https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

CAMELS-COL: A Large-Sample Hydrometeorological Dataset for Colombia

David A. Jimenez [1,2], Julian E. Meneses [2], Pedro H. Bernardes Solha[1], Alvaro Avila-Diaz[2], Benjamin Quesada[2], Bruno Melo Brentan[1], André Ferreira Rodrigues [1].

[1]Hydraulic and Water Resources Department, Federal University of Minas Gerais UFMG, 31270-901, MG, Belo Horizonte, Brazil.
[2]Earth System Science Program, School of Sciences and Engineering, Universidad del Rosario, Bogotá D.C., Colombia.

Correspondence to: Alvaro Avila Diaz (alvaro.avila@urosario.edu.co)

Abstract. Catchment Attributes and Meteorology for Large-Sample Studies (CAMELS-COL) is a large-sample hydrological dataset for Colombia that integrates daily meteorological and hydrological time series with comprehensive catchment attributes. The dataset comprises daily precipitation, evapotranspiration, and temperature from CHIRPS and MSWX satellite sources and streamflow records from 347 gauging stations covering the range 1981-2022. Additionally, CAMELS-COL provides a wide range of catchment attributes, including physiographic characteristics, climatic indices, hydrological signatures, land cover, geology, and soil properties, derived primarily from official governmental sources. CAMELS-COL follows the standardized framework of previous CAMELS datasets, such as those developed for Brazil and Chile, to ensure consistency with global hydrological datasets. By incorporating Colombian catchments across diverse hydroclimatic regions, including the Andean, Amazonian, and Caribbean basins, this dataset extends the CAMELS initiative into tropical environments, offering a unique resource for hydrological research. The analysed basins show low flood susceptibility. An analysis of the aridity index reveals that 74.7% of basins have a subhumid climate, 20.7% are semiarid, and only 4.6% are classified as humid. Flow elasticity to precipitation is highest in the Amazon and Orinoco, highlighting their greater streamflow sensitivity to rainfall changes. The Base Flow Index underscores groundwater's crucial role in stabilizing and regulating surface water, particularly in forested basins. The dataset supports studies on hydrological processes, extreme hydroclimatic events, climate change impacts, and water resource management. Public availability encourages scientific collaboration and facilitates the inclusion of Colombian catchments in continental and global hydrological analyses. The dataset is accessible at https://doi.org/10.5281/zenodo.15554735 (Jimenez et al., 2025).

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Open Access Earth System Science Data Discussions

1. INTRODUCTION

One of the primary objectives of hydrology is to understand hydrological processes in sufficient detail so that hydrological models can accurately simulate a wide range of hydrological environments across different spatial and temporal scales and under diverse environmental conditions (Gupta et al., 2014). However, this cannot be achieved solely through intensive research in a few heavily instrumented basins. In this context, large-scale hydrological research supported by extensive datasets from multiple basins enables a broader, solid formulation and understanding of hydrological processes (Addor et al., 2020).

Currently, a variety of datasets compile different hydrometeorological and physiographic information from basins in various parts of the world. In this sense, Arsenault et al., (2020) compile in "The Hydrometeorological Sandbox - École de technologies supérieure – HYSETS" the hydrological, meteorological data, and physiographic information of more than 14,000 catchments located in high latitude regions covering 1950-2018. On the other hand, Kratzert et al. (2023) compiled and standardized climatological, hydrological, and physiographic information for over 6,830 basins as part of the Caravan initiative, resulting in an aggregated and standardized representation of the CAMELS (Catchment Attributes and Meteorology for Large-sample Studies) datasets.

The CAMELS project was designed to improve the representation of hydrological processes across large spatial domains, making it possible to simulate a variety of hydrological environments under different climatic and topographic conditions. What differentiates CAMELS from other existing hydro-climatological datasets is its specific focus on catchment attributes, which include not only meteorological data but also physiographic, hydrological, geological, soil, and land cover features that are critical for hydrological modeling. CAMELS was created for the United States CAMELS-US (Addor et al., 2017), and later expanded to Chile CAMELS-CL (Alvarez-Garreton et al., 2018), Brazil CAMELS-BR (Chagas et al., 2020), United Kingdom CAMELS-UK (Coxon et al., 2020), Australia CAMELS-AUS (Fowler et al., 2021), Switzerland CAMELS-CH (Höge et al., 2023), Germany CAMELS – DE (Loritz et al., 2024), France CAMELS-FR (Delaigue et al., 2024), Denmark CAMELS-DK (Liu et al., 2024), and India CAMELS-INDIA (Mangukiya et al., 2024).

In Colombia, most hydrometeorological stations are concentrated in the Magdalena region, a mountainous area within the Andes Mountain range. Limited funding for the Institute of Hydrology, Meteorology, and Environmental Studies (IDEAM), which is responsible for monitoring and researching the country's climate conditions, has restricted hydrometeorological observation in regions such as the Amazon, Orinoco, and Pacific (Quesada and Avila-Diaz, 2023). On the other hand, as highlighted by Avila-Diaz et al. (2024), the Water National Study (ENA)—one of the main governmental resources for assessing the state of the country's water resources, updated biannually—has not incorporated updates on the country's water situation in its most recent editions. Moreover, these editions have neither included analyses related to the El Niño-Southern Oscillation (ENSO) nor comprehensively addressed the impacts of global climate change in the country.

In this context, the CAMELS-CO dataset seeks to contribute to the understanding of hydrological processes across Colombia. Colombia is the fourth largest country in South America (~1,141,748 km²) with a vast range of topographical and

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

climatological conditions, as the Pacific Ocean, the Andes Mountains, and the Amazon Rainforest influence it. To address the challenges of understanding the hydrological dynamics in this topographically and climatologically diverse country, Furthermore, this work will contribute to the understanding of regional hydrology in South America. This paper introduces a comprehensive and unique dataset covering 347 catchments across Colombia, with a higher concentration in the Andes Mountain range. The dataset includes: i) Observed daily streamflow time series from 347 stream gauges; ii) Daily meteorological data derived from satellite observations (temperature and precipitation); iii) Daily meteorological data and its derived products (e.g., potential evapotranspiration - ETo) iv) Climatic indices, v) Hydrological signatures, and iv) Catchments attributes, including land cover, geology, and soil characteristics, are calculated from official information provided by governmental institutions. One of the main contributions of CAMELS-COL is the calculation of the Curve Number using the Soil Conservation Service (SCS) method for each of the analyzed catchments, which is a parameter widely used in peak flow estimation.

All data are provided in an accessible and standardized format, adhering to the workflow of previous CAMELS datasets, which allows for direct comparison. By integrating hydrological data from multiple catchments into a single, ready-to-use resource, this dataset streamlines data collection and preprocessing, significantly reducing time and effort. It supports hydrological research in Colombia, enhances the benchmarking, calibration, and validation of hydrological models, and aids in flood and drought monitoring. Additionally, it contributes to climate change adaptation planning at the departmental level and enables global and regional hydroclimatic comparisons. The paper is structured as follows: Sect. 2 describes the study area and data collection. Sect. 3 details the derived catchment attributes, including physiographic, hydroclimatological, geological, land cover, and soil characteristics, along with a discussion of their spatial distribution. Finally, Sect. 4 summarizes the main results of the paper.

2 Data provided by CAMELS-COL

The CAMELS-COL dataset provides a comprehensive collection of hydrometeorological, physiographic, climate indices, and hydrological signatures for Colombia's main catchments, following similar works in South America, such as CAMELS-BR (Chagas et al., 2020) and CAMELS-CL (Alvarez-Garreton et al., 2018).

The dataset includes catchment identification and localization. The physiographic characteristics encompass drainage area, perimeter, Gravelius compactness index, form factor, main channel length, equivalent slope of the main channel, the concentration time by four different methodologies (Kirpich, Ven te Chow, Johnstone, and Corps Engineers) and the mean curve number of Soil Conservation Service (SCS) method. Additionally, the dataset includes the percentage distribution of land cover types from MapBiomas-Colombia (MapBiomas, 2022), as well as the lithological and soil characteristics of the catchments, derived from the Geological Map of Colombia and the Soil Map, provided by the Colombian Geological Survey and the Agustín Codazzi Geographic Institute, respectively. In all cases, priority was given to official sources data, to ensure the highest possible representativeness and reliability of the information included in CAMELS-COL.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

The hydrometeorological data includes daily streamflow observations, precipitation, potential evapotranspiration, and maximum, mean, and minimum temperatures for each catchment. The climate indices encompass the aridity index, frequency, and duration of high and low precipitation days. The hydrological signatures include the 5th and 95th daily quantile flow, runoff ratio, streamflow elasticity concerning precipitation, slope of the flow duration curve, base flow index, and the frequency and duration of high and low flow days. Table 1 describes the information included in the dataset, and the following sections present a detailed explanation of the data source, and the methodology used to estimate each index.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Open Access Earth System Science Data Discussions

Table 1 Summary of the data provided by CAMELS-COL.

Attribute class	Attribute name	Description	Units	Data source
Catchment information	gauge_id	Catchment identifier (IDEAM code of streamflow)	-	IDEAM
	gauge_department	Gauge streamflow department provided by IDEAM	-	IDEAM
	gauge_lat	Gauge streamflow latitude (WGS-1984 System)	ºN	IDEAM
	gauge_lon	Gauge streamflow longitude (WGS-1984 System)	ºE	IDEAM
	gauge_elev	Gauge streamflow elevation	m.a.s.l.	IGAC DEM
Physiograpic characteristics	area	Catchment area	km^2	IGAC DEM
	perimeter	Catchment perimeter	km	IGAC DEM
	gravelius_index	Gravelius compactness index of catchment	-	-
	factor_form	Form factor of catchment	-	-
	streng_chanel	Streng of the main chanel	km	IGAC DEM
	equi_slope	Equivalent slope of main chanel	m/m	Silveira (2005)
	cn_catchment	Mean curve number of Soil Conservation Service (SCS) method.	-	Wu et al., (2024)
	tc_kirpich	Concentration time by Kirpich Equation	hours	Silveira (2005)
	tc_chow	Concentration time by Ven te Chow Equation	hours	Silveira (2005)
	tc_Johnstone	Concentration time by Johnstone Equation	hours	Silveira (2005)
	tc_Engi_Corps	Concentration time by Corps Engineers Equation	hours	Silveira (2005)
Hydrometeorological data	pr	Daily precipitation of catchment	mm/dia	CHIRPSV2
	poten_evapo	Daily potential evapotranspiration of catchment	mm/dia	MSWX
	t_max	Daily maximum temperature of catchment	ºC	MSWX
	t_mean	Daily mean temperature of catchment	ºC	MSWX
	t_min	Daily minimum temperature of catchment	ºC	MSWX
	streamflow	Mean daily streamflow of catchment	m^3/s	IDEAM
Climatic indices.	aridity	Aridity, computed as the ratio of mean potential evapotranspiration (ETo) to mean precipitation	-	Arora (2002)
	high_prec_freq	Frequency of high precipitation days (≥ 5 times the mean daily precipitation)	days yr^{-1}	CHIRPS v2.0
	high_prec_dur	Average duration of high precipitation events (number of consecutive days ≥ 5 times the mean daily precipitation)	days	CHIRPS v2.0
	low_prec_freq	Frequency of dry days (< 1 mm day^{-1})	days yr^{-1}	CHIRPS v2.0
	low_prec_dur	Average duration of dry precipitation events (number of consecutive dry days)	days	CHIRPS v2.0

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Open Access Earth System Science Data Discussions

Attribute class	Attribute name	Description	Units	Data source
Hydrological signatures.	q_mean	Mean daily discharge	mm day^{-1}	IDEAM
	runoff_ratio	Runoff ratio, computed as the ratio of mean daily discharge to mean daily precipitation	-	
	stream_elas	Streamflow precipitation elasticity (i.e., streamflow's sensitivity to precipitation changes at the annual timescale, using the mean daily discharge as reference). See equation 7 in Sankarasubramanian et al. (2001), with the last element being P/Q not Q/P (i.e as presented previously in CAMELS-BR)	-	
	slope_fdc	Slope of the flow duration curve between the log-transformed 33rd and 66th streamflow percentiles	-	
	baseflow_index	Baseflow index, computed as the ratio of mean daily baseflow to mean daily discharge, with the hydrograph separation performed using the Ladson et al. (2013) digital filter	-	
	Q5	5% flow quantile (low flow)	mm day^{-1}	
	Q95	95% flow quantile (high flow)	mm day^{-1}	
	high_q_freq	Frequency of high-flow days (> 9 times the median daily flow)	days yr^{-1}	
	high_q_dur	Average duration of high-flow events (number of consecutive days > 9 times the median daily flow)	days	
	low_q_freq	Frequency of low-flow days (< 0.2 times the mean daily flow)	days yr^{-1}	
	low_q_dur	Average duration of low-flow days (number of consecutive days < 0.2 times the mean daily flow)	days	
Land cover characteristics.	forest_perc	Percentage of catchment covered by forest formation	%	Mapbiomas (2022)
	nat_non_forest_form_perc	Percentage of catchment covered by natural non-forest formation	%	
	agricul_livestock_perc	Percentage of catchment covered by agriculture and livestock area	%	
	non_vegeted_perc	Percentage of catchment covered by non-vegetated area	%	
	water_bodies_perc	Percentage of catchment covered by water body	%	
	non_identi_land_perc	Percentage of catchment covered by non-identified land cover	%	
Geologic characteristics.	unconso_depo_perc	Percentage of catchment covered by unconsolidated deposit	%	Gómez et al., (2023)
	hypabyssal_rock_perc	Percentage of catchment covered by hypabyssal rock	%	
	metamor_rock_perc	Percentage of catchment covered by metamorphic rock	%	
	plutonic_rock_perc	Percentage of catchment covered by plutonic rock	%	
	sedimen_rock_perc	Percentage of catchment covered by sedimentary rock	%	
	volcaniclastic_rock_perc	Percentage of catchment covered by volcaniclastic rock	%	
	volcanic_rock_perc	Percentage of catchment covered by volcanic rock	%	

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Attribute class	Attribute name	Description	Units	Data source
Soil characteristics.	inceptisols_perc	Percentage of catchment covered by inceptisols	%	IGAC
	entisols_perc	Percentage of catchment covered by entisols	%	
	andisols_perc	Percentage of catchment covered by andisols	%	
	urban_perc	Percentage of catchment covered by urban areas	%	
	alfisols_perc	Percentage of catchment covered by alfisols	%	
	vertisols_perc	Percentage of catchment covered by vertisols	%	
	aridisols_perc	Percentage of catchment covered by aridisols	%	
	water_bodies_perc	Percentage of catchment covered by water bodies	%	
	ultisols_perc	Percentage of catchment covered by ultisols	%	
	spodosols_perc	Percentage of catchment covered by spodosols	%	
	coal_mine_pit_perc	Percentage of catchment covered by coal mine pit	%	
	histosols_perc	Percentage of catchment covered by histosols	%	
	oxisols_perc	Percentage of catchment covered by oxisols	%	
	eroded_misce_perc	Percentage of catchment covered by eroded miscellaneous	%	
	rocky_misce_perc	Percentage of catchment covered by rocky miscellaneous	%	
	mollisols_perc	Percentage of catchment covered by mollisols	%	
	perpetual_snow_perc	Percentage of catchment covered by perpetual snow	%	

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

3 Methodology

3.1 Study area

The research focuses on Colombia, a country with diverse topography ranging from sea level to 5,703 meters above sea level (m.a.s.l) in the Andes Mountains. Colombia is segmented into five primary hydrological regions: the Amazon, Pacific, Orinoco, Caribbean, and Magdalena (Figure 1a). The Amazon contributes 37% of the country's surface streamflow, the Pacific 14%, the Orinoco 26%, the Caribbean 9%, and the Magdalena 14% (Bedoya, and Gonzales, 2021).

The Colombian climate exhibits substantial diversity and spatial variability. The warmest average monthly temperatures are observed in the lowlands of the Magdalena and Orinoco regions, where temperatures exceed 24°C. In contrast, the coolest temperatures are found in the mountainous areas, ranging from 8°C and 20°C Figure 1b). Rainfall patterns exhibit significant variability: the northern Caribbean is the driest region, with an average annual rainfall of 500 mm/year, while the Pacific region is the wettest, receiving over 11,000 mm/year in annual precipitation (Figure 1c). Annual evapotranspiration generally varies from 1200 to 1400 mm/year, but it rises to 1400–1600 mm/year in the lowlands of the Magdalena region and the northern Caribbean (Figure 1d). Regarding land cover, and considering the 2022 MapBiomas report, it is observed that most of Colombia's land cover is mainly associated with forest formations, agricultural and livestock areas, and natural non-forest formations (Figure 1e). According to the official geological map of Colombia (Gómez et al., 2023), the predominant Colombian lithology is mainly associated with sedimentary rocks and unconsolidated deposits (Figure 1f).

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

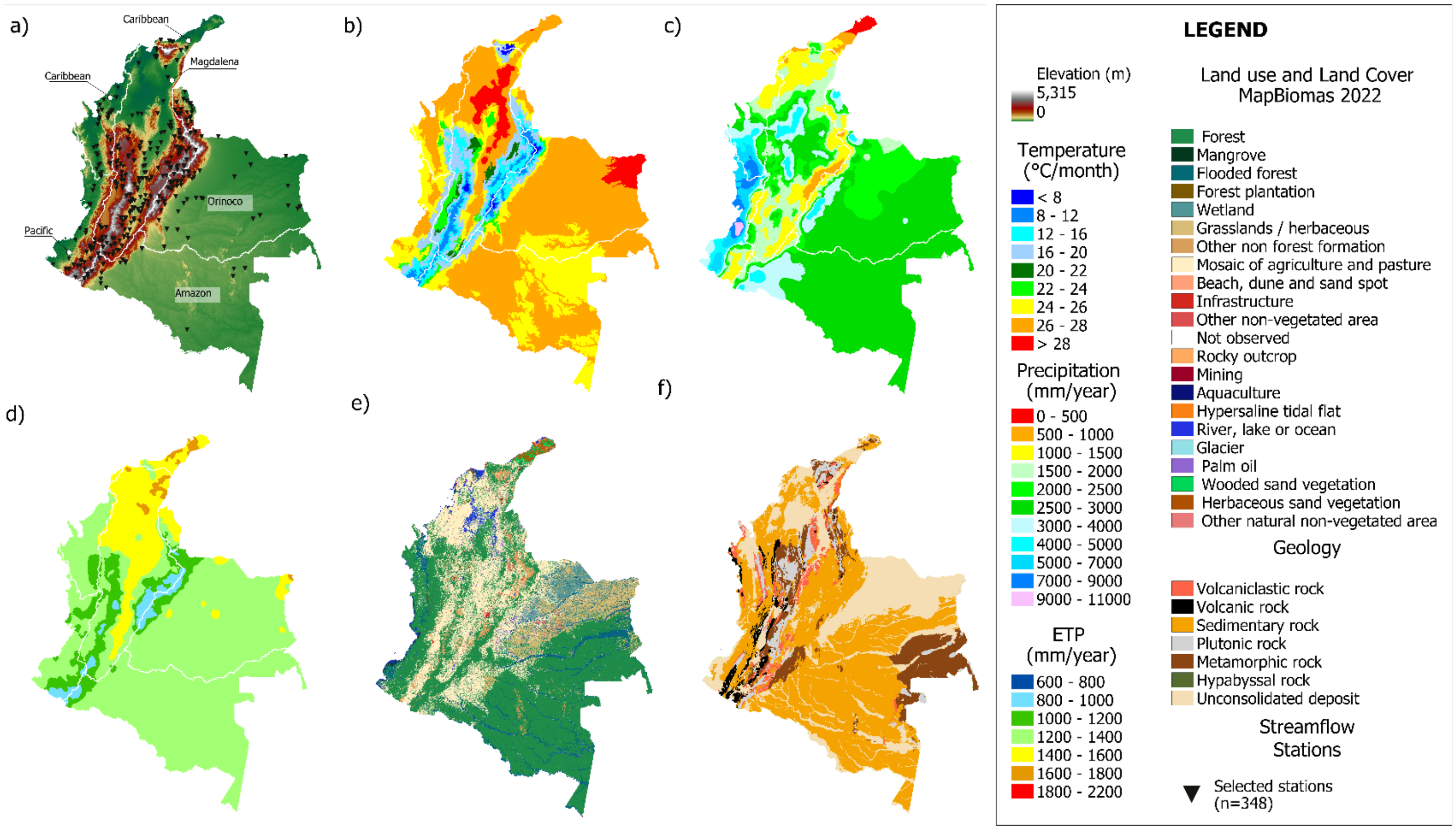

Figure 1: Study Area. a) Colombian topography and streamflow stations considered in the study. b) Mean monthly temperature in Colombia during 1981-2010 reported by the IDEAM . c) Mean annual precipitation in Colombia during 1981-2010 (mm/year) reported by the IDEAM. d) Mean annual Potential Evapotranspiration in Colombia during 1981-2010 reported by the IDEAM, and d) Mean annual precipitation in Colombia during 1981-2010 (mm/year) reported by the IDEAM. e) Colombian land cover identified in MapBiomas 2022. f) Colombian Geology Map reported by the Geological Service of Colombia.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

3.2 Data collection

3.2.1 Streamflow Data and Watershed Delineation

Daily Streamflow data is obtained from 921 gauging stations managed by IDEAM (http://dhime.ideam.gov.co/atencionciudadano/) as of February 2023. Following the methodology adopted in CAMELS-CL, CAMELS-IND, and CAMELS-DE, only streamflow stations with at least 10 years of records starting from 1980 are considered. Additionally, to ensure the representativeness of the information included in CAMELS-COL, only basins with at least 70% data availability are evaluated (see Figure 2).

Once the eligible streamflow stations are identified, the corresponding drainage areas are delineated using the Digital Elevation Model (DEM) of Colombia provided by the Instituto Geográfico Agustín Codazzi (Available at: Modelo Digital del Elevación. Colombia. STRM 30 Metros, https://www.colombiaenmapas.gov.co, accessed Jun 2024). The delineation process was carried out using the WhiteboxTools package (Lindsay, 2016), with streamflow stations serving as the outlet points for each basin.

After delineating the drainage areas, their consistency was evaluated. After delineating the drainage areas, their consistency was assessed. Basins with inconsistent delineations were excluded, particularly in cases where inconsistencies were attributed to limitations of the Digital Elevation Model (DEM) used. Additionally, in line with the methodology used in CAMELS-DE, only basins containing at least one central pixel with available gridded weather data are considered. This decision ensures consistency with previous CAMELS-datasets, which included basins where gridded or satellite-based climate data could be reliably assigned within the basin boundaries.

After this evaluation. 347 streamflow stations are selected for analysis, with 22 located in the Amazon hydrological region, 31 in the Pacific, 57 in the Orinoco, 55 in the Caribbean, and 182 in the Magdalena basin (Figure 1a). Approximately 47% of the streamflow stations are located at elevations below 500 meters above sea level (m.a.s.l.), 21% between 500 and 1000 m.a.s.l., 13% between 1000 and 1500 m.a.s.l., 9% between 1500 and 2000 m.a.s.l., 5% between 2000 and 2500 m.a.s.l., and the remaining 5% above 2500 m.a.s.l. (Figure 1a).

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

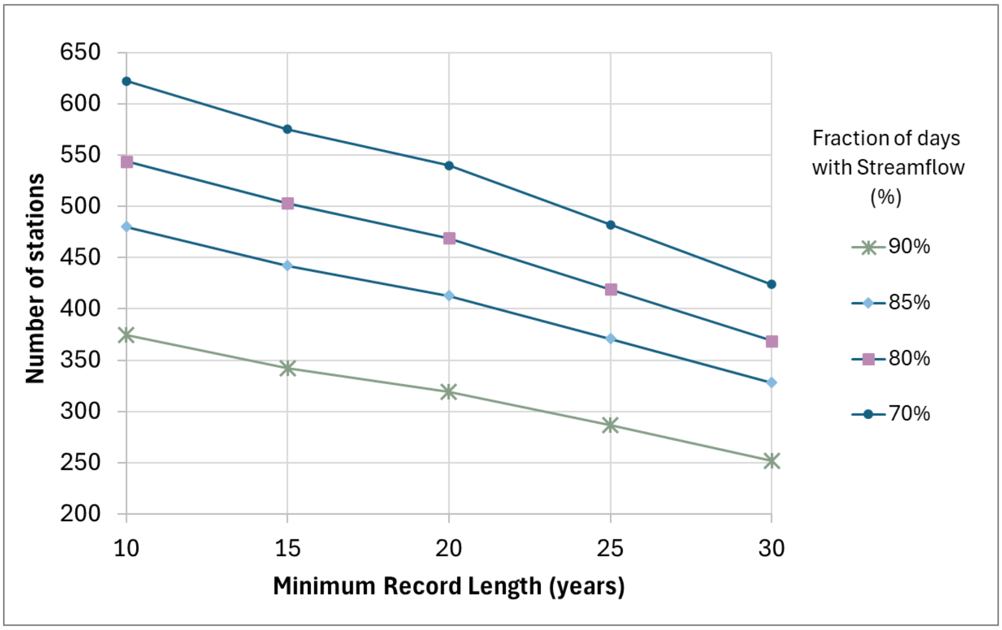

Figure 2 Number of stations with at least 10 to 30 years of daily streamflow records, categorized by different thresholds of data availability (90%, 85%, 80%, and 70% percentage of days with streamflow observations).

3.2.2 Climatological Information

To analyze the hydroclimatological conditions of each basin, precipitation and maximum and minimum temperature data are obtained from satellite sources. The maximum (T_{max}) and minimum temperatures (T_{min}) are extracted from the MSWX dataset. MSWX is a global gridded meteorological dataset with 3-hourly resolution at 0.1°×0.1° available at the www.gloh2o.org/mswx by GloH2O (Araghi et al., 2023; Beck et al., 2022). MSWX consists of four sub-products: the historical record from 1 January 1979 to approximately 5 days ago (MSWX-Past), the near real-time extension with a 3-hour latency (MSWX-NRT), the medium-range forecast ensemble covering up to 10 days (MSWX-Mid), and the seasonal forecast ensemble spanning up to 7 months (MSWX-Long) (Beck et al., 2022). At the same time, potential evapotranspiration (ETo) is estimated using satellite-derived temperature data, applying the Hargreaves method (Hargreaves and Samani, 1985). The calculation follows the equations outlined by Proutsos et al. (2024) as follows:

$$ETo = 0.0023 \cdot R_a \cdot (T + 17.8) \cdot (T_{max} - T_{min})^{0.5} \quad [1]$$

where: ETo is the potential evapotranspiration (mm day-1), R_a is the extraterrestrial radiation (mm day-1) estimated from latitude and Julian day, calculated based on latitude and Julian day, following the methodology presented by the FAO (2006) in "*Evapotranspiración del Cultivo: Guías para la Determinación de los Requerimientos de Agua de los Cultivos (Estudios*

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

FAO: Riego y Drenaje)", T, T_{max} and T_{min} are the daily mean (calculated as the average of T_{max} and T_{min}), maximum, and minimum air temperatures in ºC respectively recovered from MSWX dataset.

Precipitation information is retrieved from the Climate Hazards Group InfraRed Precipitation with Station data – CHIRPS (Available at: https://www.chc.ucsb.edu/data/chirps). It is a quasi-global (50°S-50°N) dataset with daily data available from 1981 to the present with a horizontal resolution of 5 km (Funk et al., 2015). The CHIRPS algorithm uses three data sources to calculate precipitation: a) Rainfall data from CHPclim, b) satellite-based precipitation estimates using TIR (Thermal Infrared), and c) in situ measurements from rain stations (Dinku et al., 2018). This dataset is selected for its effectiveness in accurately representing precipitation patterns across Colombia (Romero-Hernández et al., 2024; Arregocés et al., 2023; Valencia et al., 2023; Cepeda and Cañon, 2022). The dataset presents daily data from January 1981 to December 2022 because it is the common range between climatological information and streamflow series.

The average climatic conditions of each basin are estimated by calculating the mean of the pixel values located within its boundaries. No interpolation method was applied, as the goal was to use the original satellite data directly, preserving its resolution and spatial characteristics.

3.2.3 Geology

The geological characteristics of the study catchments are obtained from the "Geological Map of Colombia 2023, Scale 1:1.5M," prepared by the GSC (Gómez et al., 2023). This map was created from a raster with a spatial resolution of 30.0 m. The 193 geological units represented on the map are chronostratigraphic units grouped according to the age and lithology of the materials. For the age, the units were classified into six groups (Cenozoic, Mesozoic, Paleozoic, Neoproterozoic, Mesoproterozoic, Paleoproterozoic), with the reference being the 2022 International Chronostratigraphic Chart (Cohen et al., 2022). Regarding lithology, the units were divided into seven groups, distinguishing between rocks and deposits (Unconsolidated deposit, Sedimentary rock, Volcaniclastic rock, Volcanic rock, Hypabyssal rock, Plutonic rock, Metamorphic rock). The geological data of GCS can be accessed through their official website at https://www2.sgc.gov.co/MGC/Paginas/mgc_1_5M2023.aspx. To facilitate the interpretation of the results, only the primary lithology classification is presented in the main text; this means Unconsolidated deposit, Sedimentary rock, Volcaniclastic rock, Volcanic rock, Hypabyssal rock, Plutonic rock, Metamorphic rock. However, the feature "Geologic_characteristics" in the dataset includes the percentage distribution of each of the 193 geological units in each catchment. The lithology for each catchment is estimated by intersecting the geological dataset raster with the shape of the area.

3.2.4 Land Cover

The land cover of each catchment is estimated using data from MapBiomas Colombia for the year 2022 (MapBiomas, 2022). The classification maps, derived from Landsat mosaics, generate thematic land cover maps. These maps are part of the

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

MapBiomas initiative, which updates them whenever classification algorithms are refined. The MapBiomas Colombia dataset consists of a raster with a spatial resolution of 30.0 m. The dataset was accessed and downloaded in June 2023 from https://colombia.mapbiomas.org/segunda-coleccion-de-mapbiomas-colombia . It is important to note that, as of that date, the most recent dataset available was from 2022. Land cover estimation is performed by intersecting the MapBiomas raster with the shape of the area.

The use of MapBiomas Colombia in this study is justified by its high spatial resolution and standardized classification methodology (MapBiomas, 2022). Additionally, it is important to highlight that no official governmental entity currently provides an updated land cover dataset covering all of Colombia. In this context, MapBiomas Colombia stands out as a reliable source, as it involves researchers and specialists in remote sensing, programming, ecosystems, and land use, ensuring a scientifically robust classification. Furthermore, its methodology enables comparisons with catchments in other countries, which is particularly relevant given that this project spans multiple nations, including Brazil, Argentina, Peru, and Ecuador. The MapBiomas dataset classifies land cover into 20 types (Table 2), divided into five major groups: 1) Forest Formation, 2) Natural Non-Forest Formation, 3) Agricultural and Livestock Areas, 4) Non-Vegetated Areas, and 5) Water Bodies. To facilitate the interpretation of the results, only the five major classifications are presented in the main text. However, the feature "Land_cover_characteristics" in the dataset includes the percentage distribution of each land cover category considered in MapBiomas.

Table 2 Land Cover classification in MapBiomas Colombia 2022 – Collection 2

Forest formation	
1	Forest
2	Mangrove
3	Flooded forest
4	Wooded sand vegetation
Natural non forest formation	
5	Wetland
6	Grasslands / herbaceous
7	Hypersaline tidal flat
8	Rocky outcrop
9	Herbaceous sand vegetation
10	Other non forest formation
Agricultural and livestock area	
11	Forest plantation
12	Palm oil
13	Mosaic of agriculture and pasture
Non-vegetated area	
14	Beach, dune and sand spot
15	Infrastructure

16	Mining
17	Other non-vegetated area
Water body	
18	River, lake or ocean
19	Aquaculture
20	Glacier
Not observed	

3.2.5 Soil

The basins' soils are characterized using data from the "Soil Maps of the Colombian Territory—scale 1:100,000," published on the Open Data portal by the Agustín Codazzi Geographic Institute (IGAC), accessible at https://geoportal.igac.gov.co/contenido/datos-abiertos-agrologia. This source is selected for its official status and accurate representation of Colombian territory.

According to IGAC, these maps display the spatial distribution of soil characteristics, determined through general surveys conducted in each department. The maps provide descriptions and interpretations of soil-forming environments, the physical, chemical, mineralogical, and morphological properties of the soil, its taxonomy, drainage capacity, and spatial distribution. The soil types featured in the Soil Maps of the Colombian Territory include: Alfisols, Andisols, Aridisols, Coal Mine Pit, Entisols, Eroded Miscellaneous, Histosols, Inceptisols, Mollisols, Oxisols, Perpetual Snow, Rocky Miscellaneous, Spodosols, Ultisols, Urban areas, Vertisols, Water Bodies

3.3 Climatic index, hydrological signatures and physiographic characteristics

3.3.1 Climatic index

The Aridity Index ($\emptyset$) is calculated as the ratio of total annual potential evapotranspiration to total annual precipitation. The Aridity Index is related to climatic regimes according to the classification presented by Arora (2002) and Ponce (Ponce et al., 2000), in which the following Aridity Index defines arid, semi-arid, sub-humid, and humid regions ranges: $12 > \emptyset \geq 5$; $5 > \emptyset \geq 2$; $2 > \emptyset \geq 0.75$; and $0.75 > \emptyset \geq 0.375$, respectively.

The annual frequency of high and low precipitation days is estimated as the number of days in the year with high precipitation (Pr > 5 times the mean daily precipitation) and the number of dry days (Pr < 1 mm/day), respectively. The reported values correspond to the average estimates based on the available data. Similarly, the duration of high precipitation and dry periods per year is measured as the longest consecutive sequence of high precipitation days and dry days. These values also represent the average estimates from the available data.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

3.3.2 Hydrological signatures

Hydrologic signatures can be key in understanding streamflow patterns by capturing variations like flood peaks and baseflow. They assist in analyzing watershed processes by linking runoff generation and groundwater contributions to physical characteristics. Signatures can improve calibration and evaluation in hydrological modelling by ensuring models replicate key hydrological behaviors. Moreover, they guide regionalization studies, enabling predictions in ungauged basins by transferring knowledge from similar watersheds (McMillan, 2021).

The CAMELS-COL dataset features 11 hydrological signatures for each of the 347 catchments, being computed according to Addor et al. (2018) and described in Table 1, ensuring consistency with previous CAMELS datasets. The annual frequency of high and low flow days is estimated as the number of days in the year with high flow (Q > 9 times the median daily flow) and the number of low flow days (Q < 0.2 times the mean daily flow), respectively. The presented values reflect the average estimates based on the available data. Similarly, the duration of high and low flow events is determined as the longest consecutive number of days with high and low flows. The presented values also represent the average estimates from the available data.

3.3.3 Physiographic Characteristics

The estimation of the physiographic characteristics of the catchment is performed using the Whitebox library. The concentration times are estimated using the equations presented by Silveira (2005), with the equivalent slope applied in the estimation. The equivalent slope is used to represent the mean slope of the main river more accurately. On the other hand, the form factor and the compactness index are estimated using the equations presented by Hipólito and Carmo (2017). Table 3 presents the equations used to estimate the physiographic characteristics.

Table 3 Physiographic characteristics estimated

Geometric characteristic	Equation	Comments
Drainage area (km^2)	-	Estimated whit Whitebox library
Perimeter (km)	-	Estimated whit Whitebox library
Gravelius compactness index	$K_c = \frac{0.282\,P}{\sqrt{A}}$	-
Form factor	$K_f = \frac{A}{L^2}$	-
Strength of the main channel (L)	-	Estimated whit Whitebox library
Concentration time		
Kirpich	$T_c = 0.0663 L^{0.77} S^{-0.385}$	Typically for a small (approximately 0.5 km²), forested watershed in a mountainous region (Kirpich, 1940)
Ven te Chow	$T_c = 0.160 L^{0.64} S^{-0.320}$	t is a modification of the Kirpich equation, recommended for watersheds with mixed land cover, including some urban development but predominantly rural, with an area greater than 0.2 km² (Chow, 1964)
Johnstone	$T_c = 0.462 L^{0.50} S^{-0.250}$	This method was defined with data from 19 rural basins in the USA, with an area range of 64.8–4206 km² (Kaufmann de Almeida et al., 2017).
Corps Engineers	$T_c = 0.191 L^{0.76} S^{-0.19}$	It is commonly used in larger catchments, particularly those with complex flow networks, such as large river basins with multiple tributaries and mixed land cover. (Kaufmann de Almeida et al., 2017)
Equivalent slope	$S = \left(\frac{\sum L}{\sum \left(\frac{L_i}{\sqrt{I_i}} \right)} \right)^2 ; s_i = \frac{\Delta H_i}{L_i}$	s_i slope of segment i (m/m) L_i length of of segment i (m) ΔH_i height difference of segment i (m) L Length of main chanel (m) S Equivalent slope (m/m)

L: Length of the min channel (km), T_c: Concentration time (hours). A: Area (km^2)
P Perimeter (km)

The compactness coefficient and the form factor are fundamental indicators of a watershed's susceptibility to flooding. The compactness coefficient varies with the shape of the watershed, being higher in more irregular basins (Zävoianu, 1985). Its minimum value is 1, representing a perfectly circular basin, which promotes a greater concentration of surface runoff and,

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

consequently, a higher risk of flooding. Values between 1.25 and 1.50 indicate a medium to high tendency for flooding, while coefficients above 1.50 signal a lower susceptibility to these events (Villela and Mattos, 1975). On the other hand, the form factor expresses the relationship between the basin's shape and the generation of concentrated surface runoff. Values close to 1 indicate more circular basins, where runoff converges quickly to the outlet, increasing the risk of flooding. Lower values reflect more elongated basins, characterized by less concentrated runoff and a lower tendency for flooding (Villela and Mattos, 1975).

The time of concentration equations used in this study are selected because they are the most frequently cited in official technical documents in Colombia (e.g., CORCUENCAS (2014) and Invias (2009)). Due to the variability in the applicability of these equations, specific criteria have been established for estimating concentration times. These criteria consider the applicability range of each equation (Table 3) to ensure accurate and context-appropriate calculations. Thus, for areas smaller than 0.2 km², the time of concentration is calculated using the Kirpich equation. For areas between 0.2 and 0.5 km², the time of concentration is determined as the average of the values obtained using the Kirpich and Ven Te Chow equations. For areas between 0.5 and 64.8 km², the Ven Te Chow equation is applied. For areas ranging from 64.8 to 4206.1 km², the concentration time is calculated using the Johnstone equation, and for areas larger than 4206.1 km², the concentration time is determined using the Corps of Engineers method. These estimates serve only as an approximation of the time of concentration. Future studies could apply other equations reported in the literature; however, all the results obtained are present in the dataset.

3.3.3.1 Curve Number (CN)

The curve number (CN) is widely used in hydrology to estimate direct rainfall runoff, particularly in water resources management studies and stormwater control structure design. Developed by the U.S. Natural Resources Conservation Service (NRCS), the CN is based on soil characteristics, land use, hydrological conditions, and land cover. It plays a crucial role in modeling a watershed's hydrological response to precipitation events, helping to assess the amount of water that will infiltrate, store, or run off the surface. The CN was calculated using the ArcHydro Tools "Generate CN Raster" tool (yellow.esri.com - /archydro/), designed to generate CN rasters by integrating spatial data on soil, cover, and hydrological parameters. This tool automates the assignment of CN values in each cell, ensuring spatial coherence and homogeneous resolution. The inputs and adaptations made for its application are described below.

The first input is the DEM, whose main function is to define the spatial extent and resolution of the output raster. Previously, the DEM underwent a "Fill" correction process to remove anomalies from empty cells, ensuring a continuous and accurate representation of the raster layer. It is important to note that the DEM is not used to adjust CN values as a function of elevation, but only as a spatial reference to align and rasterize the input layers that are in vector format.

The second input corresponded to the soil layer in vector format obtained from IGAC, which was described previously. This dataset describes the drainage capacity and the soil characteristics necessary to classify the different soil types in the

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

hydrological groups (A, B, C, D). The classification is made using the data available in the attribute table, based on the specifications of the hydrological groups described by the United States Department of Agriculture (Soil Survey Staff, 2016). In this context, hydrological groups are assigned according to the following criteria: well-drained or excessively drained soils correspond to hydrological group A; either or moderately drained to hydrological group B; poorly drained to hydrological group C; while urban areas, bodies of water or soils with very poor drainage are classified in hydrological group D.

The third input corresponded to the land cover in raster format provided by MapBiomas, which is converted to vector format to guarantee compatibility with the "Generate CN Raster" tool. Subsequently, a direct relationship is established between the land cover of MapBiomas and the coverage classes proposed by the United States Army Corps of Engineers in the HEC-HMS manual (Table 4) (USACE, 2023), generated a new land cover denominated in the present work as “Equivalent land cover”.

The three previously described variables are processed in the "Generate CN Raster" tool, which provided a raster of 30 meters of spatial resolution, in which each cell represents a CN value adjusted according to the type of land cover and the assigned hydrological group.

Table 4 Land Cover Equivalents Between MapBiomas and USACE Classifications with the Proposed CN by the USACE.

Land Cover MapBiomas (ID MapBiomas)	USACE land cover (ID USACE)	A	B	C	D
River, lake, or ocean (33)	open_water (11)	98	98	98	98
	developed_open (21)	49	69	79	84
	developed_low (22)	57	72	81	86
	developed_med (23)	61	75	83	87
Infrastructure (24)	developed_high (24)	81	88	91	93
Beach,dune and sand spot (23) Other non-vegetated area (25) Not observed (27) Rocky outcrop (29) Mining (30) Glacier (34)	barren_land (31)	78	86	91	93
Forest (3)	deciduous_forest (41)	45	66	77	83
	evergreen_forest (42)	25	55	70	77
	mixed_forest (43)	36	60	73	79
Other non forest formation (13) Wooded sand vegetation (49) Herbaceous sand vegetation (50)	Shrub (52)	55	72	81	86
Grasslands/herbaceous (12)	Grassland (71)	50	69	79	84

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Open Access Earth System Science Data Discussions

Land Cover MapBiomas (ID MapBiomas)	USACE land cover (ID USACE)	A	B	C	D
	pasture (81)	49	69	79	84
Forest plantation (9) Mosaic of agriculture and pasture (21) Palm oil (35)	cultivated_crops (82)	67	78	85	89
Mangrove (5) Flooded forest (6)	woody_wetlands (90)	30	58	71	78
Wetland (11) Aquaculture (31) Hypersaline tidal flat (32)	herbaceous_wetlands (95)	30	58	71	78

3. Results

3.1 Catchment and physiographic characteristics

The analyzed catchments exhibit diverse drainage areas. 41% have drainage areas between 28 and 525 km², predominantly located in the mountainous zones of the hydrological regions. 24% of the catchments fall within the 525–1650 km² range, 23,6% have areas between 1650 and 12425 km². The remaining 11.2% correspond to catchments exceeding 12425 km², located predominantly in the Magdalena, Orinoco, and Amazonas regions (Figure 3a).

As mentioned, the compactness coefficient and the form factor are used to analyze a watershed's susceptibility to flooding. The results obtained for these indices are consistent and lead to the same conclusion: the studied basins exhibit low susceptibility to flooding. In this context, it is important to highlight that all basins have a compactness coefficient greater than 1.5 and a form factor less than 0.5. On the other hand, 16% of the analyzed catchments have concentration times ranging between 4 and 24 hours. Only three catchments present concentration times below 10 hours. However, it is important to note that these are small basins in the Andean region (Station code: 26017040, 35097010, and 35087030). A larger proportion, 46%, fall within 24 to 55 hours, while 25% of the catchments exhibit concentration times between 55 and 105 hours, and 13% exceed 105 hours (Figure 3b).

Among the analyzed catchments, 53% exhibit CN values between 60 and 70 (Figure 3c), indicating moderate infiltration capacity and lower runoff generation, typically associated with forested or well-vegetated areas (USDA, 2021). This condition is present in 46% of the stations in the Orinoco region, in 91% of the stations of the Amazon and Caribbean regions, in 53% of the stations of the Pacific region, and in 46% of the stations of the Magdalena region. In catchments where CN values are below 70, forest formations are generally the predominant land cover, which helps explain the observed patterns (Figure 4).

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

It is also worth highlighting that the Magdalena and Caribbean regions are largely dominated by Inceptisols (Figure 6). When well-preserved—especially under native forest cover—these soils play a key role in regulating streamflow and minimizing runoff. However, their capacity to perform this function is significantly reduced when replaced by agriculture or pasture (Mello et al., 2025). Thus, the prevalence of Inceptisols, along with the conservation of natural vegetation, likely contributes to the lower CN values observed in these regions.

On the other hand, 43% of the catchments have CN values ranging from 70 to 80 (Figure 3e), which are typically associated with agricultural or mixed land cover, where runoff is more pronounced but still not excessive (USDA, 2021). Most of these catchments correspond to the larger basins of the Magdalena region, where land cover is predominantly a mosaic of agriculture and pasture, contributing to higher CN values. Although these findings further reinforce that Colombia's landscape comprises forest formations, agricultural lands, and livestock areas, it also sheds light on the vulnerability of Colombia's headwater to maintain water yield services as agriculture pressure intensifies in the region (Arias and Barragán, 2020).

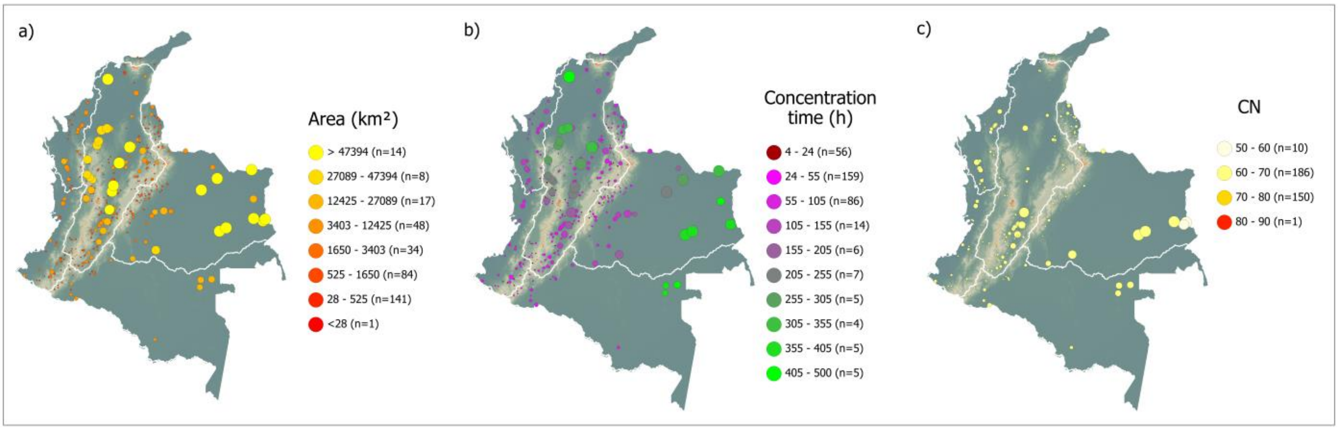

Figure 3 Physiographic characteristics of the 347 selected catchments. The size of the circles is proportional to the size of the catchment. A) Range area of the basins, b) Concentration time, c) Curve Number.

3.2 Land cover assessment

Distinctive land use and cover patterns were identified among the analyzed catchments. Forest formations predominate in 43% of the catchments, covering more than 50% of their area. This high forest coverage plays a key role in regulating hydrological processes, favoring infiltration, reducing surface runoff, and promoting water retention, which can contribute to lower susceptibility to flooding (Kan et al., 2019). On the other hand, in 37% of the catchments, agricultural and livestock activities occupy more than half of the area principally in the Magdalena region (Figure 4). These land uses tend to increase runoff due to soil compaction that, increases bulk density and penetration resistance and reduces macroporosity, infiltration rates and hydraulic conductivities, amplifying peak flows and influencing sediment transport (Martínez and Zinck, 2004; Alegre and Cassel, 1996).

Additionally, 86% of the catchments have less than 20% of their area associated with natural non-forest formations, indicating that these landscapes have a minimal role in the general hydrological dynamics. Non-vegetated areas represent less than 10% of all catchments. The water bodies occupied less of 10% of the basin's area in 99% of the cases. Furthermore, in 93% of the catchments, the land cover was fully identified, while in the remaining 7%, less than 10% of the area could not be classified by the MapBiomas project (see FigureS1 in supplementary material).

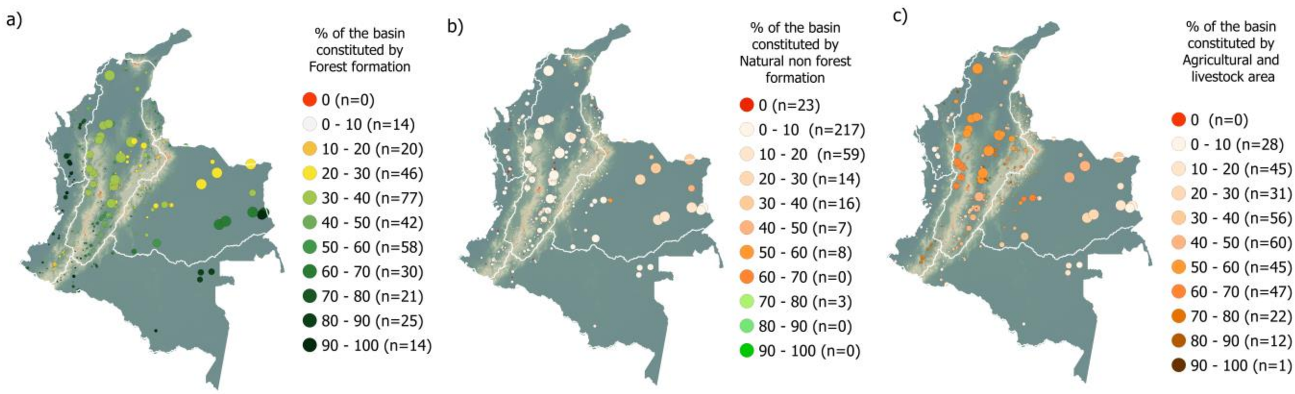

Figure 4 Land Cover of the 347 selected catchments. The size of the circles is proportional to the size of the catchment. The next land cover is considered: A) Forest formation, b) Natural nonforest formation, c) Agriculture and livestock area.

3.3 Geology assessment

The results show that in 37% of the analyzed basins, sedimentary rocks, formed by weathering pre-existing rocks (Crook and Beck, 2024), constitute more than 50% of the lithological composition (Figure 5). This lithology is present in all hydrological regions and is particularly predominant in the valleys of the Magdalena, Orinoco and Caribbean regions. Due to the heterogeneous nature of this group, which includes mudstone, poorly consolidated conglomerates, sandstones with ferruginous and clayey matrices, siltstones, and other sedimentary formations. The second predominant lithology corresponds to plutonic rocks, which comprise more than 50% of the lithological composition in 9% of the analyzed basins (Figure 5). These rocks are mainly found in the valleys of the Magdalena and Orinoco regions. Due to their coarse grain texture and low primary porosity, they have limited infiltration capacity (Watson, 1990). However, when intensely fractured, they can allow water infiltration and significantly contribute to aquifer recharge.

Metamorphic rocks are the third most common lithology, composing more than 50% of the material in 5% of the analyzed basins, and are predominantly found in the valleys of the Magdalena region (Figure 5). These rocks, formed under high pressure and temperature, generally have low primary porosity, which limits infiltration (USGS, 2023). In this sense, basins dominated

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

by metamorphic rocks may have reduced water retention capacity without fracturing, limiting its regularization capacity. On the other hand, 1% of the analyzed stations have a lithological composition dominated by more than 50% volcanic rocks (Figure 5), primarily located in the volcanic region of Colombia, especially near the Puracé, Sotará, Tolima, and Ruiz volcanoes. Characterized by rapid cooling and a fine-grained texture, volcanic rocks typically have low permeability, which may limit infiltration and water storage (Jerram, 2021). These geological characteristics highlight the limited capacity of water storage and release in the volcanic regions (In the Colombian case, primarily in the Magdalena region), which impacts streamflow (mainly in the recession period).

The smallest lithological material in the hydrographic basins includes unconsolidated deposits, hypabyssal rocks, and volcaniclastic rocks (Figure 5). Unconsolidated deposits, found across all hydrological regions, are composed of poorly cemented materials, which grant them high porosity and permeability, facilitating infiltration, aquifer recharge, and water storage (Wahlstrom, 1974). However, these characteristics depend on the other lithological compositions present. On the other hand, hypabyssal rocks, predominantly located in the Magdalena region, crystallize at intermediate depths and display medium-grained textures. Their permeability can be variable, depending on their degree of fracturing (Macheyeki et al., 2020). Finally, volcaniclastic rocks found in the volcanic region of Colombia are primarily located in the mountainous areas of the Magdalena region. Composed of volcanic fragments of varying sizes, these rocks may have areas with high initial porosity, though their consolidation over time can reduce infiltration (White and Houghton, 2006).

Although we can have an overview of geology and water storage-release processes from the presented dataset, the geological diversity present in the main hydrographic basins of Colombia underscores the need for individualized analysis of each basin in future studies, intending to account for its specific characteristics.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

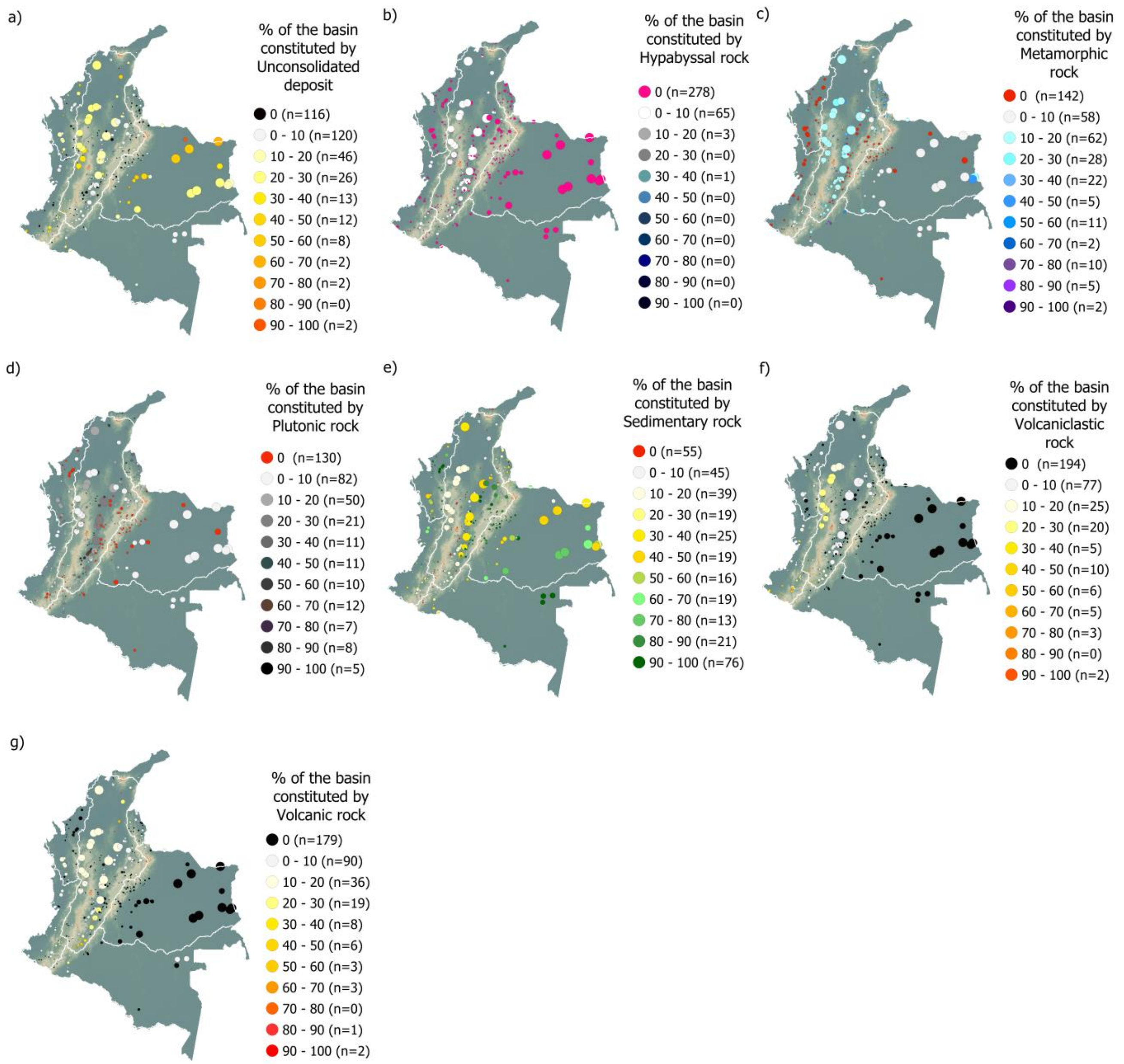

Figure 5 Geology of the 347 selected catchments. The size of the circles is proportional to the size of the catchment. The lithologies include A) unconsolidated deposits, B) hypabyssal rocks, C) metamorphic rocks, D) plutonic rocks, E) sedimentary rocks, F) volcanoclastic rocks, and G) volcanic rocks.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Open Access Earth System Science Data Discussions

3.4 Soil Assessment

The soil composition analysis in Colombia's watersheds revealed that the predominant soil types are Andisols, Entisols, and Inceptisols (Figure 6). The other soil types generally cover less than 10% of the watershed area (e.g., Alfisols and urban areas). They are sometimes absent, as with Spodosols and soils associated with open-pit coal mines. The spatial distribution of these soil types, distinct from Andisols, Entisols, and Inceptisols, is presented in the item "Spatial distribution of the Secondary soil typologies" of supplementary material (Figure S2).

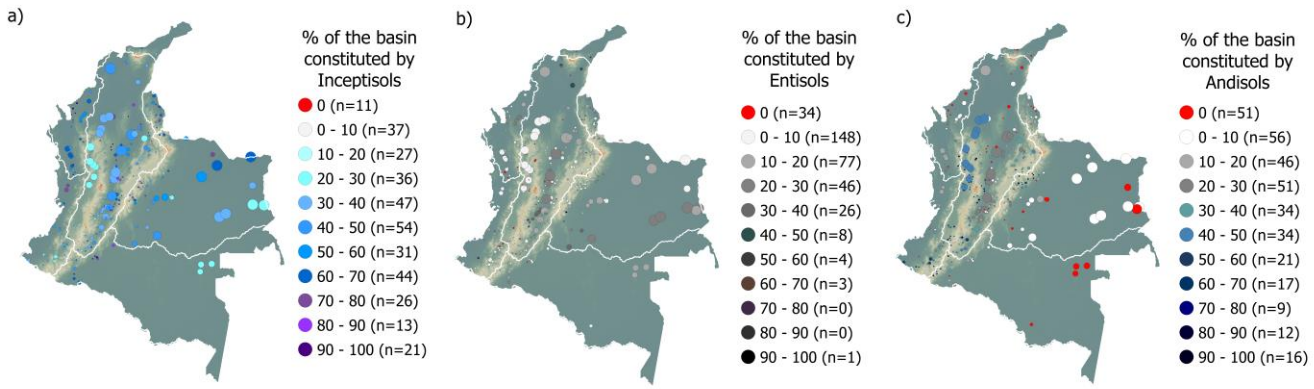

Figure 6 Main Soil Typologies of the 347 Selected Catchments: The Circle Size is Proportional to the Catchment Area: A) Inceptisols, B) Entisols, C) Andisols.

Andisols, characterized by their volcanic origin, exhibit good permeability and drainage capacity due to their fine and porous texture derived from volcanic ash (Uehara, 2005). It was identified that 22% of the watersheds have more than 40% of their area covered by Andisols, predominantly in areas near the volcanic region of the Andes, where the dominant lithology consists of volcanic and volcaniclastic rocks. Additionally, 34% of the watersheds have between 10% and 40% of their area occupied by Andisols, mainly in the Pacific and Magdalena hydrological regions.

Entisols, soils with limited pedological development that can exhibit variability in drainage capacity depending on the parent material, are commonly found in alluvial plains and areas with high erosion (Eswaran and Reich, 2005). It was observed that 23% of the analyzed watersheds have between 10% and 40% of their surface covered by this soil type, with a broad spatial distribution across all hydrological regions.

Finally, Inceptisols, which represent an intermediate stage in soil development, and whose drainage capacity depends on the underlying lithology (Foss et al., 1983), are the most prevalent in Colombia. It was observed that 39% of the studied watersheds have more than 40% of their surface covered by Inceptisols, with a concentrated distribution in the Magdalena, Pacific, and Caribbean regions, as well as in the western sector of the Amazon and Orinoco regions, near the Andes Mountain range.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Entisols and Inceptsols are poorly developed soils, with shallow horizons and steep slopes (Idaho, 2025). Therefore, they are most present in the headwater region of the Andes Mountains. These soils exhibit specific hydropedological characteristics that influence water infiltration, retention, and release. For instance, Inceptsols contribute significantly to water yield, but only when well-preserved (Mello et al., 2025) because root systems and organic matter accumulation enhance soil hydraulic properties, improving infiltration and percolation. In contrast, Entisols are shallow soils close to the parent material, which limits its water storage capacity (except in basins dominated by snowmelt processes). This makes them more susceptible to runoff formation and erosion (Kumar et al., 2000). On the other hand, Andisols improve water storage and release (Uehara, 2005), supplying streamflow in recession periods and increasing overall water yield. In summary, the hydrological responses of basins with predominance of Entisols and Inceptisols are more sensitive to land use and cover changes.

3.5 Climatological Assessment

The precipitation patterns of Colombian catchments exhibit significant spatial variability. Notably, river basins located in the western Caribbean and Pacific regions register the highest daily precipitation values (Pr > 10 mm/day). In the Amazonian catchments, the mean daily precipitation ranges between 8 mm/day and 10 mm/day. In comparison, most basins in the Magdalena region show daily precipitation values varying between 4 mm/day and 6 mm/day (Figure 7a).

Potential evapotranspiration is directly linked to temperature. In this regard, the catchments in the Pacific, Caribbean, Orinoco and Amazon regions, as well as the large basins of the Magdalena region, present the highest potential evapotranspiration values, ranging from 9 to 13 mm/day (Figure 7b). Temperature, on the other hand, shows an inverse relationship with altitude. The lowest mean daily temperatures (8°C - 17°C) are recorded in the high-mountain catchments of the Magdalena, while catchments with outlet points in the inter-Magdalena valleys predominantly experience temperatures between 17°C and 23°C. In contrast, the catchments in the Pacific, Caribbean and Amazon regions, as well as the large basins of the Magdalena region, register temperatures exceeding 23°CFinally, the highest temperatures are observed in the Orinoco region, where values exceed 29 °C (Figure 7c).

The analysis of the aridity index reveals that 74.7% of the analyzed catchments exhibit a subhumid climate regime, encompassing catchments from the Magdalena, Orinoco, and Amazon regions. Meanwhile, the eastern flank of the Andes and the Caribbean region display a semiarid climate regime, accounting for 20.7% of the analyzed catchments. Lastly, only 4.6% of the catchments fall within a humid climate regime, predominantly located on the western flank of the Caribbean and Pacific regions (Figure 7d).

Catchments with the highest occurrence of extreme precipitation days (≥ 5 times the mean daily precipitation) are concentrated in the mountainous areas of Magdalena region thus as Orinoco region. In the case of Magdalena, this pattern may be associated with orographic effects, as mountain ranges such as the Eastern Cordillera act as barriers to the southeast trade winds, promoting the formation of orographic rainfall (Ruiz et al., 2017; Espinoza et al., 2015). In contrast, the northern part of the country exhibits fewer extreme precipitation days, possibly due to the reduced influence of the Andes' orographic effects. Additionally, a direct spatial correlation is observed between the number of high-precipitation days and their duration.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Catchments with a high number of high-precipitation days also tend to experience longer consecutive sequences of such events (Figure 7 e and f).

The analyses also reveal that the Andes Mountain range catchments experience the highest number of dry days per year (>190 days/year). This can be attributed to the orographic shadow effect, where moist winds from the Atlantic and Pacific lose much of their humidity as they ascend the mountain slopes, reducing precipitation in the high-altitude areas of the Colombian Massif. Another relevant factor is the oscillation of the Intertropical Convergence Zone (ITCZ), which shifts north or south during certain periods of the year, reducing the influence of rain-generating systems in the Magdalena region (Loaiza Cerón et al., 2020).

Catchments with outlets in non-mountainous areas experience fewer dry days, ranging from 100 to 190 days. Notably, the country's largest river basins record the lowest dry days (47–100 days/year) (Figure 7g). River basins in the Pacific region experience the fewest consecutive dry days, likely due to the region's high and more evenly distributed rainfall throughout the year. In the Magdalena, Orinoco, and Amazon regions, the number of consecutive dry days ranges from 9 to 12. However, basins located in the mountainous areas of the Andes experience a higher number of consecutive dry days, exceeding 18 days (Figure 7h).

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

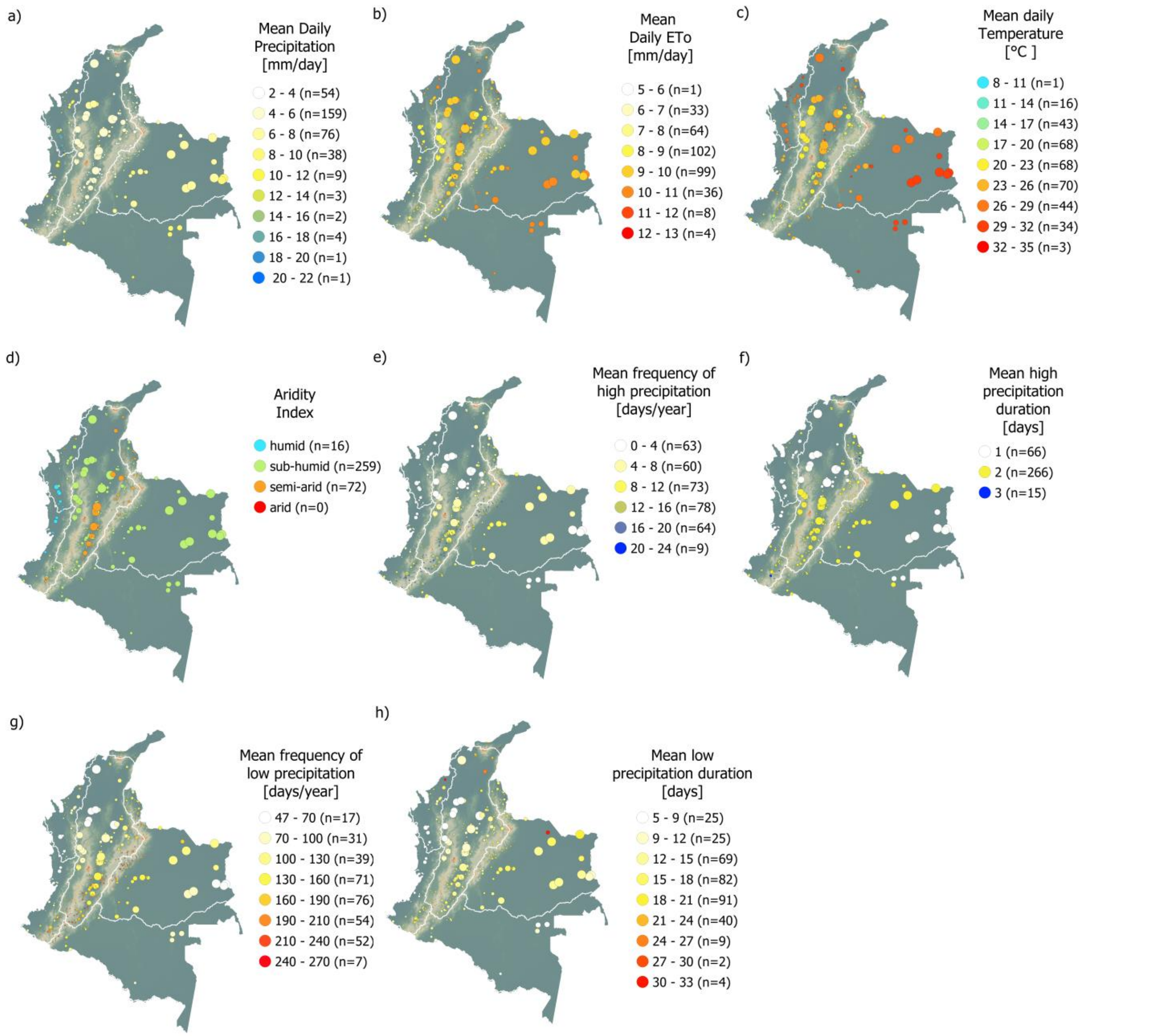

Figure 7 Climatic index of the 347selected catchments. The size of the circles is proportional to the size of the catchment. A) Mean daily precipitation in the catchment, b) Mean daily ETo in the catchment, c) Mean daily temperature in the catchment, d) Catchment aridity index, e) Mean frequency of high precipitations days, f) Mean duration of high precipitations days, g) Mean frequency of low precipitations days, h) Mean duration of low precipitations days.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

3.6 Hydrological Signatures

Figure 8a shows that 58% of the basins have an average daily flow of less than 3.3 mm/day, predominantly located in the Magdalena region. In contrast, basins in the Orinoco and Amazon regions exhibit average daily flows ranging between 3.3 and 6.3 mm/day, while those in the Pacific region present the highest values (> 6.3 mm/day). These findings indicate a relationship with daily precipitation, meaning regions with higher precipitation also tend to have higher streamflows. On the other hand, Q5 and Q95 follow a similar spatial pattern to the mean daily flow, with higher values in the Pacific basin and lower values in the Magdalena basin (Figure 8b and Figure 8c).

The runoff coefficient predominantly ranges between 0.3 and 0.7, with 56% of the analyzed basins falling within this range. In the Magdalena region, 57% of the catchments exhibit runoff coefficients between 0.4 and 0.7. The predominant runoff coefficient ranges from 0.5 to 0.7 in the Caribbean region, with 33% of the basins reporting values within this range (Figure 8d). In the Orinoco region, 35% of the basins recorded runoff coefficients greater than 1, while this proportion increases to 38% in the Pacific region and 41% in the Amazon region. In the Orinoco and Amazon regions, these basins are primarily located on the western slopes of the Andes Mountain range. Additionally, 37% of the basins in the Orinoco region present runoff coefficients between 0.5 and 0.7, while 32% of the Amazon basins report runoff coefficients between 0.7 and 1.0. Runoff coefficients exceeding 1 may be linked to an underestimation of precipitation by the CHIRPS satellite, as reported by Cavalcante et al. (2020) and Paredes-Trejo et al. (2017).

One third of Colombia’s river basins (34%), mainly in the Amazon and Orinoco regions, exhibited a streamflow elasticity ranging from 1.5 to 2.0, indicating that a 1% variation in precipitation could result in a 1.5% to 2.0% change in streamflow (Anderson et al., 2024). In the Orinoco region, elasticity ranged from 1.0 to 2.5. In the Magdalena region, values varied between 0.5 and 4.0, with a predominance of elasticities between 1.0 and 2.5 observed in 71% of the region stations. Meanwhile, in the Caribbean region, elasticities fluctuated between 0.5 and 4.0, with a predominant range of 1.0 to 2.5. In this region, 15% of the stations recorded elasticities between 1.0 and 1.5, 37% between 1.5 and 2.0, and 30% reported values above 2.0 (Figure 8e).

Regarding the flow duration curve, 53% of the basins exhibit a slope between 0.6 and 0.8, predominantly in the Magdalena, Pacific, and Caribbean regions. This suggests a balance between base flow and response to precipitation events, considering that a high slope (>0.8) indicates a flashy flow regime, while a low slope suggests a more damped flow (Chouaib et al., 2018; Sawicz et al., 2011). In contrast, basins with high slopes (>0.8) are found in the Amazon and Orinoco regions, as well as in the eastern flank of the Magdalena region, within the Andean Mountain range massif (Figure 8f).

In Colombia, 76% of river basins have a Base Flow Index (BFI) between 0.6 and 0.8, suggesting that groundwater resources play a crucial role in the stability, availability, and regulation of surface water sources. Furthermore, a correlation is observed between land cover and BFI, as basins with an index above 0.6 are predominantly covered by forests, which occupy more than 30% of the total basin area (Figure 8g).

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Considering that most of the country's basins exhibit an elasticity index greater than 1 and runoff coefficients predominantly below 0.7, any alteration in the precipitation regime, deforestation, or overexploitation of aquifers is expected to compromise Colombia's water stability and availability. Furthermore, the slope of the flow duration curve reinforces this vulnerability, especially in the Orinoco and Amazon basins, where approximately 52% of the country's groundwater reserves are concentrated (Aranguren-Díaz et al., 2024) and where high slopes of the flow duration curve (>0.8) prevail. This suggests a more variable flow regime, leading to greater susceptibility to changes in water availability in response to any alteration in precipitation, land use change, or aquifer overexploitation.

On the other hand, 61% of the basins analyzed did not record any high-flow days, while another 14% experienced only one high-flow day per year. About 15% of the basins had 2, 3, or 4 high-flow days per year, while the remaining 15% reported between 4 and 30 days of high-flow. It is observed that basins with more than 2 high-flow days per year are predominantly located on the eastern flank of the Andes Mountain range, in the transition zone between the Magdalena and Orinoco regions. A correlation is evident regarding the duration of these events with the total number of high-flow days. In this regard, 14% of the basins recorded only one consecutive day of high flows, two consecutive high flows account 14% of the basins as well, also 14% experienced between 3 and 5 consecutive days, and 21% of the basins registered events lasting more than 5 consecutive days (Figure S3 of supplementary material).

The number of low-flow days is higher than that of high-flow days. In this regard, basins in the western part of the country predominantly exhibit fewer low-flow days (0–5 days per year) compared to those in the eastern region, where values exceed 5 days per year. The highest number of low-flow days is primarily observed in the upper reaches of the Andes, particularly along the mountain range near the hydrological divide between the Magdalena, Orinoco, and Amazon regions (Eastern Cordillera), as well as in the Amazon and Orinoco regions. Additionally, a relationship is observed between the maximum number of consecutive low-flow days and the number of low-flow days per year, suggesting that basins experiencing frequent low-flow conditions also tend to have prolonged periods of consecutive low flows (Figure 8j and Figure 8k).

The observed hydrological signatures shed light on the vulnerability of most Colombian basins to changes in precipitation. This aligns with the previous statement regarding soil properties and how basin responses are driven by the land use and cover conditions. High elasticity values (mostly between 1.5 and 2.5) represent the sensitivity of streamflow to precipitation variability (Anderson et al., 2024). This is reinforced by the BFI values ranging from 0.6 to 0.8, which are lower than those observed in basins overlying productive aquifers (Collischonn and Dornelles, 2013). Thus, water storage contributions to streamflow appear to be constrained by soil water-holding capacity, depth, and release processes. This limited storage capacity is reflected in the slope of the flow duration curve (ranging mainly from 0.4 to 1.0), highlighting the basins' difficulty in sustaining streamflow (lower regulation). In summary, Colombian basins seem more sensitive to climate variability than to hydrological processes (e.g., water storage and release), i.e., climate change is likely to intensify extreme events such as floods and droughts in these regions.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

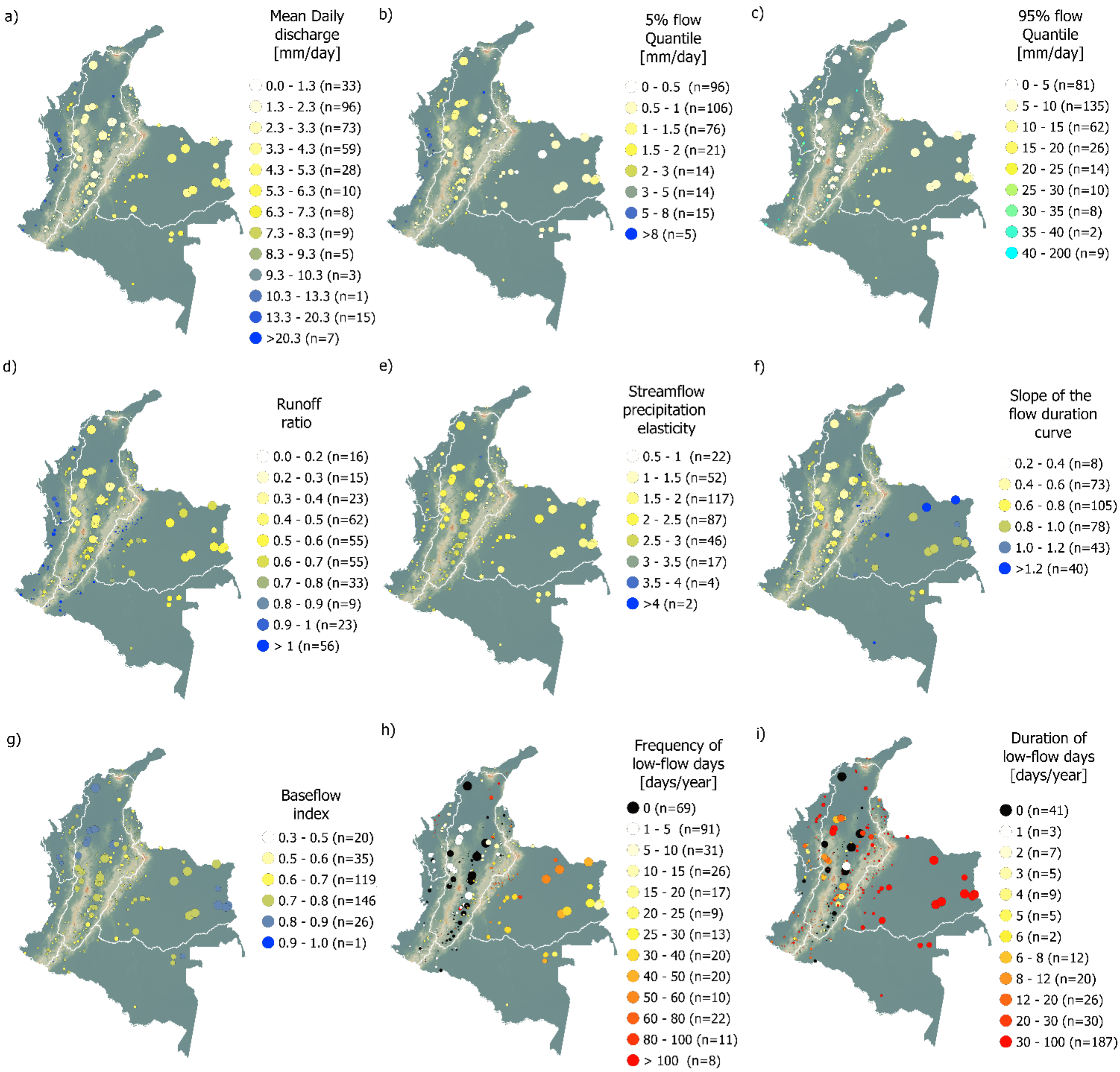

Figure 8 Hydrological signatures of the 347 selected catchments. The size of the circles is proportional to the size of the catchment. a) Mean daily discharge, b) Q5 (5th percentile daily flow), c) Q95 (95th percentile daily flow), d) Runoff ratio, e) Streamflow precipitation elasticity, f) Slope of the FDC, g) Baseflow index, h) Frequency of high-flow days, i) Duration of high-flow days, j) Frequency of low-flow days, k) Duration of low-flow days.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

4. Water Resources and Climate Change in Colombia

Colombia's agriculture and energy production are particularly sensitive to hydrological dynamics, making them dependent on both climate variability and water availability. For instance, coffee and rice crops, which are vital to both food security and export revenues, rely on rainfall seasonality and irrigation schedules. As climate change increases the rainfall spatiotemporal variability in Colombia, water availability in both time and space is likely to shift, impacting harvesting schedules and crop yields. Moreover, approximately 70% of Colombia's power generation comes from hydroelectricity (IEA, 2023), making it vulnerable to streamflow variability. In the Magdalena River Basin, projections under RCP4.5 and RCP8.5 scenarios show reduced dry season discharge in the headwater regions and increased inundation downstream due to intensified rainfall (Munar et al., 2023). This pattern is confirmed in the IDEAM (2024), which assessed the impacts of climate change across Colombia based on the CMIP6 and different emission scenarios. On the other hand, drying trends are expected in February (−0.60 mm/year) in the Caribbean region, while some departments are likely to experience increases up to 1.61 mm/year in November (Arregocés et al., 2024). However, the annual precipitation is likely to reduce by up to 44% under high-emission scenarios (IDEAM et al., 2024). Hydrological signatures (such as BFI and elasticity) highlight the sensitivity of these watersheds to climate change, as they are highly sensitive to precipitation, which in turn impacts streamflow (McMillan, 2020). BFI values indicate low to medium capacity of streamflow regularization (Figure 9g), largely due to the predominant presence of Inceptisols and Entisols, as well as the presence of porous and fractured aquifers (Figures 6 and 7). Therefore, the Magdalena and Caribbean regions are highly sensitive to climate change as their hydrological responses are more influenced by weather variability than hydrological stability.

Water availability is also dependent on the Páramo ecosystems, such as those in the Chingaza National Park. The Páramo ecosystem is over 3,000 m a.s.l and mostly in areas of Andisols (Figure 6). Andisols have good permeability and elevated porosity (Uehara, 2005), which together with the Páramo improves organic matter accumulation (due to low temperatures in high altitudes), increases water infiltration and storage, and regulates baseflow, configuring it as "natural sponges" (Herzog et al., 2011). Moreover, this ecosystem can intercept cloud and mist, increasing water input to the soil as occult precipitation (Cárdenas et al., 2017). However, this ecosystem is highly sensitive to warming and dry conditions (Cresso et al., 2020). Climate projections suggest that up to 52% of the Páramo area could become unsuitable, undermining ecosystem services such as groundwater recharge and carbon sequestration (Cresso et al., 2020). As much of the watersheds in the Andean region have elasticity > 1.0 (Figure 9e), they are sensitive to precipitation, and a slight change in rainfall can lead to intensified droughts or floods, as Páramos are likely to lose their ability to produce water (Cresso et al., 2020).

Mean temperatures are projected to increase by 0.7 to 2.5 °C under SSP2-4.5 in the mid-century across Colombia (IDEAM et al., 2024). The official Colombia's report (IDEAM et al., 2024) also projects an increase up to 5.0°C in the five hydrological regions of Colombia (Pacific, Magdalena, Orinoco, Amazon, and Caribbean) by the end of the century. Although this can be less important than precipitation changes, due to the high dependence of Colombia's watershed on precipitation (indicated by hydrological signatures), this increment potentially increases atmospheric water demand and pressures water availability. Irrigation schedule and demand changes with greater evapotranspiration to keep crop yield as observed in a densely irrigated area in Brazil (Rodrigues et al., 2024). In this sense, the projected precipitation variability and air temperature increment will likely increase water conflicts across Colombia, mainly in the watersheds that are dependent on the Andean and Páramo ecosystems' services. Moreover, the country's hydroclimate is heavily influenced by the El Niño–Southern Oscillation (ENSO), which may cause recurring anomalies such as intense droughts during El Niño and excessive flooding during La Niña (Rodriguez-Espinosa et al., 2025). However, the effects of large-scale climate anomalies in Colombia are still unclear, and more studies are necessary to improve our understanding of their impacts and the areas most affected.

CAMELS-COL provides enough information to improve continental-scale understanding on water resources response to climate change, decreasing uncertainties of current models, especially in underserved regions such as the Amazon and Orinoco basins, where data is sparse. Moreover, this dataset sheds light on the sensibility of Colombia's watersheds to weather variability, which can face more frequent floods and droughts depending on the increase or decrease of precipitation. This information can also support a coordinated national strategy for adaptive water governance, backed by scientific monitoring and updated design standards, is essential to reduce future climate risks. Moreover, it adds to the CAMELS-CH (Höge et al., 2023) and CAMELS-BR (Chagas et al., 2020) and enables advances on the South America hydrological understanding and response to climate change.

5. Conclusions

So far, large-sample hydrological studies in Colombia have lacked a comprehensive and easily accessible dataset. This study introduces CAMELS-COL, a new dataset that provides streamflow time series for 347 catchments nationwide. In addition to streamflow data, CAMELS-COL includes climatic series, physiographic characteristics, climatic indices, hydrological signatures, land cover, geology, and soil properties, primarily derived from official governmental sources.

The analysis of Colombian catchments highlights the diverse physiographic, hydrological, geological, and climatological factors influencing water dynamics across the country. In this sense, the analyzed basins exhibit significant variability in drainage area, with 41% ranging between 28 and 525 km², predominantly located in the mountainous zones of the Magdalena, Caribbean, Amazon, and Orinoco regions. The compactness coefficient and form factor suggest that most evaluated Colombian basins have low susceptibility to flooding. Additionally, basin concentration times vary widely, from less than 10 hours to over 105 hours.

Forest formations predominate in 43% of the catchments, and Agricultural and livestock land cover dominate 37%. In contrast, non-vegetated areas and water bodies have a less prominent role in the hydrological processes of the analyzed basins, as they represent less than 10% of the land cover total basin area in most cases. Regarding geological characteristics, sedimentary rocks dominate the lithological composition in 37% of the catchments. Plutonic rocks are the second most prevalent lithology, constituting more than 50% of the lithological composition in 9% of the catchments. Finally, metamorphic and volcanic rocks account for over 50% of the basin area in 5% and 1% of the analyzed catchments, respectively.

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Precipitation patterns in Colombian river basins exhibit high spatial variability. The highest daily values are observed in the Caribbean and Pacific regions, while the Magdalena region experiences lower daily precipitation. Potential evapotranspiration follows a direct relationship with temperature, with the highest values recorded in the Caribbean, Pacific, Amazon, and larger basins of the Magdalena.

An analysis of the aridity index reveals that 74.6% of the basins have a subhumid climate regime, while 20.7% exhibit semiarid conditions and only 4.6% are classified as humid. The distribution of extreme precipitation days indicates that the Magdalena basins, particularly those in mountainous areas, experience the highest frequency of heavy rainfall events while recording the greatest number of dry days.

From a hydrological perspective, the lowest mean daily streamflows are found in the Magdalena basins. In contrast, the highest values occur in the Pacific, Amazon, and Orinoco regions, aligning with precipitation distribution patterns. The runoff coefficient generally ranges from 0.3 to 0.7 across most basins; however, values exceeding 1 are observed in the Pacific, Orinoco, and Amazon regions, possibly due to underestimations in satellite-based precipitation data.

Flow elasticity in response to precipitation is highest in the Amazon and Orinoco, indicating a greater streamflow sensitivity to rainfall changes. The slope of the flow duration curve suggests that basins in the Orinoco and Amazon regions exhibit greater hydrological variability, making them more vulnerable to shifts in water availability caused by changes in precipitation patterns, deforestation, or overexploitation of aquifers. The Base Flow Index highlights the crucial role of groundwater in stabilizing and regulating surface water, particularly in basins with extensive forest cover. Finally, high-flow days are most frequent in the transitional zone between the Magdalena, Orinoco, and Amazon regions. In contrast, low-flow days are more predominant in the eastern part of the country.

The CN values of the analyzed catchments range from 50 to 82, with 54% falling between 60 and 70, indicating moderate infiltration capacity and relatively lower runoff generation. These catchments are mainly found in the Amazon, Orinoco, Pacific, and Caribbean hydrological regions, as well as in the mountainous areas of the Magdalena region, where the predominance of forest formations as the primary land cover explains the observed CN values. Conversely, another 43% of the catchments exhibit CN values between 70 and 80, typically associated with agricultural or mixed land cover, where runoff is more pronounced but not excessive. Most of these catchments correspond to the larger basins of the Magdalena region, where land cover consists mainly of a mosaic of agriculture and pasture, contributing to higher CN values.

The CAMELS-COL dataset represents a significant advancement in large-sample hydrology for Colombia, providing a standardized and accessible resource for analyzing the country's diverse hydrological, climatic, and physiographic characteristics. It enhances the benchmarking, calibration, and validation of hydrological models while supporting flood and drought monitoring efforts. Additionally, it contributes to climate change adaptation planning at the departmental level and enables global and regional hydroclimatic comparisons. Future research should integrate this dataset with dynamic hydrological models to improve water resource management strategies and assess the long-term impacts of climate change on Colombian river basins.

6. Data availability

The CAMELS-COL dataset is freely available at https://doi.org/10.5281/zenodo.15554735 (Jimenez et al., 2025). The files provided are (i) the Geologic_characteristics, Land cover characteristics, Soil characteristics, Climatic indices, Hydrological signatures and Physiograpic Characteristics in *.xlsx files (ii) the daily time series in zip files, (iii) the catchment boundaries of the catchments in zip file (iv) a readme file.

7. Author contributions

DAJ, BMB, and AFR initiated the investigation. JEM and AAD estimated the catchments' Curve Numbers (CN), while PHBS calculated the hydrological signatures. DAJ, AAD, and BQ performed the analysis of climatological information. FR and BMB designed the conceptual framework. All authors contributed to the interpretation of the results, the discussion of the findings, and the editing of the manuscript.

8. Competing interests

The authors declare that they have no conflict of interest.

9. Acknowledgments

The authors thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for the financial support provided through the doctoral scholarship (process number 14920), which has been essential for the development of this research. The authors also acknowledge the funding from Universidad del Rosario for the project “Extremos hidroclimatológicos en Colombia durante 1980 al 2100 – EXHIDROC”.

10. References

Addor, N., Do, H.X., Alvarez-Garreton, C., Coxon, G., Fowler, K., Mendoza, P.A., 2020. Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges. Hydrol. Sci. J. 65, 712–725. https://doi.org/10.1080/02626667.2019.1683182

Addor, N., Nearing, G., Prieto, C., Newman, A.J., Le Vine, N., Clark, M.P., 2018. A Ranking of Hydrological Signatures Based on Their Predictability in Space. Water Resour. Res. 54, 8792–8812. https://doi.org/10.1029/2018WR022606

Addor, N., Newman, A.J., Mizukami, N., Clark, M.P., 2017. The CAMELS data set: catchment attributes and meteorology for large-sample studies. Hydrol. Earth Syst. Sci. 21, 5293–5313. https://doi.org/10.5194/hess-21-5293-2017

Alegre, J.C., Cassel, D.K., 1996. Dynamics of soil physical properties under alternative systems to slash-and-burn. Agric. Ecosyst. Environ., Alternatives to Slash-and-Burn Agriculture 58, 39–48. https://doi.org/10.1016/0167-8809(95)00654-0

Alvarez-Garreton, C., Mendoza, P.A., Boisier, J.P., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J., Ayala, A., 2018. The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies – Chile dataset. Hydrol. Earth Syst. Sci. 22, 5817–5846. https://doi.org/10.5194/hess-22-5817-2018

Araghi, A., Martinez, C.J., Olesen, J.E., 2023. Evaluation of MSWX gridded data for modeling of wheat performance across Iran. Eur. J. Agron. 144, 126769. https://doi.org/10.1016/j.eja.2023.126769

Arias, G.A., Barragán, R.M., 2020. Análisis de la expansión de la frontera agrícola y propuesta de solución a eventuales conflictos en la localidad de Sumapaz. Universidad Distrial Francisco José de Caldas.

Arora, V.K., 2002. The use of the aridity index to assess climate change effect on annual runoff. J. Hydrol. 265, 164–177. https://doi.org/10.1016/S0022-1694(02)00101-4

Arregocés, H.A., Gómez, D., Castellanos, M.L., 2024. Annual and monthly precipitation trends: An indicator of climate change in the Caribbean region of Colombia. Case Stud. Chem. Environ. Eng. 10, 100834. https://doi.org/10.1016/j.cscee.2024.100834

Arregocés, H.A., Rojano, R., Pérez, J., 2023. Validation of the CHIRPS dataset in a coastal region with extensive plains and complex topography. Case Stud. Chem. Environ. Eng. 8, 100452. https://doi.org/10.1016/j.cscee.2023.100452

Arsenault, R., Brissette, F., Martel, J.-L., Troin, M., Lévesque, G., Davidson-Chaput, J., Gonzalez, M.C., Ameli, A., Poulin, A., 2020. A comprehensive, multisource database for hydrometeorological modeling of 14,425 North American watersheds. Sci. Data 7, 243. https://doi.org/10.1038/s41597-020-00583-2

Avila-Diaz, A., Guzmán-Escalante, J., Quesada, B., 2024. Gestión del agua en Colombia: ¿Estamos preparados para los retos del cambio climático? | Universidad del Rosario [WWW Document]. URL https://urosario.edu.co/revista-nova-et-vetera/columnistas/gestion-del-agua-en-colombia-estamos-preparados-para-los-retos (accessed 10.30.24).

Beck, H.E., Pan, M., Dutra, E., Miralles, D.G., 2022. Global 3-Hourly 0.1° Bias-Corrected Meteorological Data Including Near-Real-Time Updates and Forecast Ensembles.

Bedoya, M.A., Gonzales, M.A., 2021. Valoración de servicios ecosistémicos hidrológicos de los parques nacionales naturales de Colombia - Macrocuenca Magdalena - Cauca. Bogotá.

Cárdenas, M.F., Tobón, C., Buytaert, W., 2017. Contribution of occult precipitation to the water balance of páramo ecosystems in the Colombian Andes. Hydrol. Process. 31, 4440–4449. https://doi.org/10.1002/hyp.11374

Cepeda, A.E., Cañon, B.J., 2022. Performance of high-resolution precipitation datasets CHIRPS and TerraClimate in a Colombian high Andean Basin. Geocarto Int. 37, 17382–17402. https://doi.org/10.1080/10106049.2022.2129816

Chagas, V.B.P., Chaffe, P.L.B., Addor, N., Fan, F.M., Fleischmann, A.S., Paiva, R.C.D., Siqueira, V.A., 2020. CAMELS-BR: hydrometeorological time series and landscape attributes for 897 catchments in Brazil. Earth Syst. Sci. Data 12, 2075–2096. https://doi.org/10.5194/essd-12-2075-2020

Chow, V.T., 1964. Handbook of Applied Hydrology: A Compendium of Water-resources Technology.

Cohen, K.M., Finney, S.C., Gibbard, P.L., Fan, J.-X., 2022. The ICS International Chronostratigraphic Chart. Episodes 36, 199–204. https://doi.org/10.18814/epiiugs/2013/v36i3/002

CORCUENCAS, 2014. Formulación POMCA Río Luisa y Otros Directos al Magdalena Plan de Ordenación y Manejo de la Cuenca Hidrográfica, Hidrología. Ibagué.

Coxon, G., Addor, N., Bloomfield, J.P., Freer, J., Fry, M., Hannaford, J., Howden, N.J.K., Lane, R., Lewis, M., Robinson, E.L., Wagener, T., Woods, R., 2020. CAMELS-GB: hydrometeorological time series and landscape attributes for 671 catchments in Great Britain. Earth Syst. Sci. Data 12, 2459–2483. https://doi.org/10.5194/essd-12-2459-2020

Cresso, M., Clerici, N., Sanchez, A., Jaramillo, F., 2020. Future Climate Change Renders Unsuitable Conditions for Paramo Ecosystems in Colombia. Sustainability 12, 8373. https://doi.org/10.3390/su12208373

Crook, W.A.K., Beck, C.K., 2024. Sedimentary rock | Definition, Formation, Examples, & Characteristics | Britannica [WWW Document]. URL https://www.britannica.com/science/sedimentary-rock (accessed 2.5.25).

Delaigue, O., Guimarães, G.M., Brigode, P., Génot, B., Perrin, C., Soubeyroux, J.-M., Janet, B., Addor, N., Andréassian, V., 2024. CAMELS-FR dataset: A large-sample hydroclimatic dataset for France to explore hydrological diversity and support model benchmarking. Earth Syst. Sci. Data Discuss. 1–27. https://doi.org/10.5194/essd-2024-415

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., Ceccato, P., 2018. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Q. J. R. Meteorol. Soc. 144, 292–312. https://doi.org/10.1002/qj.3244

Espinoza, J.C., Chavez, S., Ronchail, J., Junquas, C., Takahashi, K., Lavado, W., 2015. Rainfall hotspots over the southern tropical Andes: Spatial distribution, rainfall intensity, and relations with large-scale atmospheric circulation. Water Resour. Res. 51, 3459–3475. https://doi.org/10.1002/2014WR016273

Eswaran, H., Reich, P.F., 2005. WORLD SOIL MAP, in: Hillel, D. (Ed.), Encyclopedia of Soils in the Environment. Elsevier, Oxford, pp. 352–365. https://doi.org/10.1016/B0-12-348530-4/00019-9

FAO, 2006. Evapotranspiración del cultivo. Guías para la determinación de los requerimiento de agua de los cultivos.

Foss, J.E., Moormann, F.R., Rieger, S., 1983. Chapter 10 Inceptisols, in: Wilding, L.P., Smeck, N.E., Hall, G.F. (Eds.), Developments in Soil Science, Pedogenesis and Soil Taxonomy. Elsevier, pp. 355–381. https://doi.org/10.1016/S0166-2481(08)70621-4

Fowler, K.J.A., Acharya, S.C., Addor, N., Chou, C., Peel, M.C., 2021. CAMELS-AUS: hydrometeorological time series and landscape attributes for 222 catchments in Australia. Earth Syst. Sci. Data 13, 3847–3867. https://doi.org/10.5194/essd-13-3847-2021

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci. Data 2, 150066. https://doi.org/10.1038/sdata.2015.66

Gómez, J., Montes, N.E., Marín, E., 2023. Mapa Geológico de Colombia 2023. Escala 1:1 500 000.

Gupta, H.V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., Andréassian, V., 2014. Large-sample hydrology: a need to balance depth with breadth. Hydrol. Earth Syst. Sci. 18, 463–477. https://doi.org/10.5194/hess-18-463-2014

Hargreaves, H.G., Samani, A.Z., 1985. Reference Crop Evapotranspiration from Temperature. Appl. Eng. Agric. 1, 96–99. https://doi.org/10.13031/2013.26773

Herzog, k. S., Martinez, R., Jorgense, M.P., Holm, T., 2011. Climate Change and Biodiversity in the Tropical Andes (1 No. 1).

Hipólito, R.J., Vaz, C., Álvaro, 2017. Hidrologia e recursos hídricos, 3rd ed. IJf press, Lisboa, Portugal.

Höge, M., Kauzlaric, M., Siber, R., Schönenberger, U., Horton, P., Schwanbeck, J., Floriancic, M.G., Viviroli, D., Wilhelm, S., Sikorska-Senoner, A.E., Addor, N., Brunner, M., Pool, S., Zappa, M., Fenicia, F., 2023. CAMELS-CH: hydro-meteorological time series and landscape attributes for 331 catchments in hydrologic Switzerland. Earth Syst. Sci. Data 15, 5755–5784. https://doi.org/10.5194/essd-15-5755-2023

Idaho, U., 2025. Inceptisols | Soil & Water Systems | University of Idaho [WWW Document]. URL https://www.uidaho.edu/cals/soil-orders/inceptisols (accessed 3.30.25).

IDEAM, PNUD, GEF-CBIT, 2024. Escenarios de Cambio Climático: De la cuarta comunicación nacional de Colombia. Bogotá.

IEA, 2023. Colombia 2023. Energy Policy Review (No. 1).

Invias, 2009. Manual de drenajes para carreteras, 1st ed. Colombia.

Jerram, D., 2021. Introducing Volcanology: A Guide to Hot Rocks, 2nd ed. Liverpool University Press. https://doi.org/10.2307/jj.12949053

Jimenez, D.A., Menesses, J.E., Solha, B.H.P., Avila-Diaz, A., Quesada, B., Brentan, B.M., Rodrigues, A.F., 2025. CAMELS-COL: hydrometeorological time series and catchment attributes for 347catchments in Colombia. https://doi.org/10.5281/zenodo.15554735

Kan, X., Cheng, J., Hu, X., Zhu, F., Li, M., 2019. Effects of Grass and Forests and the Infiltration Amount on Preferential Flow in Karst Regions of China. Water 11, 1634. https://doi.org/10.3390/w11081634

Kaufmann de Almeida, I., Kaufmann Almeida, A., Garcia Gabas, S., Alves Sobrinho, T., 2017. Performance of methods for estimating the time of concentration in a watershed of a tropical region. Hydrol. Sci. J. 62, 2406–2414. https://doi.org/10.1080/02626667.2017.1384549

Kirpich, Z.P., 1940. Time of concentration in small agricultural watersheds. Civ. Eng. 10, 362.

Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., Shalev, G., Matias, Y., 2023. Caravan - A global community dataset for large-sample hydrology. Sci. Data 10, 61. https://doi.org/10.1038/s41597-023-01975-w

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

Kumar, A., Murty, V.V.N., Kawachi, T., Maeda, S., Fazal, M.A., 2000. Runoff Simulation in Small Agricultural Watershed of Vertic Inceptisols of Indian Soils. J. Rainwater Catchment Syst. 5, 39–43. https://doi.org/10.7132/jrcsa.KJ00003257842

Lindsay, J.B., 2016. Whitebox GAT: A case study in geomorphometric analysis. Comput. Geosci. 95, 75–84. https://doi.org/10.1016/j.cageo.2016.07.003

Liu, J., Koch, J., Stisen, S., Troldborg, L., Højberg, A.L., Thodsen, H., Hansen, M.F.T., Schneider, R.J.M., 2024. CAMELS-DK: Hydrometeorological Time Series and Landscape Attributes for 3330 Catchments in Denmark. Earth Syst. Sci. Data Discuss. 1–30. https://doi.org/10.5194/essd-2024-292

Loaiza Cerón, W., Andreoli, R.V., Kayano, M.T., Ferreira de Souza, R.A., Jones, C., Carvalho, L.M.V., 2020. The Influence of the Atlantic Multidecadal Oscillation on the Choco Low-Level Jet and Precipitation in Colombia. Atmosphere 11, 174. https://doi.org/10.3390/atmos11020174

Loritz, R., Dolich, A., Acuña Espinoza, E., Ebeling, P., Guse, B., Götte, J., Hassler, S.K., Hauffe, C., Heidbüchel, I., Kiesel, J., Mälicke, M., Müller-Thomy, H., Stölzle, M., Tarasova, L., 2024. CAMELS-DE: hydro-meteorological time series and attributes for 1555 catchments in Germany. https://doi.org/10.5194/essd-2024-318

Macheyeki, A.S., Li, X., Kafumu, D.P., Yuan, F., 2020. Chapter 1 - Elements of exploration geochemistry, in: Macheyeki, A.S., Li, X., Kafumu, D.P., Yuan, F. (Eds.), Applied Geochemistry. Elsevier, pp. 1–43. https://doi.org/10.1016/B978-0-12-819495-9.00001-3

Mangukiya, N.K., Kumar, K.B., Dey, P., Sharma, S., Bejagam, V., Mujumdar, P.P., Sharma, A., 2024. CAMELS-INDIA: hydrometeorological time series and catchment attributes for 472 catchments in Peninsular India. Earth Syst. Sci. Data Discuss. 1–43. https://doi.org/10.5194/essd-2024-379

MapBiomas, 2022. MapBiomas – Colección 2 de la Serie anual de Mapas de Cobertura y Uso del Suelo de Colombia.

Martínez, L.J., Zinck, J.A., 2004. Temporal variation of soil compaction and deterioration of soil quality in pasture areas of Colombian Amazonia. Soil Tillage Res. 75, 3–18. https://doi.org/10.1016/j.still.2002.12.001

McMillan, H., 2020. Linking hydrologic signatures to hydrologic processes: A review. Hydrol. Process. 34, 1393–1409. https://doi.org/10.1002/hyp.13632

McMillan, H.K., 2021. A review of hydrologic signatures and their applications. WIREs Water 8, e1499. https://doi.org/10.1002/wat2.1499

Mello, C.R., Gúzman, J.A., Vieira, N.P.A., Viola, M.R., Beskow, S., Guo, L., Alvarenga, L.A., Rodrigues, A.F., 2025. Mitigating Severe Hydrological Droughts in the Brazilian Tropical High-land Region: A Novel Land Use Strategy under Climate Change. Int. Soil Water Conserv. Res. https://doi.org/10.1016/j.iswcr.2025.03.005

Munar, A.M., Mendez, N., Narvaez, G., Campo Zambrano, F., Motta-Marques, D., Lyra Fialho Brêda, J.P., Santos Fleischmann, A., Angarita, H., 2023. Modelling the climate change impacts on river discharge and inundation extent in the Magdalena River basin – Colombia. Hydrol. Sci. J. 68, 1286–1300. https://doi.org/10.1080/02626667.2023.2215932

Ponce, V.M., Pandey, R.P., Ercan, S., 2000. Characterization of Drought across Climatic Spectrum. J. Hydrol. Eng. 5, 222–224. https://doi.org/10.1061/(ASCE)1084-0699(2000)5:2(222)

Proutsos, N.D., Fotelli, M.N., Stefanidis, S.P., Tigkas, D., 2024. Assessing the Accuracy of 50 Temperature-Based Models for Estimating Potential Evapotranspiration (PET) in a Mediterranean Mountainous Forest Environment. Atmosphere 15, 662. https://doi.org/10.3390/atmos15060662

Quesada, B., Avila-Diaz, A., 2023. Urgency of understanding and responding to hydroclimatic impacts in Colombia. PLOS Clim. 2, e0000188. https://doi.org/10.1371/journal.pclm.0000188

Rodrigues, A.F., Brentan, B.M., Ottoni, M.V., Amorim, J.S., Mourão, M.A.A., Curi, N., Avanzi, J.C., de Mello, C.R., 2024. Has unsustainable groundwater use induced low flow regimes in the Urucuia Aquifer System? An urgent call for integrated water management. J. Environ. Manage. 370, 122979. https://doi.org/10.1016/j.jenvman.2024.122979

Rodriguez-Espinosa, C.A., Valencia, C.F., Ramirez, C.G., Valencia, C.D., 2025. Dynamic effect of climate change on flood damage cost in the Andean region of Colombia using an ARDL-ECM model and climate change projections. Sustain. Cities Soc. 125, 106309. https://doi.org/10.1016/j.scs.2025.106309

Romero-Hernández, C.M., Avila-Diaz, A., Quesada, B., Medeiros, F., Cerón, W.L., Guzman-Escalante, J., Ocampo-Marulanda, C., Rodrigues Torres, R., Zuluaga, C.F., 2024. Bias-corrected high-resolution precipitation datasets

https://doi.org/10.5194/essd-2025-200
Preprint. Discussion started: 23 June 2025

assessment over a tropical mountainous region in Colombia: A case of study in Upper Cauca River Basin. J. South Am. Earth Sci. 140, 104898. https://doi.org/10.1016/j.jsames.2024.104898
Ruiz, M.J.F., Serna, C.J., Zapatas, L.H.J., 2017. Atlas de viento de Colombia.
Silveira, A., 2005. Desempenho de Fórmulas de Tempo de Concentração em Bacias Urbanas e Rurais. Rev. Bras. Recur. Hídricos 10, 5–29. https://doi.org/10.21168/rbrh.v10n1.p5-29
Uehara, G., 2005. VOLCANIC SOILS, in: Hillel, D. (Ed.), Encyclopedia of Soils in the Environment. Elsevier, Oxford, pp. 225–232. https://doi.org/10.1016/B0-12-348530-4/00031-X
USACE, 2023. Creating a Curve Number Grid and Computing Subbasin Average Curve Number Values [WWW Document]. URL https://www.hec.usace.army.mil/confluence/hmsdocs/hmsguides/gis-tutorials-and-guides/creating-a-curve-number-grid-and-computing-subbasin-average-curve-number-values (accessed 2.13.25).
USDA, U.S.D. of A., 2021. Part 650 Engineering Field Handbook National Engineering Handbook, in: Estimating Runoff Volume and Peak Discharge. pp. 650–2.18.
USGS, 2023. What are metamorphic rocks? | U.S. Geological Survey [WWW Document]. URL https://www.usgs.gov/faqs/what-are-metamorphic-rocks (accessed 2.5.25).
Valencia, S., Marín, D.E., Gómez, D., Hoyos, N., Salazar, J.F., Villegas, J.C., 2023. Spatio-temporal assessment of Gridded precipitation products across topographic and climatic gradients in Colombia. Atmospheric Res. 285, 106643. https://doi.org/10.1016/j.atmosres.2023.106643
Villela, S.M., Mattos, A., 1975. Hidrologia aplicada. McGraw-Hill, São Paulo.
Wahlstrom, E.E. (Ed.), 1974. Chapter 2 - Rocks and Unconsolidated Deposits: Origin, Composition, and Properties, in: Developments in Geotechnical Engineering, DAMS, DAM Foundations, and Reservoir Sites. Elsevier, pp. 31–51. https://doi.org/10.1016/B978-0-444-41236-2.50007-7
Watson, J.V., 1990. PLUTONIC ROCKSPlutonic rocks, in: Petrology. Springer US, Boston, MA, pp. 464–465. https://doi.org/10.1007/0-387-30845-8_191
White, J.D.L., Houghton, B.F., 2006. Primary volcaniclastic rocks. Geology 34, 677–680. https://doi.org/10.1130/G22346.1
Wu, Q., Ramirez Avila, J.J., Yang, J., Ji, C., Fang, S., 2024. High-resolution Annual Dynamic dataset of Curve Number from 2008 to 2021 over Conterminous United States. Sci. Data 11, 207. https://doi.org/10.1038/s41597-024-03044-2
Zävoianu, I. (Ed.), 1985. Chapter VI Basin Shape, in: Developments in Water Science, Morphometry of Drainage Basins. Elsevier, pp. 103–113. https://doi.org/10.1016/S0167-5648(08)70418-0