https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Open Access Earth System Science Data Discussions

GRIT-ADB: A Global Attribute Database for the GRIT Hydrography

Boen Zhang[1,6], Michel Wortmann[2,1], Yinxue Liu[1,3], Frederik Kratzert[4], Simon Moulds[5], and Louise Slater[1]

[1]School of Geography and the Environment, University of Oxford, Oxford, UK
[2]European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, UK
[3]Geography and Environment, Loughborough University, Loughborough, UK
[4]Google Research, Vienna, Austria
[5]School of Geosciences, University of Edinburgh, Edinburgh, UK
[6]School of Geography and Planning, Sun Yat-sen University, Guangzhou, China

Correspondence: Boen Zhang (boen.zhang@ouce.ox.ac.uk) and Louise Slater (louise.slater@ouce.ox.ac.uk)

Abstract. Global hydro-environmental databases provide essential information for large-scale hydrological, ecological, geomorphological, and Earth system analyses. Most existing global databases are built upon convergent river representations that do not explicitly capture bifurcating, multi-channel river systems. In addition, these databases primarily characterise long-term climatological means or static representations of environmental conditions derived from earlier-generation global datasets, limiting their applicability for time-varying analyses of hydroclimatic and geomorphological processes. Here we present GRIT-ADB, a new attribute database for the vectorised Global River Topology (GRIT), a new river hydrography dataset created from a 30 m resolution river mask and terrain data that provides a topology-explicit and physically realistic representation of river networks including divergent flow pathways. GRIT-ADB provides standardised hydro-environmental information for 19.6 million km of rivers and streams. It currently comprises 64 time-varying (multi-dimensional embeddings) and 35 static variables (>300 attributes), spanning five categories: hydrology, physiography, climate, land cover and use, and soils and geology. Attributes are derived by aggregating and harmonising data from state-of-the-art global datasets and are accumulated along the river network from headwaters to basin outlets, while preserving the topology of divergent and complex flow pathways. The attributes are linked to multiple GRIT scales, including hierarchically-nested subbasins, individual river reaches of up to 1 km long, and coarser-scale river segments of several kilometres long, providing a flexible framework that can accommodate future extensions of GRIT and additional attributes. By combining a standardised attribute framework with explicit representation of bifurcating river hydrography, GRIT-ADB enables improved large-scale yet high-resolution analyses of river connectivity, hydrological extremes, hydro-ecological processes, and climate impacts in complex river systems, supporting a wide range of global hydrological and environmental applications.

1 Introduction

In river basins worldwide, shifting hydroclimatic extremes, increasing hydroclimate variability, and growing human pressures are driving the development of innovative strategies for water resources assessment, planning, and risk management (e.g.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Mullin, 2020; Rosa and Sangiorgio, 2025). These approaches are increasingly underpinned by data-driven analyses of static and dynamic hydro-environmental variables, enabled by globally consistent datasets of static and dynamic catchment attributes that provide homogeneous information across political boundaries and support hydrologic analysis in data-scarce regions (Linke et al., 2019; Jones et al., 2025; Frame et al., 2022).

Such datasets (e.g. HydroATLAS; Linke et al., 2019) depend on river hydrography products that provide a topology-aware representation of the global river network, linking individual river reaches and segments to their contributing catchment areas. Regional examples include NHD/NHDPlus (Moore et al., 2019) and EU-Hydro (EEA, 2025), while global datasets include MERIT-Hydro (Yamazaki et al., 2019), HydroSHEDS (Lehner et al., 2008), SWORD (Altenau et al., 2021), and Hydrography90m (Amatulli et al., 2022). The global hydro-environmental attribute database HydroATLAS is built upon HydroSHEDS while Environment90m (García-Márquez et al., 2026) is built upon Hydrography90m (Amatulli et al., 2022). These datasets offer physically consistent representations of river topology, enabling network-based hydrological analysis across both gauged and ungauged regions. Together, these hydrographic and attribute datasets have substantially advanced large-scale hydrologic research and applications by providing standardised, spatially explicit foundations for modelling and analysis.

Notwithstanding their advances, the quality, timeliness and spatial resolution of existing hydro-environmental attribute databases are limited by various factors. First, attribute coverage is uneven, and some key variables are derived from earlier global products or long-term climatological estimates. These inputs often represent conditions from past decades, which may not fully capture present-day hydro-environmental regimes under ongoing global change. For example, discharge estimates in HydroATLAS are based on 1971-2000 simulations from the 0.5-degree WaterGAP model (Linke et al., 2019), while precipitation estimates in NHDPlus are derived for the same 1971-2000 period (Moore et al., 2019). Such temporal aggregation constrains the representation of present-day variability and extremes.

The second key limitation relates to the assumption of strictly convergent river flow inherent to most river hydrography datasets, including EU-Hydro, HydroRIVERS (Lehner and Grill, 2013) and MERIT-Hydro. This simplification breaks down in numerous basins, especially in lowland floodplains and deltaic environments, where flow bifurcation and distributary systems are common. Consequently, widely used hydro-environmental datasets inherit a convergence-only structure, limiting their ability to describe flow regimes in lowland floodplains and deltaic environments. To address this limitation, we previously developed the Global River Topology (GRIT) dataset (Wortmann et al., 2025c), which provides a vectorised branching river network representing 19.6 million km of streams and rivers with drainage area greater than 50 km^2 and over 67,000 bifurcations. GRIT is constructed by combining a 30 m Landsat-derived river mask with elevation-generated stream networks and by adopting the 30 m FABDEM digital terrain model (Hawker et al., 2022) in place of conventional SRTM derivatives, enabling a detailed representation of river-system complexity at the planetary scale.

To address the structural and attribute limitations of existing hydrological databases, we introduce GRIT-ADB, a new global hydro-environmental attribute database built on the vectorised GRIT hydrography. GRIT-ADB provides a single, comprehensive, consistently organised and fully-global database comprising 35 static and 64 time-varying hydro-environmental attributes at segment and reach scales (Table 1). The attributes are derived from publicly available state-of-the-art datasets and grouped into five categories: hydrology, physiography, climate, land cover and land use, and soils and geology. Together, these features

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Open Access Earth System Science Data Discussions

establish GRIT-ADB as a topology-aware and globally-consistent foundation for large-scale hydrological analysis across both gauged and ungauged river systems.

Table 1: Hydro-environmental attributes provided in GRIT-ADB. "Res. (G/V)" indicates data type: gridded (G; spatial resolution given) or vector (V). Five categories of attributes are included in GRIT-ADB. The "Year" column indicates the period of data used for calculation. All attributes are available at the spatial scales of both river reach and river segment.

Category	Variable	Source data	Res. (G/V)	Year	Temporal type	Citation
Hydrology	Reservoir volume	GRanD v1.3	V	most recent	static	Lehner et al. (2011)
Hydrology	Lake area	GLAKES	V	1984–2019	static	Pi et al. (2022)
Hydrology	Drainage area	GRIT	V	most recent	static	Wortmann et al. (2025c)
Hydrology	Number of bifurcations	GRIT	V	most recent	static	Wortmann et al. (2025c)
Hydrology	River width	GRWL	V	1984–2015	static	Allen and Pavelsky (2018)
Soils & geology	Permeability	GLHYMPS v2	V	1965-2018	static	Huscroft et al. (2018)
Soils & geology	Porosity	GLHYMPS v2	V	1965-2018	static	Huscroft et al. (2018)
Soils & geology	Absolute depth to bedrock	DTB	G: 250 m	1965-2016	static	Shangguan et al. (2017)
Soils & geology	Sand fraction in soil	SoilGrids v2	G: 250 m	most recent	static	Poggio et al. (2021)
Soils & geology	Clay fraction in soil	SoilGrids v2	G: 250 m	most recent	static	Poggio et al. (2021)
Soils & geology	Silt fraction in soil	SoilGrids v2	G: 250 m	most recent	static	Poggio et al. (2021)
Soils & geology	Bulk density	SoilGrids v2	G: 250 m	most recent	static	Poggio et al. (2021)
Soils & geology	Cation exchange capacity	SoilGrids v2	G: 250 m	most recent	static	Poggio et al. (2021)
Soils & geology	Coarse fragments	SoilGrids v2	G: 250 m	most recent	static	Poggio et al. (2021)
Soils & geology	Nitrogen	SoilGrids v2	G: 250 m	most recent	static	Poggio et al. (2021)
Soils & geology	pH	SoilGrids v2	G: 250 m	most recent	static	Poggio et al. (2021)
Soils & geology	Soil organic carbon	SoilGrids v2	G: 250 m	most recent	static	Poggio et al. (2021)
Climate	Aridity index	Global-AI_PET_v3	G: 30 arcseconds	1970–2000	climatology	Zomer et al. (2022)
Climate	PET	Global-AI_PET_v3	G: 30 arcseconds	1970–2000	climatology	Zomer et al. (2022)
Climate	Mean air temperature	MSWX	G: 0.1 degree	1991-2020	climatology	Beck et al. (2022)
Climate	Air temperature seasonality	MSWX	G: 0.1 degree	1991-2020	Climatology	(Beck et al., 2022)
Climate	Mean annual precipitation	MSWEP v2	G: 0.1 degree	1991-2020	climatology	Beck et al. (2019)
Climate	Precipitation seasonality	MSWEP v2	G: 0.1 degree	1991-2020	climatology	Woods (2009)
Climate	Climate zones	Köppen–Geiger zones	G: 30 arcseconds	1991-2020	climatology	Beck et al. (2018)
Land cover	Satellite embedding	AlphaEarth	G: 10 m	2017-2024	time series	Brown et al. (2025)
Land cover	Urban area extent	ESA CCI LC	G: 300 m	2020	static	C3S (2019)
Land cover	Cropland extent	ESA CCI LC	G: 300 m	2020	static	C3S (2019)
Land cover	Forest cover extent	ESA CCI LC	G: 300 m	2020	static	C3S (2019)
Land cover	Water body extent	ESA CCI LC	G: 300 m	2020	static	C3S (2019)
Land cover	Grassland extent	ESA CCI LC	G: 300 m	2020	static	C3S (2019)

Continued on next page

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Open Access Earth System Science Data Discussions

Table 1 (continued)

Category	Variable	Source data	Res. (G/V)	Year	Temporal type	Citation
Physiography	Elevation	FABDEM	G: 30 m	most recent	static	Hawker et al. (2022)
Physiography	Stream length	GRIT	V	most recent	static	Wortmann et al. (2025c)
Physiography	Stream gradient	FABDEM	G: 30 m	most recent	static	Hawker et al. (2022)
Physiography	Strahler order	GRIT	V	most recent	static	Wortmann et al. (2025c)
Physiography	Sinuosity	GRIT	V	most recent	static	Wortmann et al. (2025c)
Physiography	Azimuth	GRIT	V	most recent	static	Wortmann et al. (2025c)

2 Methods

The spatial sub-basin and river reach geometries used in GRIT-ADB are derived from the Global River Topology (GRIT) hydrography (Wortmann et al., 2025c). GRIT comprises three primary vector datasets, including network segments, network reaches and catchments. Network segments are conceptually consistent with those in conventional river networks (Figure 1a). Each segment is subdivided into equal-length reaches between segment nodes, each shorter than 1 km (Figure 1b), enabling the extraction of hydro-environmental attributes for every 1 km reach of the GRIT network. GRIT (the underlying hydrography dataset) provides the geometric and topological structure of the river network, but it does not include external hydro-environmental variables. The only attributes available within GRIT are those derived directly from the line geometry and network topology (connectivity of the network), such as stream length, sinuosity, azimuth, and Strahler order. GRIT also inherits river width attributes from GRWL, which is a key input to the GRIT methodology (Wortmann et al., 2025c).

To link hydro-environmental attributes to the GRIT hydrography, river reaches and their catchment boundaries are used as the fundamental spatial units. Segment-scale estimates are obtained by aggregating reach-scale statistics. Given the approximately 20 million river reaches and associated catchment boundaries, the attributes are extracted for 61 geographic domains (Figure A1) in parallel. We perform the extraction and aggregation of summary statistics in the Equal Earth projection. Source data are first clipped to the extent of each domain and re-projected to have the same projection as the GRIT vector dataset. After computing local statistics, the results are appended to the vectorised sub-basin polygons and river reaches via their unique identifier codes (IDs).

The hydro-environmental attributes are drawn from global state-of-the-art raster and vector datasets (Table 1). These datasets have complete global coverage, high spatial resolution, documented precision, and are being actively maintained and updated. We only consider datasets that are free and publicly available, and available under licenses that give permission to use and distribute derivatives. Attributes in GRIT-ADB span five categories: hydrology, physiography, climate, land cover and land use, and soils and geology. In addition to the hydrology attributes computed directly from GRIT and GRWL, we also calculate the contributing lake area and reservoir volume for each river reach based on GRanD v1.3 and GLAKES datasets (Lehner et al., 2011; Pi et al., 2022). The GRanD dataset contains over 7,300 large dams and reservoir polygons with associated attributes such as storage capacity, compiled from multiple national and international dam inventories. The GLAKES dataset provides

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Open Access Earth System Science Data Discussions

a global inventory of lakes with associated geometric and temporal information, enabling quantification of upstream lake area contributions to each reach. Soil attributes are drawn from the SoilGrids v2 dataset (Poggio et al., 2021). Precipitation and temperature data are from the Multi-Source Weighted-Ensemble Precipitation (MSWEP) v2 and Multi-Source Weather (MSWX) datasets (Beck et al., 2019, 2022). The GLobal HYdrogeology MaPS (GLHYMPS) dataset is used to extract the geological attributes (Huscroft et al., 2018). Land cover distribution for 2020 is taken from the corresponding ESA CCI Land Cover dataset (C3S, 2019).

To capture temporal dynamics in land cover, we additionally use the AlphaEarth Satellite Embedding dataset (Brown et al., 2025). This global, analysis-ready product is generated by Google DeepMind's AlphaEarth Foundations model and provides annual 64-dimensional embedding vectors at 10 m spatial resolution from 2017 to present. The embeddings are derived from multi-sensor Earth observation data, including optical and radar imagery, and encode the spectral, spatial, and temporal properties of each location, facilitating large-scale land-cover mapping, change detection, and similarity analysis without direct reliance on raw satellite imagery.

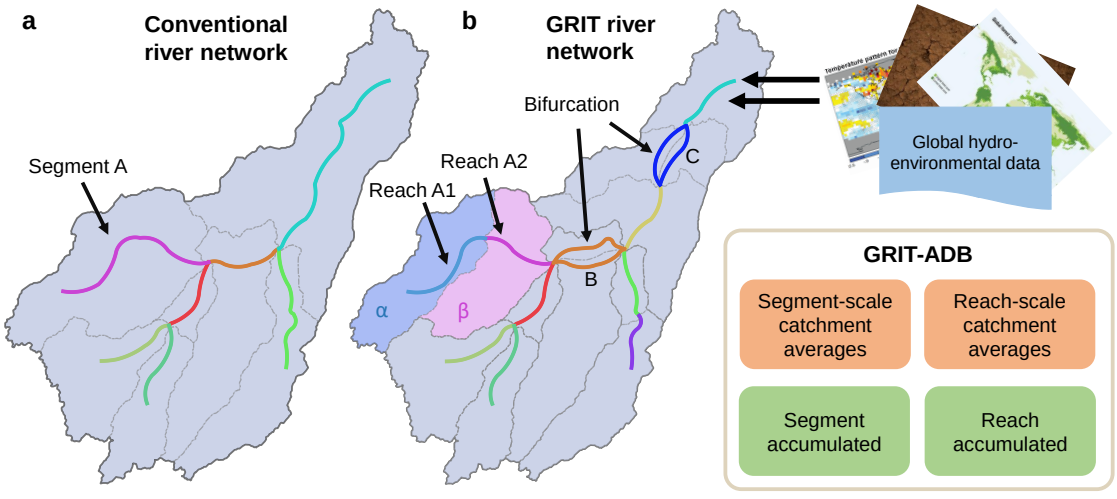


Figure 1. Conceptual diagram of GRIT-ADB. Unlike conventional river networks (a), GRIT subdivides segments (e.g. Segment A) into equal-length reaches (e.g. Reach A1 and Reach A2), each shorter than 1 km (b). GRIT explicitly represents bifurcations, illustrated by reaches B and C, where flow may split into multiple branches. Hydro-environmental attributes extracted from global datasets are summarised at multiple spatial scales in GRIT-ADB, including reach- and segment-scale catchment averages, and upstream-accumulated statistics for both reaches and segments. The catchment average for Reach A2 uses only area α. The upstream-accumulated statistic includes both areas α and β.

GRIT-ADB provides local and upstream attribute statistics at reach and segment levels. Depending on the spatial characteristics of the variable, local statistics are computed either for the midpoint of each reach or over the local contributing area of each catchment using area-weighted aggregation. The midpoint-based extraction is implemented using the nearest-neighbour method, while catchment-based aggregation is performed using the exactextract Python package. Upstream statistics are computed by accumulating local statistics along the GRIT river network, combining the contributions from all upstream reaches

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

or catchments with drainage area-weighted aggregation. The accumulation statistic depends on the variable, and these details, along with their units and scaling factors, are summarised in Table A1. For example, some attributes are aggregated as area-weighted means over the entire sub-basin (such as mean air temperature), whereas others are summed (such as total lake area). Because global coastline definitions differ among source datasets, slight mismatches occur between the land extents of input attribute datasets and the GRIT land mask. Missing attributes for coastal reaches/subbasins are filled using the nearest downstream or upstream value. For river reaches that bifurcate into multiple downstream branches, river widths from GRWL (Allen and Pavelsky, 2018) are used as partitioning weights in sum aggregations, such that the distributed values are proportional to downstream channel width.

3 Dataset overview

GRIT-ADB provides averages computed both within catchments and along river reaches, and for each spatial unit reports statistics at local and upstream scales. For example, Figure 2 shows upstream-mean fractional sand in soil and local-mean subsurface permeability for each river reach (Fig. 2a) and catchment (Fig. 2b) in GRIT-ADB. The inset map in the top panel shows that rivers in the downstream Amazon River Basin are substantially bifurcating, which is explicitly represented in the GRIT hydrography with corresponding attributes provided for these bifurcating rivers in GRIT-ADB. In the following section we provide an initial exploration of GRIT-ADB, focusing on global river physiography, lake connectivity, and the relationship between river characteristics and hydroclimatology.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Open Access Earth System Science Data Discussions

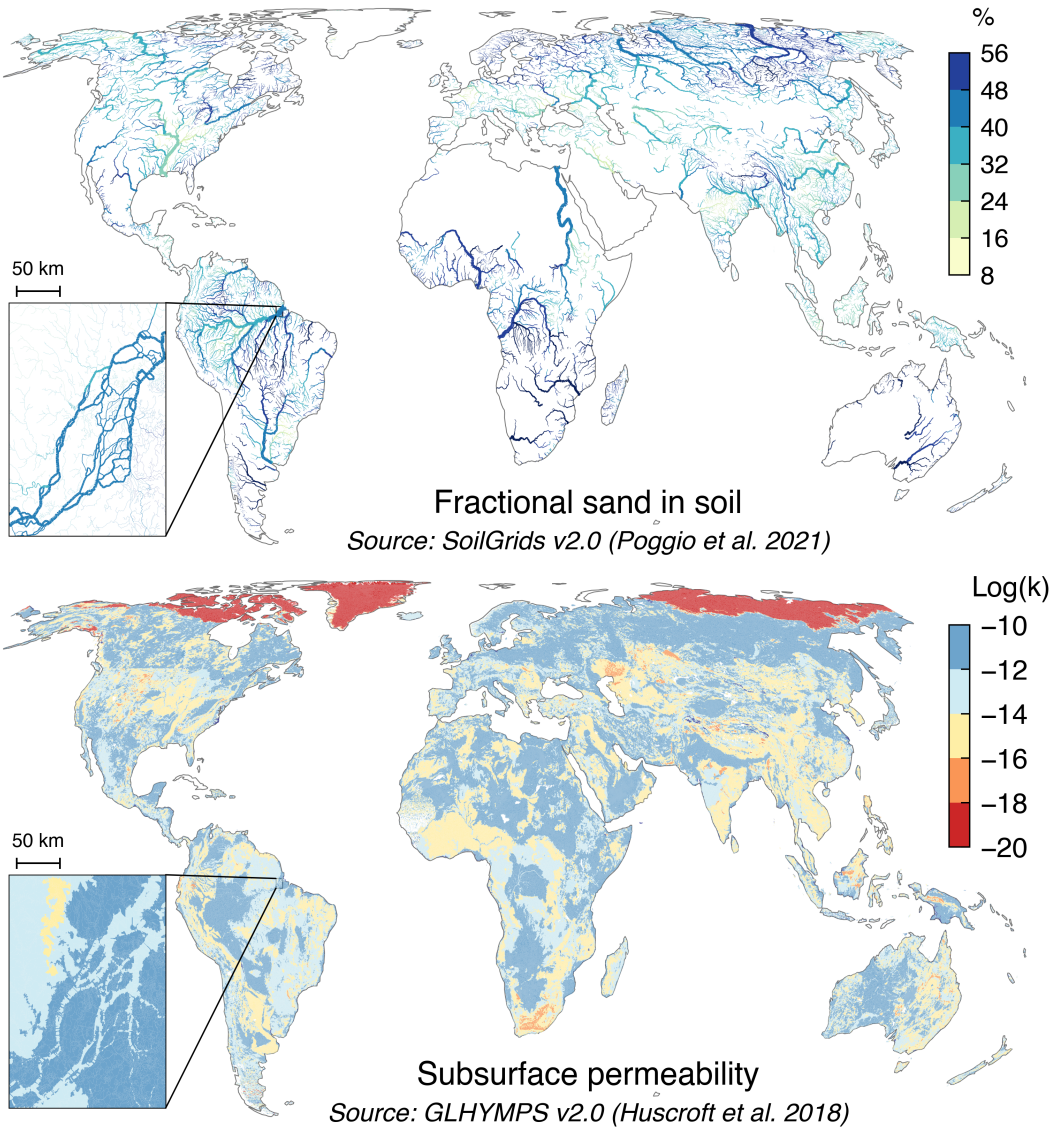


Figure 2. Example geological attributes of GRIT-ADB. Top panel: upstream mean fractional sand in soil per river reach. Bottom panel: local mean subsurface permeability per sub-basin. Local mean estimates are computed as area-weighted averages based on the fraction of each grid cell covered. Upstream mean values are calculated by propagating local-mean estimates downstream through the GRIT river network and computing drainage-area–weighted averages for each reach.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Open Access Earth System Science Data Discussions

3.1 Global river physiography

GRIT-ADB provides the physiographical characteristics of global rivers, including sinuosity, river width, and stream gradient. Although considerable effort has been devoted to understanding river morphology and its relation to other river attributes (Altenau et al., 2021; Frasson et al., 2019), most previous studies did not explicitly account for river bifurcations. By explicitly representing river bifurcations, GRIT-ADB offers the opportunity to address this knowledge gap. Here we present a simple analysis of global relationships among river width, stream gradient, catchment area, and sinuosity based on GRIT-ADB (Figure 3). We find that as catchment area and stream gradient increase, the range of possible sinuosity values decreases (Figure 3a and 3b). Stream gradient decreases with increasing river width (Figure 3c).

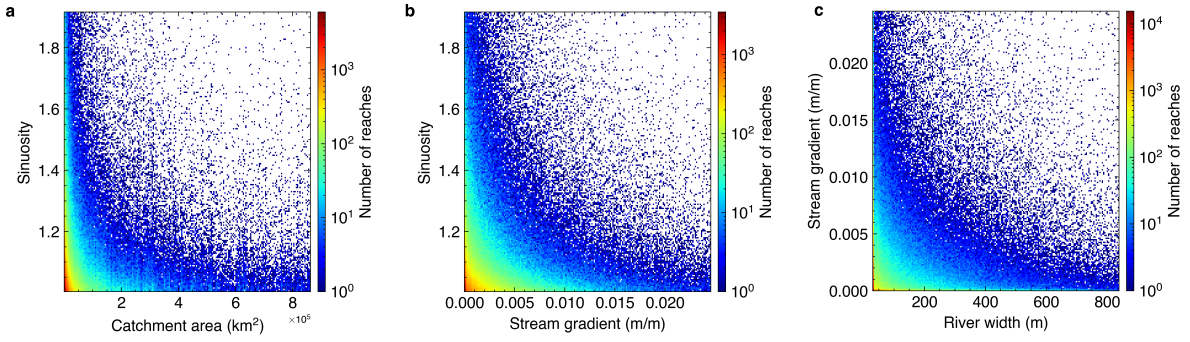


Figure 3. Global relationships between river attributes in GRIT-ADB. All panels are generated using global river reaches outside inland lakes with GRWL widths greater than 30 m, as well as their downstream segments. Colours represent the number of river reaches per bin. Panels show (a) catchment area versus sinuosity, (b) stream gradient versus sinuosity, (c) river width versus stream gradient.

3.2 Lake connectivity

Hydrological attributes are essential for characterising river behaviour. Figure 4 provides illustrative examples of hydrological attributes in GRIT-ADB. Panel (a) shows upstream accumulated lake area for each river reach, normalised by contributing drainage area. Across the global river network, the mean upstream lake area fraction is 1.2%, reflecting generally limited but spatially heterogeneous lake coverage within contributing drainage areas. Some river basins are hydrologically connected to more lake areas, such as the Ob and Yenisei River Basins in Siberia, the Nile River Basin, and the Congo River Basin, as well as most rivers in northwestern North America. Comparison of the inset maps (Figure 4a1 and 4a2) indicates that the Congo River Basin is more strongly influenced by its upstream lake areas compared to the Amazon River Basin. Figure 4b presents the number of upstream river bifurcations for each river reach, also normalised by contributing drainage area for spatial comparability. Results show that the Ganges-Brahmaputra Delta and several rivers in northeastern Siberia exhibit a high number of upstream river bifurcations. In the Ganges–Brahmaputra Delta (Figure 4b2), the Brahmaputra River exhibits more upstream bifurcations per unit drainage area compared to the Ganges River, indicating a more highly branched upstream network.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

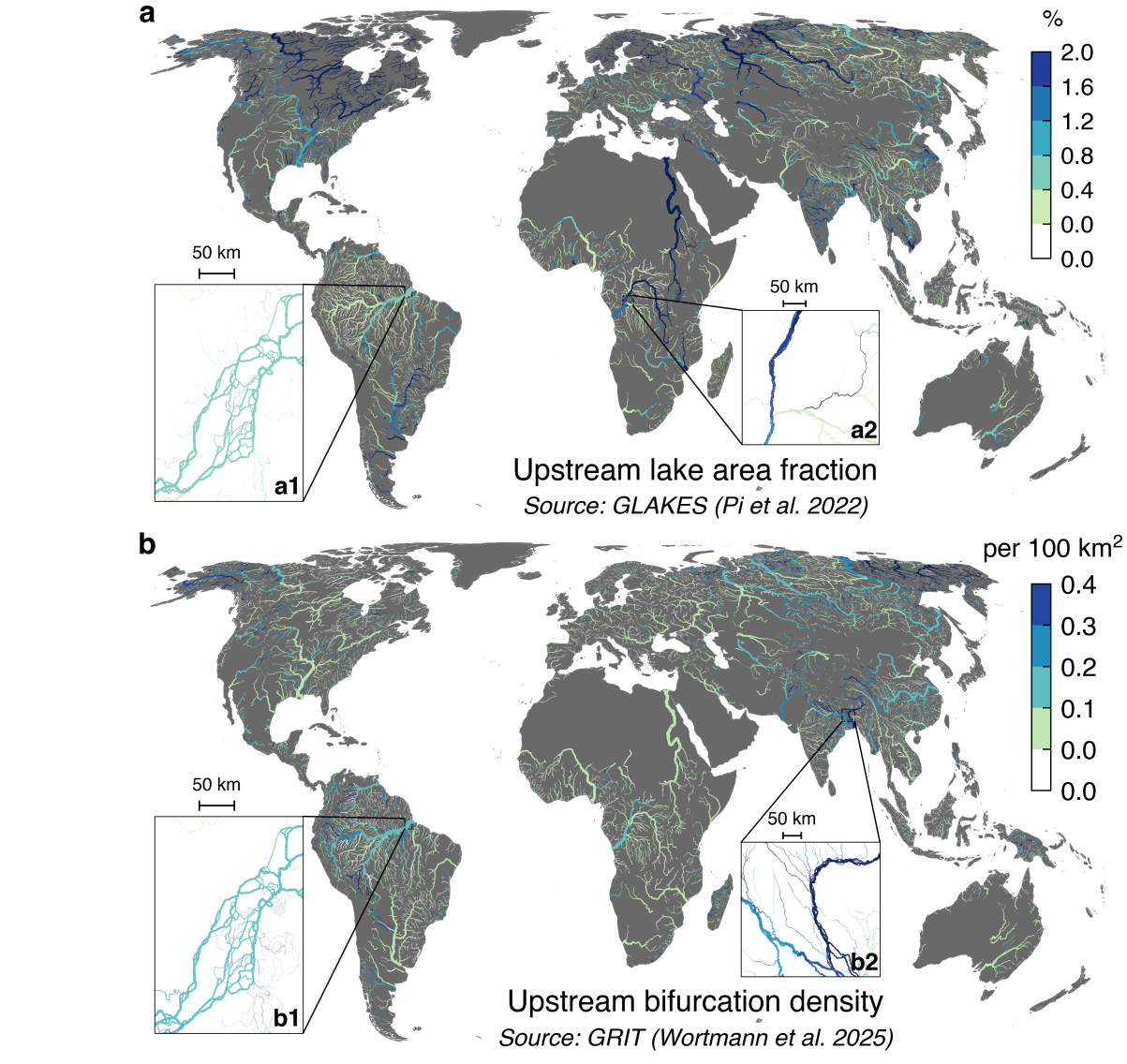
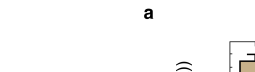


Figure 4. Example hydrological attributes of GRIT-ADB. Top panel: percentage of total upstream lake area relative to the contributing drainage area for each river reach. Bottom panel: total number of upstream river bifurcations per 100 km^2 of upstream drainage area.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026



Open Access Earth System Science Data Discussions

3.3 Relationship with hydroclimatology

We also examined river attributes across climate regions (Figure 5). Rivers in polar regions have substantially lower permeability compared to other climate regions, while variability of permeability is greatest in the cold region (Figure 5a). Soil organic carbon is highest in river basins within the cold region (Figure 5b). Compared to other climate regions, a larger proportion of the drainage network in cold regions is hydrologically connected to lakes, consistent with glacially influenced, low-relief landscapes characteristic of these regions that favour lake formation and hydrological regulation (Figure 5d). River basins in arid and tropical regions exhibit greater depth to bedrock compared to other climate regions (Figure 5e). No visible differences across climate regions are observed for stream gradient or sand fraction in soil (Figure 5c and 5f). Additional comparisons of river attributes are shown in the Supplementary Material.

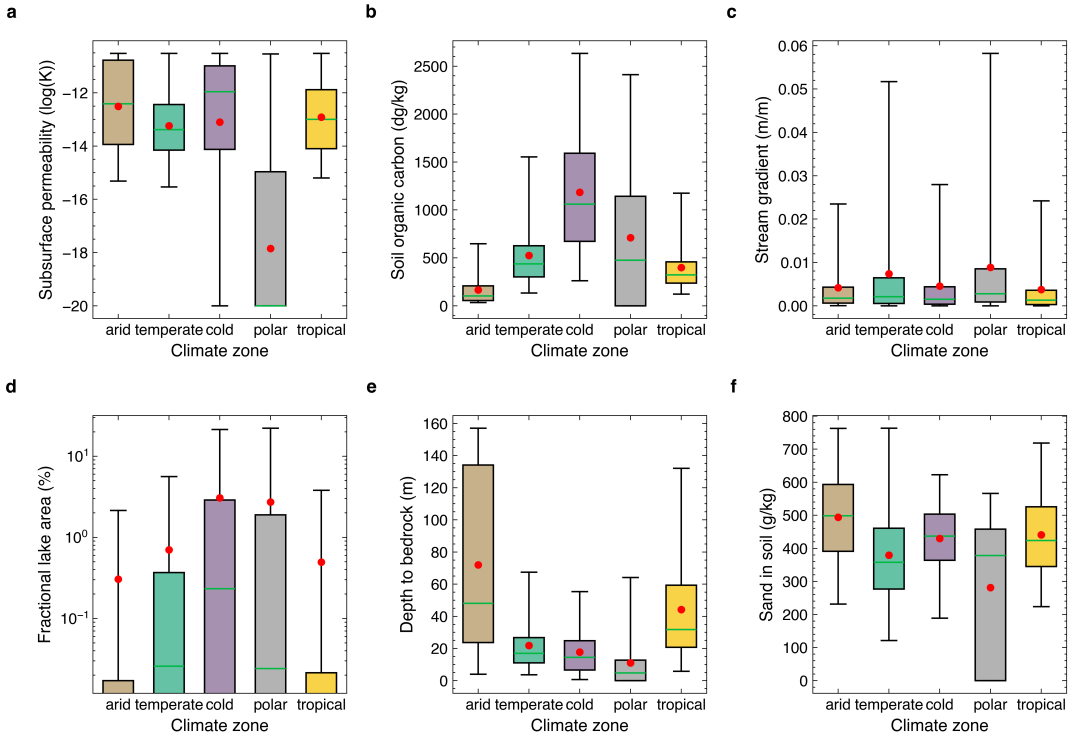


Figure 5. Boxplots of example attributes of GRIT-ADB across climate regions. Stream gradient in panel (c) is a local statistic; all other attributes are upstream statistics derived by routing through the GRIT river network. Panels (a), (b), (e), and (f) represent upstream area-weighted mean values, while panel (d) shows fractional lake area, defined as the percentage of total upstream lake area relative to contributing drainage area for each river reach. Whiskers extend to the 2.5th and 97.5th percentiles.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

3.4 Satellite embeddings

In addition to static ESA CCI land-cover attributes for the year 2020 (approximately the mid-point of many global observational streamflow records), GRIT-ADB also provides annual time series of 64-dimensional local-mean AlphaEarth satellite embeddings per sub-basin for 2017 to 2024. These embeddings are compact, data-driven summaries of multi-sensor Earth observation data that capture evolving environmental, land-surface, and anthropogenic characteristics. Unlike categorical land-cover maps, these embeddings encode continuous patterns related to vegetation dynamics, moisture conditions, surface properties, and human activity, enabling richer characterisation of sub-basin states and their temporal changes. Using the Weser River basin in Germany as an example, Figure 6a–j shows the spatial distribution of the first ten embedding dimensions in 2017, each exhibiting distinct, non-redundant spatial patterns across sub-basins that reflect different underlying environmental gradients. Figure 6k–t present the corresponding changes in these dimensions between 2024 and 2017, demonstrating how different latent features evolve unevenly across the basin. For example, the southern part of the basin exhibits a noticeable increase in the third embedding dimension between 2017 and 2024 (Figure 6m). Together, these panels illustrate the potential of AlphaEarth embeddings to capture spatial heterogeneity and temporal dynamics in river basin systems, reflecting variations in hydrological conditions and land-surface processes within a high-dimensional, continuous framework.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

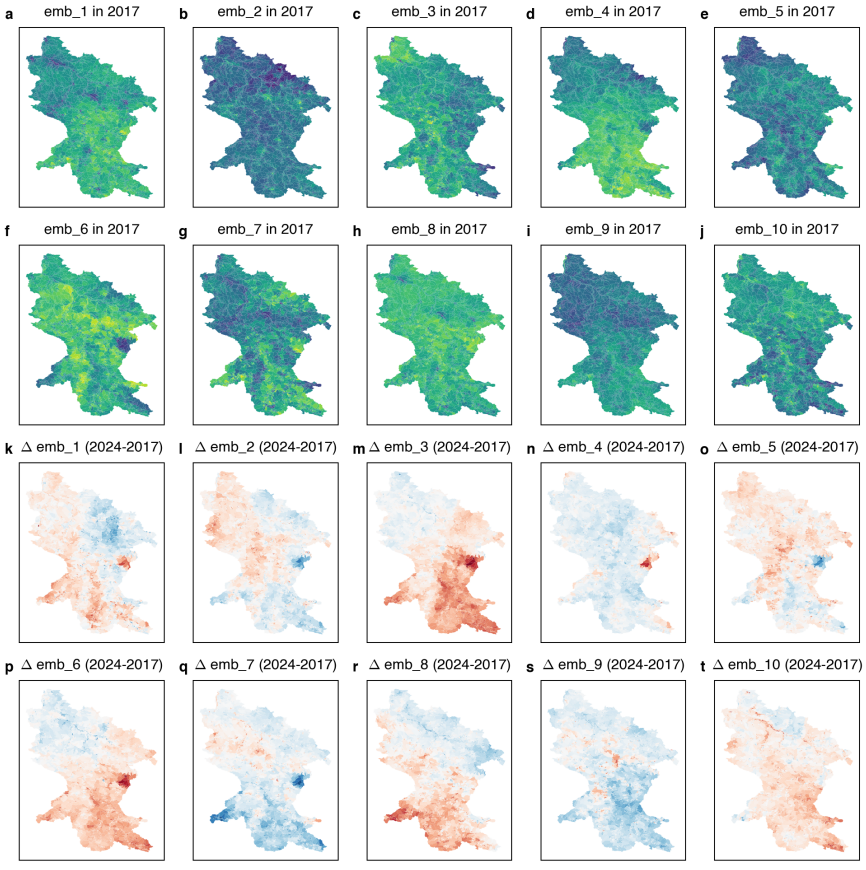


Figure 6. Spatial patterns of ten selected AlphaEarth embedding dimensions (local mean) across subbasins in the Weser River basin, Germany. Local mean estimates are computed as area-weighted averages based on the fraction of each grid cell within each subbasin. Panels (a–j) show the distribution of the first ten embedding dimensions for the year 2017, illustrating distinct spatial structures captured by each latent feature. Panels (k–t) depict the corresponding changes (Δ) in each embedding dimension between 2024 and 2017, highlighting spatially heterogeneous temporal dynamics across the basin. Colours represent normalised embedding values (normalised per dimension), with blue-to-green (a–j) indicating low to high values and red-to-blue (k–t) indicating negative to positive changes.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

4 Discussion

Global-scale hydro-environmental attribute databases play a central role in both large-scale hydrologic analyses and large-sample hydrology, particularly as machine learning methods are increasingly applied across extensive catchment datasets (Slater et al., 2025). HydroATLAS is one of the most widely used and comprehensive global hydro-environmental attribute databases, serving as a foundational resource for large-scale hydrologic and ecohydrological research (Linke et al., 2019), yet it has several key limitations. Here we introduce GRIT-ADB, a new global hydro-environmental attribute database designed to advance understanding of global river characteristics. As an illustration, we compare selected attributes in GRIT-ADB and HydroATLAS. Figure 7a and 7b present the total upstream sum reservoir volume in the Pearl River Delta derived from GRIT-ADB and HydroATLAS, respectively. GRIT-ADB captures braided river channels and their hydrological connectivity to upstream reservoirs, whereas this connectivity is not explicitly represented in HydroATLAS. This difference reflects the underlying hydrography: HydroATLAS is based on HydroSHEDS, which assumes a strictly dendritic network structure and does not represent bifurcations. Using GRIT-ADB, many reaches in the Pearl River Delta exceed 10,000 million cubic meters of upstream reservoir storage (Figure 7a). In contrast, HydroATLAS indicates that only a single mainstem exceeds this threshold, while nearly all other river segments show negligible upstream storage (Figure 7b). Such differences have implications for risk assessment of downstream hydrologic extremes.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

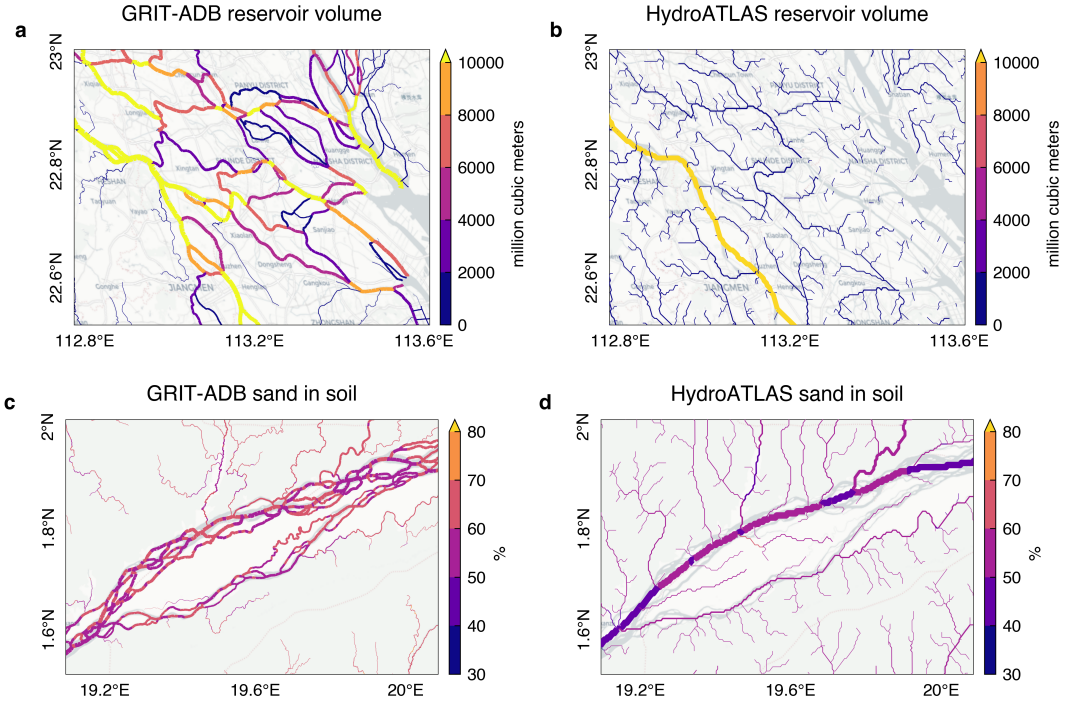


Figure 7. Comparison of example attributes in GRIT-ADB and HydroATLAS. Panels (a) and (b) present upstream accumulated reservoir volume in the Pearl River Delta from GRIT-ADB and HydroATLAS, respectively. The underlying reservoir datasets are GRanD v1.3 (GRIT-ADB) and GRanD v1.1 (HydroATLAS). Panels (c) and (d) present the local mean sand fraction in soil in the Congo River Basin from GRIT-ADB and HydroATLAS, respectively. The corresponding soil texture dataset sources are SoilGrids v2.0 (GRIT-ADB) and SoilGrids1km (HydroATLAS).

Figure 7c and 7d compare the local-mean fractional sand in soil in the Congo River Basin derived from GRIT-ADB and HydroATLAS. Both datasets derive soil texture information from SoilGrids, albeit different versions (SoilGrids v2.0 for GRIT-ADB and SoilGrids1km for HydroATLAS). SoilGrids v2.0 has a finer spatial resolution (250 m) than SoilGrids1km, and GRIT-ADB computes local mean statistics for river reaches with lengths of no more than 1 km, whereas HydroATLAS aggregates attributes at the river-segment scale. As a result, GRIT-ADB better captures spatial heterogeneity in soil texture within river basins. This improved representation is particularly important for high-resolution, network-based hydrologic modelling and prediction, especially in urbanised and coastal regions where soil texture exhibits substantial spatial variability.

Table 2 summarises the characteristics of GRIT-ADB relative to existing global hydro-environmental databases, including HydroATLAS and Environment90m. While Environment90m offers a substantially finer spatial representation, with over 700 million subbasins, neither HydroATLAS nor Environment90m represent river bifurcations. GRIT-ADB provides integrated information on reservoirs and lakes, similar to HydroATLAS but absent in Environment90m. Both HydroATLAS and GRIT-ADB provide pre-calculated upstream statistics for most attributes, including soil and geology variables, land-cover variables,

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

climate variables (with the exception of climate zones), and hydrological variables, whereas such information is not available in Environment90m. Overall, the datasets have different strengths, with GRIT-ADB emphasising detailed hydrographic structure and connectivity, while the others provide broader hydro-climatic and environmental attributes.

Table 2. Comparison of GRIT-ADB and existing global hydro-environmental databases (HydroATLAS and Environment90m). The comparison is restricted to global-scale, spatially continuous, harmonised, multi-domain databases providing integrated environmental attributes; regional or gauge-based datasets are not included.

Characteristic	GRIT-ADB-v1	HydroATLAS	Environment90m
Number of subbasins	> 20 million	> 1 million	> 700 million
Number of variables (expanded attributes)	98 (>300)	56 (281)	104 (>500)
Sub-basins	✓	✓	✓
Represents river bifurcations	✓	✗	✗
Includes time-varying attributes	✓	✗	✓
Pre-calculated upstream statistics	✓	✓	✗
Reservoir and lake attributes	✓	✓	✗

Although GRIT-ADB provides high-resolution river attributes and represents river networks more accurately than previous databases (Table 2), it has some important limitations. First, hydrological attributes pertaining to discharges, such as reanalysis-derived streamflow signatures, are not included in the current version of GRIT-ADB. This is because existing global-scale discharge datasets like GloFAS do not explicitly account for river bifurcations in their routing schemes, which can bias discharge representation, especially in downstream river sections where bifurcations are common. Future work will develop a global-scale river discharge dataset that explicitly accounts for river bifurcations.

Second, although the GRIT hydrography is high-resolution (30 m), most existing global-scale hydro-environmental datasets remain comparatively coarse. For example, the Multi-Source Weighted-Ensemble Precipitation (MSWEP) dataset is one of the most widely used precipitation datasets (Beck et al., 2019), but its spatial resolution is 0.1 degrees. We also emphasise that uncertainties and biases present in the underlying hydro-environmental datasets will propagate into GRIT-ADB. For example, the training data used to produce SoilGrids v2 exhibits strong spatial biases towards Europe and North America, with a limited number of soil profiles in Asia, Africa and South America (Poggio et al., 2021). As high-resolution products become available, they will be included in future versions of GRIT-ADB. To facilitate this, we have developed an efficient, automated workflow that can easily be adapted to new datasets.

Lastly, the representation of small streams in GRIT is currently constrained by the minimum contributing upstream drainage threshold of 50km^2. As a result, small headwater channels below this threshold are not included. Future versions of GRIT and GRIT-ADB will have a higher river network drainage density (and thus more complete representation of small streams) by lowering the minimum drainage threshold used to initiate rivers in GRIT.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

5 Conclusions

GRIT-ADB is a global hydro-environmental attribute database covering 19.6 million km of rivers and their associated catchments, built on the GRIT hydrography. The database provides 35 static and 64 time-varying variables (>300 attributes) across hydrology, physiography, climate, land cover and land use, and soils and geology, with both local and upstream statistics at reach, segment, and catchment scale.

GRIT-ADB differs from existing global databases by explicitly representing river bifurcations, enabling more accurate estimation of upstream-accumulated attributes in multi-channel systems. Because river networks integrate upstream processes, improved representation of connectivity reduces biases in downstream attribute estimates, with implications for modelling floods, connectivity, and ecohydrological dynamics, particularly in floodplains and deltaic systems where bifurcations are prevalent.

By combining topology-explicit hydrography with harmonised, multi-domain attributes derived from recent, high-resolution global datasets, GRIT-ADB provides a consistent, high-resolution basis for large-scale hydrological analysis across both gauged and ungauged basins. The framework is scalable and reproducible, enabling incorporation of updated datasets and extension to additional variables for continental to global hydrologic analysis and modelling, and supporting large-sample hydrology applications, including machine learning and other data-intensive approaches.

6 Data availability

The GRIT-ADB dataset for the Congo River basin is freely available at https://doi.org/10.5281/zenodo.19363178 (Zhang et al., 2026) as an example dataset. The full GRIT-ADB dataset will be deposited in a public repository and made openly available before final publication. The GRIT vector version is available at https://doi.org/10.528 1/zenodo.17435232 (Wortmann et al., 2025b) while the raster version is available at https://doi.org/10.5281/zenodo.15715535 (Wortmann et al., 2025a).

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Open Access Earth System Science Data Discussions

Appendix A: Supplementary Figures

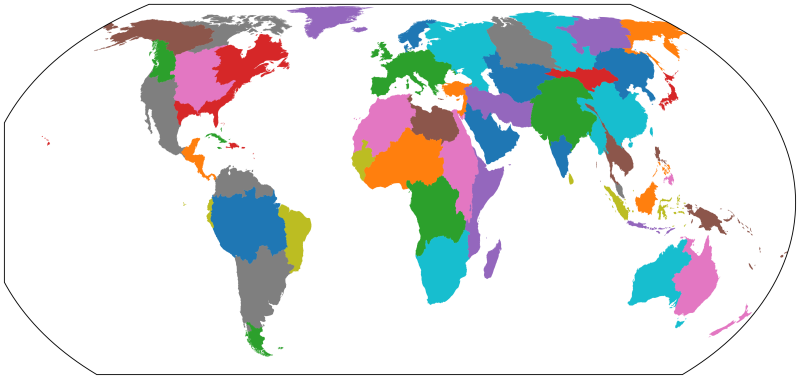


Figure A1. The 61 GRIT geographic domains used to extract hydro-environmental attributes

Appendix B: Supplementary Tables

Table A1: Summary of attributes, statistics type, units, and scaling factors in GRIT-ADB.

Attribute	Long name	Statistics	Units	Scale factor
dam_vol	Reservoir volume	Local & upstream	million cubic meters	1
lake_area	Lake area	Local & upstream	km^2	1
permeability	Subsurface permeability	Local & upstream	log(K)	1.00×10^{-6}
porosity	Subsurface porosity	Local & upstream	%	1.00×10^{-6}
BDTICM	Depth to bedrock	Local & upstream	m	0.01
sand	Sand in soil	Local & upstream	g/kg	1
clay	Clay in soil	Local & upstream	g/kg	1
silt	Silt in soil	Local & upstream	g/kg	1
bdod	Bulk density	Local & upstream	$kg\,dm^{-3}$	1
cec	Cation exchange capacity	Local & upstream	cmol(c)/kg	1
cfvo	Coarse fragments	Local & upstream	cm^3/dm^3	1
nitrogen	Nitrogen	Local & upstream	g/kg	1

Continued on next page

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Attribute	Long name	Statistics	Units	Scale factor
ocd	Organic carbon density	Local & upstream	hg m^{-3}	1
phh2o	pH water	Local & upstream	–	0.1
soc	Soil organic carbon	Local & upstream	dg/kg	1
ai	Aridity index	Local & upstream	–	0.0001
et0	Potential evapotranspiration	Local & upstream	mm	1
elevation	Elevation	Local & upstream	m	1
stream_gradient	Stream gradient	Local	m/km	1000
darea	Drainage area	Upstream	km^2	1
koppen	Climate classification	Local	–	1
cropland	Cropland extent	Local & upstream	%	100
forest	Forestland extent	Local & upstream	%	100
water	Water body extent	Local & upstream	%	100
grassland	Grassland extent	Local & upstream	%	100
urban	Urban extent	Local & upstream	%	100
emb_	Satellite embeddings	Local & upstream	–	1
strahler_order	Strahler order	Local	–	1
stream_length	Upstream stream length	Local	m	1
sinuosity	Sinuosity	Local	–	1
azimuth	Azimuth	Local	–	1
river_width	River width	Local	m	1
n_bif	Upstream number of bifurcations	Upstream	–	1
p_ann_mean	Mean annual precipitation	Local & upstream	mm	1
p_seasonality	Precipitation seasonality	Local & upstream	–	1
t_mean	Mean air temperature	Local & upstream	degreeC	1
t_seasonality	Air temperature seasonality	Local & upstream	–	1

Author contributions. L.S. initiated the work, secured funding, and supervised the research. B.Z., M.W., L.S., and co-authors contributed to the conceptualisation and design of the GRIT-ADB dataset. B.Z. led the development of the dataset, including methodology, data processing,

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

and analysis. M.W. developed the GRIT hydrography. Y.L. and F.K. contributed to the extraction of specific attributes in GRIT-ADB. B.Z. analysed the data, interpreted the results, and wrote the first draft. All authors contributed to interpretation and revision of the manuscript.

235 *Competing interests.* The authors declare that they have no competing interests

Acknowledgements. This work was funded by a FLF Fellowship (MR/V022008/1 and UKRI2054) to L.S.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

References

Allen, G. H. and Pavelsky, T. M.: Global extent of rivers and streams, Science, 361, 585–588, https://doi.org/10.1126/science.aat0636, 2018.

Altenau, E. H., Pavelsky, T. M., Durand, M. T., Yang, X., Frasson, R. P. d. M., and Bendezu, L.: The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A Global River Network for Satellite Data Products, Water Resources Research, 57, e2021WR030 054, https://doi.org/10.1029/2021WR030054, 2021.

Amatulli, G., Garcia Marquez, J., Sethi, T., Kiesel, J., Grigoropoulou, A., Üblacker, M. M., Shen, L. Q., and Domisch, S.: Hydrography90m: a new high-resolution global hydrographic dataset, Earth System Science Data, 14, 4525–4550, https://doi.org/10.5194/essd-14-4525-2022, 2022.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F.: Present and future Köppen–Geiger climate classification maps at 1-km resolution, Scientific Data, 5, 180 214, https://doi.org/10.1038/sdata.2018.214, 2018.

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° precipitation: Methodology and quantitative assessment, Bulletin of the American Meteorological Society, 100, 473–500, https://doi.org/10.1175/BAMS-D-17-0138.1, 2019.

Beck, H. E., van Dijk, A. I. J. M., Larraondo, P. R., McVicar, T. R., Pan, M., Dutra, E., and Miralles, D. G.: MSWX: Global 3-Hourly 0.1° Bias-Corrected Meteorological Data Including Near-Real-Time Updates and Forecast Ensembles, Bulletin of the American Meteorological Society, 103, E710–E732, https://doi.org/10.1175/BAMS-D-21-0145.1, 2022.

Brown, C. F., Kazmierski, M. R., Pasquarella, V. J., Rucklidge, W. J., Samsikova, M., Zhang, C., Shelhamer, E., Lahera, E., Wiles, O., Ilyushchenko, S., Gorelick, N., Zhang, L. L., Alj, S., Schechter, E., Askay, S., Guinan, O., Moore, R., Boukouvalas, A., and Kohli, P.: AlphaEarth Foundations: An embedding field model for accurate and efficient global mapping from sparse label data, arXiv preprint arXiv:2507.22291, https://doi.org/10.48550/arXiv.2507.22291, 2025.

C3S: Land cover classification gridded maps from 1992 to present derived from satellite observation, https://doi.org/10.24381/cds.006f2c9a, accessed 15-Dec-2025, 2019.

EEA: EU-Hydro River Network Database, Copernicus Land Monitoring Service hydrographic dataset, https://copernicus.discomap.eea.europa.eu/arcgis/rest/services/EUHydro/EU_Hydro_RiverNetworkDatabase/MapServer, river network and drainage layers for EEA39 countries, 2025.

Frame, J. M., Kratzert, F., Klotz, D., Gauch, M., Shalev, G., Gilon, O., Qualls, L. M., Gupta, H. V., and Nearing, G. S.: Deep learning rainfall–runoff predictions of extreme events, Hydrology and Earth System Sciences, 26, 3377–3392, https://doi.org/10.5194/hess-26-3377-2022, 2022.

Frasson, R. P. d. M., Pavelsky, T. M., Fonstad, M. A., Durand, M. T., Allen, G. H., Schumann, G., et al.: Global relationships between river width, slope, catchment area, meander wavelength, sinuosity, and discharge, Geophysical Research Letters, 46, 3252–3262, https://doi.org/10.1029/2019GL082027, 2019.

García-Márquez, J. R., Grigoropoulou, A., Tomiczek, T., Schürz, M., Bremerich, V., Torres-Cambas, Y., Buurman, M., Bego, K., Amatulli, G., and Domisch, S.: Environment90m – globally standardized environmental variables for freshwater science at high spatial resolution, Earth System Science Data, 18, 1541–1559, https://doi.org/10.5194/essd-18-1541-2026, 2026.

Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C., and Neal, J.: A 30 m global map of elevation with forests and buildings removed, Environmental Research Letters, 17, 024 016, https://doi.org/10.1088/1748-9326/ac4d4f, 2022.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Huscroft, J., Gleeson, T., Hartmann, J., and Börker, J.: Compiling and mapping global permeability of the unconsolidated and consolidated earth: GLHYMPS 2.0, Geophysical Research Letters, 45, 1897–1904, https://doi.org/10.1002/2017GL075860, 2018.

Jones, E. R., van Vliet, M. T., and Kratzert, F.: Caravan-Qual: A global scale integration of water quality observations into a large-sample hydrology dataset, EarthArXiv preprint, https://doi.org/10.31223/X54J39, preprint describing an open access water quality and hydrology dataset, 2025.

Lehner, B. and Grill, G.: Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, Hydrological Processes, 27, 2171–2186, https://doi.org/10.1002/hyp.9740, 2013.

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, Eos, Transactions, American Geophysical Union, 89, 93–94, https://doi.org/10.1029/2008eo100001, 2008.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., and Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, Frontiers in Ecology and the Environment, 9, 494–502, https://doi.org/10.1890/100125, 2011.

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., and Anand, M.: Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution, Scientific Data, 6, 283, https://doi.org/10.1038/s41597-019-0300-6, 2019.

Moore, R. B., McKay, L. D., Rea, A. H., Bondelid, T. R., Price, C. V., Dewald, T. G., and Johnston, C. M.: User's Guide for the National Hydrography Dataset Plus High Resolution (NHDPlus HR), Tech. Rep. 2019–1096, U.S. Geological Survey, https://doi.org/10.3133/ofr20191096, 2019.

Mullin, M.: The effects of drinking water service fragmentation on drought-related water security, Science, 368, 274–277, https://doi.org/10.1126/science.aba7353, 2020.

Pi, X., Luo, Q., Feng, L., Xu, Y., Tang, J., Liang, X., Ma, E., Cheng, R., Fensholt, R., Brandt, M., Cai, X., Gibson, L., Liu, J., Zheng, C., Li, W., and Bryan, B. A.: Mapping global lake dynamics reveals the emerging roles of small lakes, Nature Communications, 13, 5777, https://doi.org/10.1038/s41467-022-33239-3, 2022.

Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., and Rossiter, D.: SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty, SOIL, 7, 217–240, https://doi.org/10.5194/soil-7-217-2021, 2021.

Rosa, R. L. and Sangiorgio, M.: Global water gaps under future warming levels, Nature Communications, https://doi.org/10.1038/s41467-025-56517-2, 2025.

Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., and Dai, Y.: Mapping the global depth to bedrock for land surface modeling, Journal of Advances in Modeling Earth Systems, 9, 65–88, https://doi.org/10.1002/2016MS000686, 2017.

Slater, L., Blougouras, G., Deng, L., Deng, Q., Ford, E., Hoek van Dijke, A., Huang, F., Jiang, S., Liu, Y., Moulds, S., et al.: Challenges and opportunities of ML and explainable AI in large-sample hydrology, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 383, https://doi.org/10.1098/rsta.2024.0287, 2025.

Woods, R. A.: Analytical model of seasonal climate impacts on snow hydrology: Continuous snowpacks, Advances in Water Resources, 32, 1465–1481, https://doi.org/10.1016/j.advwatres.2009.06.011, 2009.

Wortmann, M., Slater, L., Hawker, L., Liu, Y., and Neal, J.: Global River Topology (GRIT) raster datasets, https://doi.org/10.5281/zenodo.15715535, 2025a.

Wortmann, M., Slater, L., Hawker, L., Liu, Y., and Neal, J.: Global River Topology (GRIT) vector datasets, https://doi.org/10.5281/zenodo.17435232, 2025b.

https://doi.org/10.5194/essd-2026-279
Preprint. Discussion started: 26 June 2026

Wortmann, M., Slater, L. J., Hawker, L. P., Liu, Y., Neal, J. C., Zhang, B., Ashworth, P. J., Boothroyd, R., Cloke, H., and Delorme, P.: Global River Topology (GRIT): A bifurcating river hydrography, Water Resources Research, https://doi.org/10.1029/2024WR038308, 2025c.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., and Pavelsky, T. M.: MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset, Water Resources Research, 55, 5053–5073, https://doi.org/10.1029/2019WR024873, 2019.

Zhang, B., Slater, L., Wortmann, M., Liu, Y., and Moulds, S.: GRIT-ADB: A Global Hydro-Environmental Attribute Database for the GRIT Hydrography, https://doi.org/10.5281/zenodo.19363178, 2026.

Zomer, R. J., Xu, J., and Trabucco, A.: Version 3 of the Global Aridity Index and Potential Evapotranspiration Database, Scientific Data, 9, 409, https://doi.org/10.1038/s41597-022-01493-1, 2022.