence of lakes, wetlands and other land-covers. In combination, such greater within-catchment heterogeneity may dampen direct P-R relationships in the larger catchments. In contrast to the Semi-Obs datasets, however, the pure reanalysis and ESM datasets do not capture this difference between larger- and smaller-scale catchments (Supporting Figure S1).

Overall, the comparison between Figure 3 and Supporting Figure S1 shows that the SM-R co-variation is still strong, also when considering non-weighted statistics with greater influence of small catchments. This may be due to the relatively close hydraulic connection of the whole root-zone SM (considered here) with the groundwater table and its variations, which in turn determine variations in hydraulic gradient and groundwater flow to streams and total R. Groundwater flow distance to nearest downgradient stream depends on stream density within each catchment, which may not differ much between smaller and larger catchments in the same hydro-climatic region (Darracq et al., 2010).

Supporting Figure S2 further shows co-variation results for just near-surface SM data from ESA-CCI SM-v04.5 (European Space Agency-Climate Change Initiative soil moisture product version 4.5 (Dorigo et al., 2017; Gruber et al., 2017, 2019) while keeping all other data as in the Semi-Obs2 dataset. ESA-CCI is a global remotely-sensed product that measures SM only in the top soil layer at the surface, whereas the other reanalysis products capture SM over the root zone. Result comparison shows considerably closer P-SM co-variations when considering just the surface SM (Supporting Figure S2a) instead of the whole root-zone SM (Figure 3.b) and the opposite for SM-R co-variations. These comparative results are as expected from the surface SM having more (less) direct hydraulic connection to P at the surface (the underlying groundwater table) than the whole root-zone SM. Thereby the surface SM is more (less) directly related than the root-zone SM to variations/anomalies of P and associated infiltration to SM at the surface (of groundwater table depth and associated groundwater flow contributions to total R).

Furthermore, for monthly co-variations, Figure 4 and Figure 5 show average monthly r^2 values for the northern and the southern hemisphere, with different seasonal characteristics, respectively, for one semi-observational dataset (Semi-Obs1), one reanalysis product (GLDAS), and one ESM (MPI-ESM); complete results for all datasets are shown in Supporting Figures S3-8. Results for Semi-Obs1 (Figure 4-5a) show consistently stronger monthly co-variation for SM-R than for P-R and the other variable pairs in the northern hemisphere, except for the dry VWL catchments in the southern hemisphere, where the co-variation of SM-R is smaller than that of SM-ET. The GLDAS reanalysis results (Figure 4b) are also largely consistent with those of Semi-Obs1, but with closer SM-R and P-R co-variations in the southern hemisphere (Figure 5b). The MPI-ESM, however, only weakly captures the data-given monthly co-variation patterns, and particularly overestimates the P-R co-variation (Figure 4-5c). Overall, the reanalysis and ESMs capture the stronger SM-ET co-variations in drier (more water limited) catchments, especially in the northern hemisphere (Figure 4 and Supporting Figures S5-8). In the southern hemisphere, monthly r^2 values differ considerably from those in the Semi-Obs datasets for both the reanalysis and the ESMs, which may be partly due to the relative lack of data for this hemisphere both for our comparison and for general reanalysis and ESM development and testing.

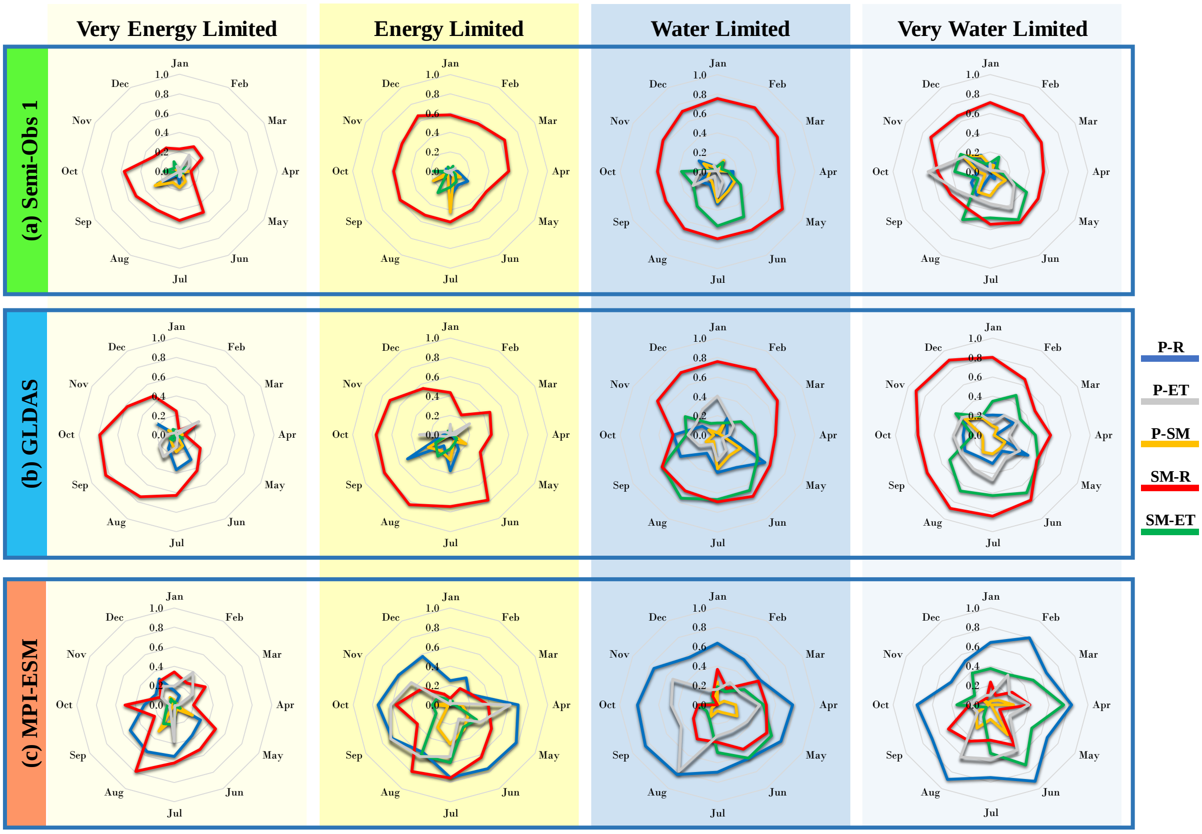

Figure 4. – Monthly r^2 values from linear regression of normalized anomalies ($NorX_{m,y}$, Eq. (1)) of variable pairs: precipitation-runoff (P-R) (blue), precipitation-evapotranspiration (P-ET) (grey), precipitation-soil moisture (P-SM) (yellow), soil moisture-runoff (SM-R) (red), and soil moisture-evapotranspiration (SM-ET) (green) in (a) semi-observational dataset 1 (Semi-Obs 1), (b) GLDAS reanalysis product, and (c) MPI-ESM Earth System Model (ESM). Results shown are for very energy-limited, energy-limited, water-limited, and very water-limited catchment classes in the northern hemisphere. See Table 1 for reanalysis product and ESM acronyms and source references.

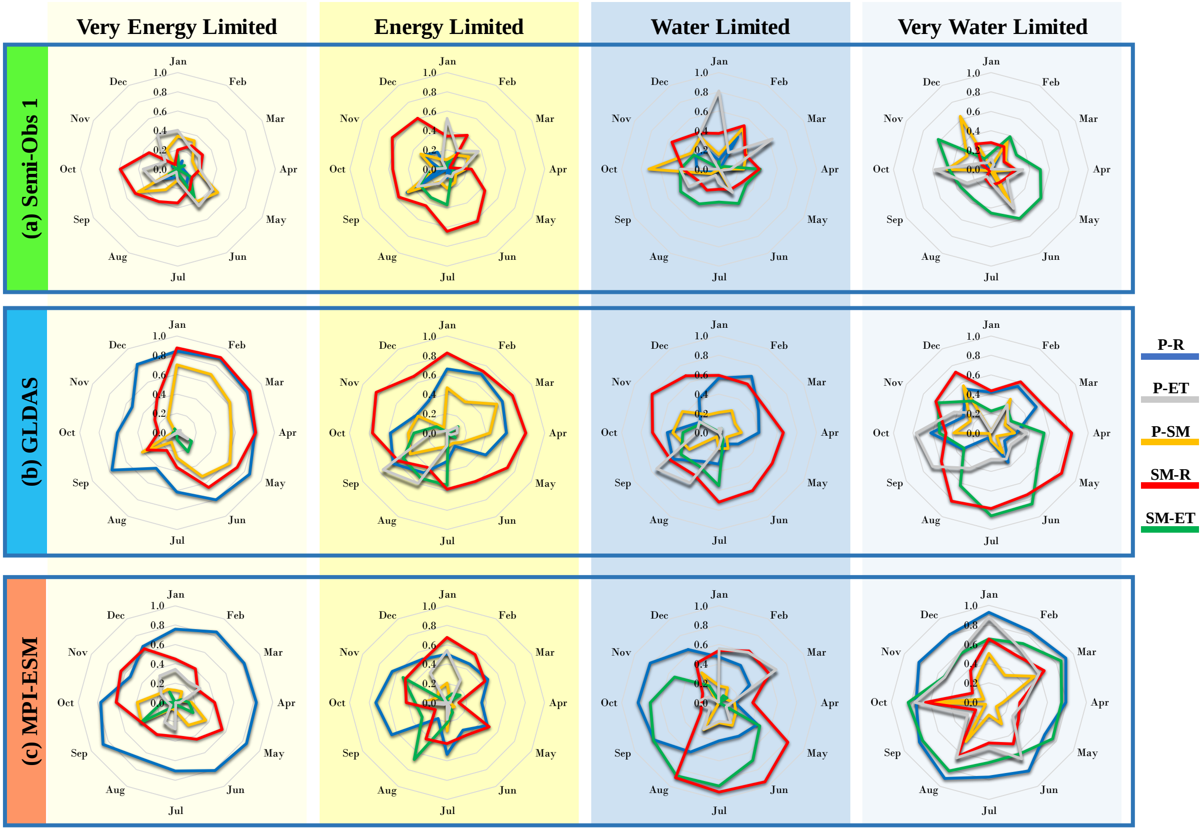

Figure 5. – Monthly r^2 values from linear regression of normalized anomalies ($NorX_{m,y}$, Eq. (1)) of variable pairs: precipitation-runoff (P-R) (blue), precipitation-evapotranspiration (P-ET) (grey), precipitation-soil moisture (P-SM) (yellow), soil moisture-runoff (SM-R) (red), and soil moisture-evapotranspiration (SM-ET) (green) in (a) semi-observational dataset 1 (Semi-Obs 1), (b) GLDAS reanalysis product, and (c) MPI-ESM Earth System Model (ESM). Results shown are for very energy-limited, energy-limited, water-limited, and very water-limited catchment classes in the southern hemisphere. See Table 1 for reanalysis product and ESM acronyms and source references.

Figure 6 finally shows reanalysis and ESM performance relative to the semi-observational datasets for both individual hydroclimatic variables (Figure 6a) and their pair-wise co-variations (Figure 6b) across all 6405 study catchments. The reanalysis products exhibit overall lower relative error (MAE/SD) than the ESMs and, among the latter, EC-Earth and MIROC have the lowest MAE/SD for individual variables. ESM errors are greatest for SM followed by R, while the lowest errors are found for T and ET. Similar conclusions are reached based on Taylor diagrams (Supporting Figure S9), and these individual variable findings are also consistent with those in other studies showing generally better agreement between ESMs and observation data for T than for hydrological variables (Asokan et al., 2016; Arvid Bring et al., 2015; Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S.C., Collins, W., 2013; Woldemeskel et al., 2012). For the co-variations of variables, Figure 6b shows highest reanalysis and ESM errors relative to data for the SM-R relationship, followed by the P-ET or P-R relationships, while the SM-ET relationship is associated with the smallest co-variation errors. Comparison between Figure 6a and Figure 6b shows that relatively good performance (small error) for individual hydroclimatic variables does not necessarily lead to good performance also in the variable co-variation patterns. This means that models may rank differently based on individual variable performance than on co-variation performance. For instance, ERA5 is ranked 1st for R and SM individually (based on Figure 6a), while it shows the third worst performance and MPI-ESM is the 1st ranked model for the SM-R covariation relationship (Figure 6-b).

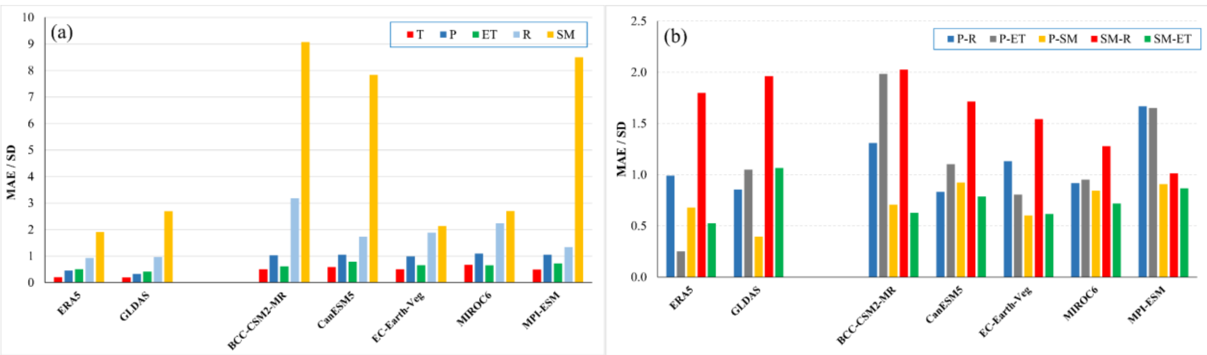

Figure 6. – Performance assessment of reanalysis products and Earth System models (ESMs) in terms of mean absolute error (MAE) relative to values in the reference dataset Semi-Obs 1 (Table 1) divided by the Semi-Obs 1 standard deviation (SD). Results are shown for (a) individual variable values, and (b) co-variations between variable pairs, as quantified by the coefficient of determination (r^2) in linear regression for each pair. The variables considered are precipitation (P), runoff (R), temperature (T), evapotranspiration (ET) and soil moisture (SM). See Table 1 for reanalysis product and ESM acronyms and source references.

Conclusion

A main answer to the first research question of this study is that the observation-based datasets robustly show the closest large-scale hydrological co-variations to be those between root-zone SM and the blue freshwater runoff flux R. This strong co-variation pattern emerges predominantly and consistently across the multiple study catchments in different world parts and climates. Co-variations are overall weaker than those of the SM-R relationship for both the blue P-R and the green P-ET flux connections. The blue P-R relationship is stronger than the green P-ET relationship for relatively wet, energy-limited climate conditions, while the latter is stronger only for very dry and very water limited conditions. The role of SM as possible main regulator of green ET flux variations also increases for drier and more water-limited conditions, while the role of P for SM variations increases, and that of SM for R variations decreases for SM measured at smaller depth, closer to the surface.

For the second research question of this study, reanalysis products and ESMs are found to exhibit their greatest errors relative to observation-based datasets for the dominant SM-R co-variation relationship, and the separate SM and R variables. ESM performance assessment for hydrology often focuses on individual variables, while our study shows that individual-variable performance of an ESM may be good while its variable co-variation performance may be poor. Overall, ESM results for large-scale SM-R and P-R co-variations exhibit considerable biases and large differences among the models, which mostly perform better for SM-ET and P-ET co-variations. The findings of this study can contribute to improve modeling capabilities and advance fundamental understanding of essential large-scale hydrological co-variation patterns for water resource security, and water-related hazards and land-atmosphere interactions.

References

Albert, J. S., Destouni, G., Duke-Sylvester, S. M., Magurran, A. E., Oberdorff, T., Reis, R. E., Winemiller, K. O., & Ripple,

W. J. (2020). Scientists' warning to humanity on the freshwater biodiversity crisis. *Ambio*. https://doi.org/10.1007/s13280-020-01318-8

Asokan, S. M., Rogberg, P., Bring, A., Jarsjö, J., & Destouni, G. (2016). Climate model performance and change projection for freshwater fluxes: Comparison for irrigated areas in Central and South Asia. *Journal of Hydrology: Regional Studies*, *5*, 48–65. https://doi.org/10.1016/j.ejrh.2015.11.017

Beaudoing, H. and M. Rodell, N. (2019). *GLDAS Noah Land Surface Model L4 monthly 0.25 x 0.25 degree V2.0*. Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC).

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The Relative Importance of Different Flood-Generating Mechanisms Across Europe. *Water Resources Research*, *55*(6), 4582–4593. https://doi.org/10.1029/2019WR024841

Blöschl, G. (2006). Hydrologic synthesis: Across processes, places, and scales. *Water Resources Research*, *42*(3), 2–4. https://doi.org/10.1029/2005WR004319

Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H. G., Sivapalan, M., Stumpp, C., Toth, E., Volpi, E., Carr, G., Lupton, C., Salinas, J., Széles, B., Viglione, A., Aksoy, H., … Zhang, Y. (2019). Twenty-three unsolved problems in hydrology (UPH)–a community perspective. *Hydrological Sciences Journal*, *64*(10), 1141–1158. https://doi.org/10.1080/02626667.2019.1620507

Bosson, E., Sabel, U., Gustafsson, L. G., Sassner, M., & Destouni, G. (2012). Influences of shifts in climate, landscape, and permafrost on terrestrial hydrology. *Journal of Geophysical Research Atmospheres*, *117*(5), 1–12. https://doi.org/10.1029/2011JD016429

Bring, A., Goldenberg, R., Kalantari, Z., Prieto, C., Ma, Y., Jarsjö, J., & Destouni, G. (2019). Contrasting Hydroclimatic Model-Data Agreements Over the Nordic-Arctic Region. *Earth's Future*, *7*(12), 1270–1282. https://doi.org/10.1029/2019EF001296

Bring, Arvid, Asokan, S. M., Jaramillo, F., Jarsjö, J., Levi, L., Pietroń, J., Prieto, C., Rogberg, P., & Destouni, G. (2015). Implications of freshwater flux data from the CMIP5 multimodel output across a set of Northern Hemisphere drainage basins. *Earth's Future*, *3*(6), 206–217. https://doi.org/10.1002/2014EF000296

Brocca, L., Melone, F., Moramarco, T., & Morbidelli, R. (2010). Spatial-temporal variability of soil moisture and its estimation across scales. *Water Resources Research*, *46*(2), 1–14. https://doi.org/10.1029/2009WR008016

Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., Mackay, D. S., & Maxwell, R. M. (2015). Hydrological partitioning in the critical zone: Recent advances and opportunities for developing transferable understanding of water cycle dynamics. *Water Resources Research*, 1–28. https://doi.org/10.1002/2015WR017096.Received

Copernicicus Climate Change Service (C3S). (2017). *ERA5: Fifth generation of ECMWF atmospheric reanalysis of the global climate. Copernicus Climate Change Service Climate Data Store (CDS)*. Copernicus Climate Change Service Climate Data Store (CDS). https://doi.org/https://cds.climate.copernicus.eu/cdsapp#!/home

Cvetkovic, V., Carstens, C., Selroos, J. O., & Destouni, G. (2012). Water and solute transport along hydrological pathways. *Water Resources Research*, *48*(6), 1–15. https://doi.org/10.1029/2011WR011367

Darracq, A., Destouni, G., Persson, K., Prieto, C., & Jarsjö, J. (2010). Scale and model resolution effects on the distributions of advective solute travel times in catchments. *Hydrological Processes*, *24*(12), 1697–1710. https://doi.org/10.1002/hyp.7588

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., … Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Destouni, G., Asokan, S. M., & Jarsj, J. (2010). Inland hydro-climatic interaction: Effects of human water use on regional climate. *Geophysical Research Letters*, *37*(18), 1–6. https://doi.org/10.1029/2010GL044153

Destouni, G., Jaramillo, F., & Prieto, C. (2013). Hydroclimatic shifts driven by human water use for food and energy production. *Nature Climate Change*, *3*(3), 213–217. https://doi.org/10.1038/nclimate1719

Destouni, G., & Verrot, L. (2014). Screening long-term variability and change of soil moisture in a changing climate. *Journal of Hydrology*, *516*(1), 131–139. https://doi.org/10.1016/j.jhydrol.2014.01.059

Do, H. X., Gudmundsson, L., Leonard, M. & Westra, S. (2018). *PANGAEA, https://doi.org/10.1594/PANGAEA.887477*. PANGAEA. https://doi.org/https://doi.org/10.1594/PANGAEA.887477

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., … Lecomte, P. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, *203*, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

EC-Earth Consortium. (2019). *EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 ScenarioMIP*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.727

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Falkenmark, M., & Rockström, J. (2006). The new blue and green water paradigm: Breaking new ground for water resources planning and management. *Journal of Water Resources Planning and Management*, *132*(3), 129–132. https://doi.org/10.1061/(ASCE)0733-9496(2006)132:3(129)

Fan, Y., & van den Dool, H. (2008). A global monthly land surface air temperature analysis for 1948-present. *Journal of Geophysical Research Atmospheres*, *113*(1), 1–18. https://doi.org/10.1029/2007JD008470

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S.C., Collins, W., et al. (2013). *Evaluation of climate models, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of* *the Intergovernmental Panel on Climate Change*.

Ghajarnia, N., Kalantari, Z., Orth, R., & Destouni, G. (2020). Close co-variation between soil moisture and runoff emerging from multi-catchment data across Europe. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-61621-y

Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(12), 6780–6792. https://doi.org/10.1109/TGRS.2017.2734070

Gruber, A., Scanlon, T., Van Der Schalie, R., Wagner, W., & Dorigo, W. (2019). Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. *Earth System Science Data*, *11*(2), 717–739. https://doi.org/10.5194/essd-11-717-2019

Gudmundsson, L., Do, H. X., Leonard, M. & Westra, S. (2018). *The Global Streamflow Indices and Metadata Archive (GSIM) - Part 2: Time Series Indices and Homogeneity Assessment*. PANGAEA. https://doi.org/https://doi.org/10.1594/PANGAEA.887470

Gudmundsson, L., Rego, F. C., Rocha, M., & Seneviratne, S. I. (2014). Predicting above normal wildfire activity in southern Europe as a function of meteorological drought. *Environmental Research Letters*, *9*(8). https://doi.org/10.1088/1748-9326/9/8/084008

Hawley, M. E., Galloway, J. N., Keene, W. C. (1988). Standard error calculations for non-seasalt constituents in marine precipitation. *Water Air Soil Pollution*, *42*(2), 87–102. https://doi.org/doi:10.1007/BF00282393

Kalantari, Z., Ferreira, C. S. S., Koutsouris, A. J., Ahmer, A. K., Cerdà, A., & Destouni, G. (2019). Assessing flood probability for transportation infrastructure based on catchment characteristics, sediment connectivity and remotely sensed soil moisture. *Science of the Total Environment*, *661*, 393–406. https://doi.org/10.1016/j.scitotenv.2019.01.009

Khatami, S., Peel, M. C., Peterson, T. J., & Western, A. W. (2019). Equifinality and Flux Mapping: A New Approach to Model Evaluation and Process Representation Under Uncertainty. *Water Resources Research*, *55*(11), 8922–8941. https://doi.org/10.1029/2018WR023750

Koohi, S., Azizian, A., Brocca, L. (2019). Evaluating the efficiency of earth2Observe reanalysis models and VIC-3L for estimation of runoff. *Journal of Water and Soil Resources Conservation*, *8*(4), 118–131.

Koutsouris, A. J., Destouni, G., Jarsjö, J., & Lyon, S. W. (2010). Hydro-climatic trends and water resource management implications based on multi-scale data for the Lake Victoria region, Kenya. *Environmental Research Letters*, *5*(3). https://doi.org/10.1088/1748-9326/5/3/034005

Langbein, W. B. (1949). Annual runoff in the United States. *Geological Survey Circular*, *52*. https://pubs.er.usgs.gov/publication/cir52

Madani, K., & Lund, J. R. (2010). Estimated impacts of climate warming on California's high-elevation hydropower. *Climatic Change*, *102*(3), 521–538. https://doi.org/10.1007/s10584-009-9750-8

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., & Verhoest, N. E. C. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture.

Geoscientific Model Development, *10*(5), 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Mastrotheodoros, T., Pappas, C., Molnar, P., Burlando, P., Manoli, G., Parajka, J., Rigon, R., Szeles, B., Bottazzi, M., Hadjidoukas, P., & Fatichi, S. (2020). More green and less blue water in the Alps during warmer summers. *Nature Climate Change*, *10*(2), 155–161. https://doi.org/10.1038/s41558-019-0676-5

Maxwell, R. M., & Condon, L. E. (2016). Connections between groundwater flow and transpiration partitioning. *Science*, *353*(6297), 377–380. https://doi.org/10.1126/science.aaf7891

Maxwell, R. M., & Kollet, S. J. (2008). Interdependence of groundwater dynamics and land-energy feedbacks under climate change. *Nature Geoscience*, *1*(10), 665–669. https://doi.org/10.1038/ngeo315

Medellín-Azuara, J., Harou, J. J., Olivares, M. A., Madani, K., Lund, J. R., Howitt, R. E., Tanaka, S. K., Jenkins, M. W., & Zhu, T. (2007). Adaptability and adaptations of California's water supply system to dry climate warming. *Climatic Change*, *87*(1 SUPPL). https://doi.org/10.1007/s10584-007-9355-z

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, *15*(2), 453–469. https://doi.org/10.5194/hess-15-453-2011

Moshir Panahi, D., Kalantari, Z., Ghajarnia, N., Seifollahi-Aghmiuni, S., & Destouni, G. (2020). Variability and change in the hydro-climate and water resources of Iran over a recent 30-year period. *Scientific Reports*, *10*(1), 1–9. https://doi.org/10.1038/s41598-020-64089-y

Naumann, G., Spinoni, J., Vogt, J. V., & Barbosa, P. (2015). Assessment of drought damages and their uncertainties in Europe. *Environmental Research Letters*, *10*(12). https://doi.org/10.1088/1748-9326/10/12/124013

Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stoll, J., Folini, D. S., Tegen, I., Wieners, K.-H., Mauritsen, T., Stemmler, I., Barthel, S., Bey, I., Daskalakis, N., Heinold, B., Kokkola, H., Partridge, D., Rast, S., Schmidt, H., Schutgens, N., Stanelle, T., … Lohmann, U. (2019). *HAMMOZ-Consortium MPI-ESM1.2-HAM model output prepared for CMIP6 AerChemMIP*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.1621

Orth, R., & Destouni, G. (2018). Drought reduces blue-water fluxes more strongly than green-water fluxes in Europe. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06013-7

Orth, R., Destouni, G., Jung, M., & Reichstein, M. (2020). Large-scale biospheric drought response intensifies linearly with drought duration in arid regions. *Biogeosciences*, *17*(9), 2647–2656. https://doi.org/10.5194/bg-17-2647-2020

Raymond, C., Horton, R. M., Zscheischler, J., Martius, O., AghaKouchak, A., Balch, J., Bowen, S. G., Camargo, S. J., Hess, J., Kornhuber, K., Oppenheimer, M., Ruane, A. C., Wahl, T., & White, K. (2020). Understanding and managing connected extreme events. *Nature Climate Change*, *10*(7), 611–621. https://doi.org/10.1038/s41558-020-0790-4

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., & Toll, D. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Saft, M., M. C. Peel, A. W. Western, and L. Z. (2016). Predicting shifts in rainfall-runoff partitioning during multiyear drought: Roles of dry period and catchment characteristics. *Water Resources Research*, *52*, 9290–9305. https://doi.org/doi:10.1002/2016WR019525

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf, and M. Z. (2016). *GPCC Full Data Reanalysis Version 7.0: Monthly Land-Surface Precipitation from Rain Gauges built on GTS based and Historic Data*. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/https://doi.org/10.5065/D6000072

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010). Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3–4), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Swart, N. C., Cole, J. N. S., Kharin, V. V, Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D., … Sigmond, M. (2019). *CCCma CanESM5 model output prepared for CMIP6 ScenarioMIP*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.1317

Tatebe, H., & Watanabe, M. (2018). *MIROC MIROC6 model output prepared for CMIP6 CMIP abrupt-4xCO2*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.5411

Thirel, G., Andréassian, V., & Perrin, C. (2015). De la nécessité de tester les modèles hydrologiques sous des conditions changeantes. *Hydrological Sciences Journal*, *60*(7–8), 1165–1173. https://doi.org/10.1080/02626667.2015.1050027

Thorslund, J., Jarsjo, J., Jaramillo, F., Jawitz, J. W., Manzoni, S., Basu, N. B., Chalov, S. R., Cohen, M. J., Creed, I. F., Goldenberg, R., Hylin, A., Kalantari, Z., Koussis, A. D., Lyon, S. W., Mazi, K., Mard, J., Persson, K., Pietro, J., Prieto, C., … Destouni, G. (2017). Wetlands as large-scale nature-based solutions: Status and challenges for research, engineering and management. *Ecological Engineering*, *108,* 489–497. https://doi.org/10.1016/j.ecoleng.2017.07.012

Törnqvist, R., Jarsjö, J., Pietroń, J., Bring, A., Rogberg, P., Asokan, S. M., & Destouni, G. (2014). Evolution of the hydro-climate system in the Lake Baikal basin. *Journal of Hydrology*, *519*(PB), 1953–1962. https://doi.org/10.1016/j.jhydrol.2014.09.074

Van Loon, A. F. (2015). Hydrological drought explained. *Wiley Interdisciplinary Reviews: Water*, *2*(4), 359–392. https://doi.org/10.1002/wat2.1085

Vaze, J., Post, D. A., Chiew, F. H. S., Perraud, J. M., Viney, N. R., & Teng, J. (2010). Climate non-stationarity - Validity of calibrated rainfall-runoff models for use in climate change studies. *Journal of Hydrology*, *394*(3–4), 447–457. https://doi.org/10.1016/j.jhydrol.2010.09.018

Woldemeskel, F. M., Sharma, A., Sivakumar, B., & Mehrotra, R. (2012). An error estimation method for precipitation and temperature projections for future climates. *Journal of Geophysical Research Atmospheres*, *117*(22), 1–13. https://doi.org/10.1029/2012JD018062

Wu, T., Chu, M., Dong, M., Fang, Y., Jie, W., Li, J., Li, W., Liu, Q., Shi, X., Xin, X., Yan, J., Zhang, F., Zhang, J., Zhang, L., & Zhang, Y. (2018). *BCC BCC-CSM2MR model output prepared for CMIP6 CMIP piControl*. Earth System Grid

Federation. https://doi.org/10.22033/ESGF/CMIP6.3016

Dee, D. P. *et al.* The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597 (2011).

Koohi, S., Azizian, A., Brocca, L. Evaluating the efficiency of earth2Observe reanalysis models and VIC-3L for estimation of runoff. *J. Water Soil Resour. Conserv.* **8**, 118–131 (2019).

Hawley, M. E., Galloway, J. N., Keene, W. C. Standard error calculations for non-seasalt constituents in marine precipitation. *Water Air Soil Pollut.* **42**, 87–102 (1988).

Rochford, P. PeterRochford/SkillMetricsToolbox. (2020).

Data availability

Soil moisture and ET data from the GLEAM-3.3a model are available at https://www.gleam.eu/. GPCC-V7 precipitation data are available from the NOAA website https://www.esrl.noaa.gov/psd/data/gridded/data.gpcc.html. Temperature data from GHCN_CAMS are available at https://www.esrl.noaa.gov/psd/data/gridded/data.ghcncams.html, and GSIM streamflow and metadata at https://doi.pangaea.de/10.1594/PANGAEA.887470 and https://doi.pangaea.de/10.1594/PANGAEA.887477. ERA5 data are generated using Copernicus Climate Change Service Information (2019) and available at https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5. GLDAS data are generated by National Aeronautics and Space Administration Goddard Earth Science Data Information and Services Center (GES DISC) and available at https://disc.gsfc.nasa.gov/datasets/GLDAS_NOAH025_M_2.0/summary?keywords=GLDAS. All CMIP6 climate model data are generated as part of the internationally-coordinated Coupled Model Intercomparison Project Phase 6 (CMIP6) and are available at https://pcmdi.llnl.gov/CMIP6/. The ESA-CCI SM-v04.5 soil moisture data is generated by the European Space Agency (ESA) and is available at http://www.esa-soilmoisture-cci.org/node/145.

Author contribution

All authors have contributed substantially to this work as specified below: N.G. compiled the data, created all Figures and Tables, and was responsible for analysing the data and writing the paper. Z.K. contributed to the result analysis and the writing. G.D. conceived and led the study and contributed to the result analysis and the writing. All authors contributed to the development of the analysis approach and discussion of the content.

Supplementary Information: Supporting Figure 1-9

Competing interests: The authors declare no conflict of interest