Data description to "CAMELS-IND: hydrometeorological time series and catchment attributes for 228 catchments in Peninsular India"

Nikunj K. Mangukiya*[1], Kanneganti Bhargav Kumar*[2], Pankaj Dey[1], Shailza Sharma[2], Vijaykumar Bejagam[1], Pradeep P. Mujumdar[2,3], Ashutosh Sharma[1,4]

* Equal contribution

[1] Department of Hydrology, Indian Institute of Technology Roorkee, Roorkee, 247667, Uttarakhand, India

[2] Department of Civil Engineering, Indian Institute of Science, Bangalore 560012, Karnataka, India

[3] Interdisciplinary Centre for Water Research, Indian Institute of Science, Bangalore 560012, Karnataka, India

[4] International Centre of Excellence for Dams, Indian Institute of Technology Roorkee, Roorkee, 247667, Uttarakhand, India

Correspondence to: Ashutosh Sharma (ashutosh.sharma@hy.iitr.ac.in)

Data Description

This is a data description summary for the CAMELS-IND data set that corresponds to the official CAMELS-IND publication (https://doi.org/10.5194/essd-17-461-2025). The summary provides an overview of the content of the dataset in this repository and, therefore, comprises the tables shown in the article. For details about the data sources and further explanations, refer to the respective sections in the article.

Following is the file structure of the "*CAMELS_IND_All_Catchments.zip*".

```
└── CAMELS_IND_All_Catchments.zip /
    ├── attributes_csv
    ├── attributes_txt
    ├── catchment_mean_forcings
    ├── shapefile_catchment
    ├── streamflow_timeseries
    ├── CAMELS_IND_data_description.pdf
    ├── change_logs_and_disclaimer.txt
    └── filter_catchments.py
```

1. Catchments shapefile and gauge locations

The folder '*shapefiles_catchment*' contains a GIS shapefile of major river basins, catchments, and gauge locations covered by CAMELS-IND. This comprises 472 catchments located in 15 major river basins (CWC basin code of 03 to 17) of Peninsular

India. Each sub-folder within the '*shapefile_catchment*' contains basin shapefile, gauge station shapefile within the basin, and catchment shapefile within the basin. The '*shapefile_catchment*' also includes sub-folder '*merged*', which has the shapefile of all gauge stations and the shapefile of all catchments in single files.

The file structure of the '*shapefile_catchment*' folder is as follows:

```
└── shapefile_catchment/
    ├── 03_godavari/
    │   ├── 03_godavari.shp
    │   ├── 03_godavari_stations.shp
    │   └── 03_godavari_catchments.shp
    ├── 04_krishna/
    ├── 05_cauvery/
    ├── 06_subernarekha/
    ├── 07_brahmani_baitarni/
    ├── 08_mahanadi/
    ├── 09_pennar/
    ├── 10_mahi/
    ├── 11_sabarmati/
    ├── 12_narmada/
    ├── 13_tapi/
    ├── 14_wfr_tapi_tadri/
    ├── 15_wfr_tadri_kanyakumari/
    ├── 16_efr_mahanadi_pennar/
    ├── 17_efr_pennar_kanyakumari/
    └── merged/
        ├── all_gauge_stations.shp
        └── all_catchments.shp
```

2. Daily meteorological time series

The folder '*catchment_mean_forcing*' contains the mean meteorological time series for each catchment from 1 January 1980 to 31 December 2020.

The file structure of the '*catchment_mean_forcing*' folder is as follows:

```
└── catchment_mean_forcings/
    ├── 03001.csv
    ├── ...
    ├── ...
    ├── ...
    └── 17001.csv
```

Table 1. *Catchment mean meteorological variables available as daily time series in CAMELS-IND*

Variable name	Description	Unit	Data Source	Reference
prcp	precipitation	mm/day	IMD	(Pai et al., 2014)
tmax	maximum temperature	°C	IMD	(Srivastava et al., 2009)
tmin	minimum temperature	°C	IMD	(Srivastava et al., 2009)
tavg	averaged temperature	°C	(tmax + tmin) / 2	
srad_lw	surface downward long-wave radiation flux	w/m^2	IMDAA	(Rani et al., 2021)
srad_sw	surface downward short-wave radiation flux	w/m^2	IMDAA	(Rani et al., 2021)
wind_u	U-component of wind (10 m)	m/s	IMDAA	(Rani et al., 2021)
wind_v	V-component of wind (10 m)	m/s	IMDAA	(Rani et al., 2021)
wind	averaged wind speed (10 m)	m/s	$\sqrt{(wind\_u^2 + wind\_v^2)}$	
rel_hum	relative humidity (2 m)	%	IMDAA	(Rani et al., 2021)
evap_canopy	evaporation rate from canopy	mm/day	IMDAA	(Rani et al., 2021)
evap_surface	evaporation rate from the soil surface	mm/day	IMDAA	(Rani et al., 2021)
pet	potential evapotranspiration (from 1981 to 2020)	mm/day		(Singer et al., 2021)
pet_gleam	potential evapotranspiration	mm/day	GLEAM	(Miralles et al., 2011)
aet_gleam	actual evapotranspiration	mm/day	GLEAM	(Miralles et al., 2011)
sm_lvl1	soil moisture of layer 1 (0-0.1 m below ground)	kg/m^2	IMDAA	(Rani et al., 2021)
sm_lvl2	soil moisture of layer 2 (0.1-0.35 m below ground)	kg/m^2	IMDAA	(Rani et al., 2021)
sm_lvl3	soil moisture of layer 3 (0.35-1 m below ground)	kg/m^2	IMDAA	(Rani et al., 2021)
sm_lvl4	soil moisture of layer 4 (1-3 m below ground)	kg/m^2	IMDAA	(Rani et al., 2021)

3. Available observed and LSTM-based predicted streamflow time series

The folder '*streamflow_timeseries*' contains available observed and regionally trained LSTM-based hydrological model predicted streamflow for all 472 catchments covering a period from 1 January 1980 to 31 December 2020. The observed streamflow time series are compiled from the online portal, India – Water Resources Information System (India-WRIS; https://indiawris.gov.in/wris/#/). A regionally trained LSTM-based hydrological model is adopted from Mangukiya et al. (2023).

The file structure of the '*streamflow_timeseries*' folder is as follows:

```
└── streamflow_timeseries/
    ├── lstm_pred_streamflow.csv
    └── streamflow_observed.csv
```

The 'lstm_pred_streamflow.csv' and 'streamflow_observed.csv' file has columns of 'year', 'month', 'day', and streamflow time series (in m^3/s) for each catchment.

4. Catchment attributes

The folder '*attributes_csv*' and '*attributes_txt*' contain static catchment attributes in '*csv*' and '*text*' format, respectively. Each folder comprises eight files representing different categories of attributes as follows:

```
└── attributes_csv /
    ├── camels_ind_name.csv
    ├── camels_ind_topo.csv
    ├── camels_ind_clim.csv
    ├── camels_ind_hydro.csv
    ├── camels_ind_land.csv
    ├── camels_ind_soil.csv
    ├── camels_ind_geol.csv
    └── camels_ind_anth.csv
```

```
└── attributes_txt /
    ├── camels_ind_name.txt
    ├── camels_ind_topo.txt
    ├── camels_ind_clim.txt
    ├── camels_ind_hydro.txt
    ├── camels_ind_land.txt
    ├── camels_ind_soil.txt
    ├── camels_ind_geol.txt
    └── camels_ind_anth.txt
```

Table 2. *Catchment attributes provided in CAMELS-IND*

File name	Attributes description
camels_ind_name	7 attributes representing gauge name and identifier
camels_ind_topo	16 attributes representing location and topography
camels_ind_clim	42 attributes representing climate indices
camels_ind_hydro	73 attributes representing hydrological signatures
camels_ind_land	13 attributes representing land cover characteristics
camels_ind_soil	28 attributes representing soil characteristics
camels_ind_geol	7 attributes representing geological characteristics
camels_ind_anth	25 attributes representing anthropogenic influence in the catchment

Table 3. *Catchment-specific static attributes available in CAMELS-IND*

Attribute class	Attribute name	Description	Unit	Data Source / Reference
Gauge name and identifier	gauge_id	gauge station identifier (5-digit; first 2 digits are CWC basin code and last 3 digits are station number)	-	
	ghi_stn_id	unique id used to identify a station, 10 characters long	-	GHI (Goteti, 2023)
	cwc_site_name	name of the station	-	CWC
	river_basin	name of the river basin	-	
	cwc_river	river/tributary	-	
	ghi_group	ghi assigned group (G1 or G2)	-	GHI (Goteti, 2023)
	flow_availability	percentage duration for which streamflow data is available between 1980-2020	%	CWC

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
Location and topography	cwc_lat	latitude of the station	°	CWC
	cwc_lon	longitude of the station	°	
	ghi_lat	latitude of the ghi relocated station	°	GHI (Goteti, 2023)
	ghi_lon	longitude of the ghi relocated station	°	
	elev_mean	catchment mean elevation	m	SRTM DEM 90m
	elev_median	catchment median elevation	m	
	elev_min	catchment min elevation	m	
	elev_max	catchment max elevation	m	
	slope_mean	catchment mean slope	%	
	slope_median	catchment median slope	%	
	slope_min	catchment min slope	%	
	slope_max	catchment max slope	%	
	cwc_area	catchment drainage area	km^2	CWC
	ghi_area	catchment drainage area	km^2	GHI (Goteti, 2023)
	gauge_elevation	elevation of the gauging station	m	SRTM DEM 90m
	dpsbar	catchment mean drainage path slope	m/km	
Climate indices	p_mean	mean daily precipitation	mm/day	IMD
	p_max	maximum daily precipitation	mm/day	
	p_mean_anum	annual average total precipitation	mm	
	p_monthly_variability	variation in precipitation patterns throughout the year (higher values indicate greater variation)	-	
	p_annual_variability	variation in annual precipitation patterns (higher values indicate greater variation)	-	
	p_unif	how uniformly the precipitation is distributed in a year, 0 if the annual maximum precipitation is uniformly distributed throughout the year, 1 if the annual maximum precipitation occurs in a single day	-	
	high_prec_freq	frequency of high precipitation days (≥ 5 times the mean daily precipitation)	days/year	
	high_prec_dur	average duration of high precipitation events (number of consecutive days ≥ 5 times the mean daily precipitation)	days	

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
	max_high_prec_dur	maximum number of consecutive days with precipitation ≥ 5 times the mean daily precipitation	days	IMD
	high_prec_timing	season during which most high precipitation days (≥ 5 times the mean daily precipitation) occur	season	
	low_prec_freq	frequency of dry days (precipitation < 1 mm/day)	days/year	
	low_prec_dur	average duration of dry periods (number of consecutive days < 1 mm/day)	days	
	max_low_prec_dur	maximum number of consecutive days with precipitation < 1 mm/day	days	
	low_prec_timing	Season during which most dry days (< 1 mm/day) occur	season	
	asynchronicity	asynchronicity between the annual precipitation and PET cycles, where high values represent high relative magnitude and phase differences	-	(Feng et al., 2019)
	tmin_mean	mean daily minimum temperature	°C	IMD
	tmax_mean	mean daily maximum temperature	°C	
	pet_mean	mean daily potential evapotranspiration	mm/day	(Singer et al., 2021)
	pet_min	minimum daily potential evapotranspiration	mm/day	
	pet_max	maximum daily potential evapotranspiration	mm/day	
	pet_mean_anum	annual average total potential evapotranspiration	mm	
	pet_gleam_mean	mean daily average potential evapotranspiration	mm/day	(Miralles et al., 2011)
	aet_gleam_mean	mean daily average actual evapotranspiration	mm/day	
	evap_canopy_mean	mean daily evaporation rate from the canopy	mm/day	IMDAA
	evap_canopy_min	minimum daily evaporation rate from the canopy	mm/day	
	evap_canopy_max	maximum daily evaporation rate from the canopy	mm/day	
	evap_canopy_anum	annual average total evaporation from the canopy	mm	
	evap_surface_mean	mean daily evaporation rate from the soil surface	mm/day	
	evap_surface_min	minimum daily evaporation rate from the soil surface	mm/day	

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
	evap_surface_max	maximum daily evaporation rate from the soil surface	mm/day	
	evap_surface_anum	annual average total evaporation from the soil surface	mm	
	aridity_p_pet	aridity index (P/PET; ratio of mean annual precipitation over the mean annual potential evapotranspiration)	-	
	aridity_pet_aet	aridity index [(PET-AET)/PET; a ratio of the deficit between potential and actual evapotranspiration over potential evapotranspiration]	-	
	ai_mean	spatially averaged aridity index of the catchment	-	(Trabucco and Zomer, 2018)
	rel_hum_mean	mean daily relative humidity (2 m)	%	IMDAA
	srad_lw_mean	mean daily surface downward long-wave radiation flux	w/m^2	
	srad_sw_mean	mean daily surface downward short-wave radiation flux	w/m^2	
	wind_mean	mean daily wind speed (10 m)	m/s	
	sm_lvl1_mean	mean daily soil moisture in layer 1 (0-0.1 m below ground)	kg/m^2	
	sm_lvl2_mean	mean daily soil moisture in layer 2 (0.1-0.35 m below ground)	kg/m^2	
	sm_lvl3_mean	mean daily soil moisture in layer 3 (0.35-1 m below ground)	kg/m^2	
	sm_lvl4_mean	mean daily soil moisture in layer 4 (1-3 m below ground)	kg/m^2	
Hydrological signatures	q_mean	mean daily streamflow of the catchment	mm/day	IndiaWRIS
	runoff_ratio	runoff ratio (ratio of mean daily streamflow to the mean daily precipitation of catchment)	-	
	streamflow_elas	streamflow precipitation elasticity (i.e., the sensitivity of streamflow to changes in precipitation at the annual timescale, using the mean daily discharge as reference)	-	Eq. (7) in (Sankarasubramanian et al., 2001)
	slope_fdc	slope of the flow duration curve between the log-transformed 33rd and 66th streamflow percentiles	-	(Addor et al., 2017)
	bfi	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	-	
	q_cv	variability of daily streamflow values (coefficient of variation)	%	IndiaWRIS

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
	q_10	first decile of mean daily streamflow (the value below which 10% of the observations fall)	mm/day	IndiaWRIS
	q_25	first quartile of mean daily streamflow (the value below which 25% of the observations fall)	mm/day	
	q_50	median of mean daily streamflow (the value below which 50% of the observations fall)	mm/day	
	q_75	third quartile of mean daily streamflow (the value below which 75% of the observations fall)	mm/day	
	q_90	90th percentile of mean daily streamflow (the value below which 90% of the observations fall; High flows)	mm/day	
	q_zero	frequency of days with zero flow	days/year	
	q_low_days	mean number of consecutive days with flow less than 25th percentile mean daily flow	days	
	freq_q_low	frequency of days with low flows (flow less than 25th percentile mean daily flow)	days/year	
	q_high_days	mean number of consecutive days with a flow more than the 95th percentile mean daily flow	days	
	freq_q_high	frequency of days with high flows (flow more than 95th percentile mean daily flow)	days/year	
	annual_q	mean annual flow volume in the catchment	MCM/year	
	mean_anum_flow	mean annual flow volume in the catchment (computed for 1950 to 2020)	MCM/year	GHI (Goteti, 2023)
	cen_time	center timing, corresponds to a day of the year (doy) at which 50% of annual flow is reached	Day	
	gini_flow	uniformity of flow over the days in a year; 0 indicates equal flow throughout the year, and 1 indicates all flow occurred in a single day	-	
	annual_max_1day	mean annual 1-day maximum flow	m^3/s	IndiaWRIS
	annual_max_3day	mean annual 3-day maximum flow	m^3/s	
	annual_max_7day	mean annual 7-day maximum flow	m^3/s	

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
	annual_max_30day	mean annual 30-day maximum flow	m^3/s	IndiaWRIS
	annual_max_90day	mean annual 90-day maximum flow	m^3/s	
	annual_min_7day	mean annual 7-day minimum flow	m^3/s	
	month_1day_max	month of 1-day maximum flow for the majority of the years	calendar month	
	month_1day_min	month of 1-day minimum flow for the majority of the years	calendar month	
	doy_min_flow	the day of the year (doy) at which minimum streamflow occurred	Day	
	doy_max_flow	the day of the year (doy) at which maximum streamflow occurred	Day	
	doy_min_flow_7	the day of the year (doy) at which a minimum 7-day streamflow occurred	Day	
	doy_max_flow_7	the day of the year (doy) at which the maximum 7-day streamflow occurred	Day	
	mean_jan_flow	mean monthly flow volume of January in the catchment (computed for 1950 to 2020)	MCM/month	GHI (Goteti, 2023)
	mean_feb_flow	mean monthly flow volume of February in the catchment (computed for 1950 to 2020)	MCM/month	
	mean_mar_flow	mean monthly flow volume of March in the catchment (computed for 1950 to 2020)	MCM/month	
	mean_apr_flow	mean monthly flow volume of April in the catchment (computed for 1950 to 2020)	MCM/month	
	mean_may_flow	mean monthly flow volume of May in the catchment (computed for 1950 to 2020)	MCM/month	
	mean_jun_flow	mean monthly flow volume of June in the catchment (computed for 1950 to 2020)	MCM/month	
	mean_jul_flow	mean monthly flow volume of July in the catchment (computed for 1950 to 2020)	MCM/month	
	mean_aug_flow	mean monthly flow volume of August in the catchment (computed for 1950 to 2020)	MCM/month	
	mean_sep_flow	mean monthly flow volume of September in the catchment (computed for 1950 to 2020)	MCM/month	
	mean_oct_flow	mean monthly flow volume of October in the catchment (computed for 1950 to 2020)	MCM/month	

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
	mean_nov_flow	mean monthly flow volume of November in the catchment (computed for 1950 to 2020)	MCM/month	GHI (Goteti, 2023)
	mean_dec_flow	mean monthly flow volume of December in the catchment (computed for 1950 to 2020)	MCM/month	
	cv_jan_flow	variability of daily streamflow values in January	%	IndiaWRIS
	cv_feb_flow	variability of daily streamflow values in February	%	
	cv_mar_flow	variability of daily streamflow values in March	%	
	cv_apr_flow	variability of daily streamflow values in April	%	
	cv_may_flow	variability of daily streamflow values in May	%	
	cv_jun_flow	variability of daily streamflow values in June	%	
	cv_jul_flow	variability of daily streamflow values in July	%	
	cv_aug_flow	variability of daily streamflow values in August	%	
	cv_sep_flow	variability of daily streamflow values in September	%	
	cv_oct_flow	variability of daily streamflow values in October	%	
	cv_nov_flow	variability of daily streamflow values in November	%	
	cv_dec_flow	variability of daily streamflow values in December	%	
	mean_swmn_flow	mean flow volume of the southwest monsoon season (June, July, Aug, Sept) in the catchment (computed for 1950 to 2020)	MCM/season	GHI (Goteti, 2023)
	mean_atmn_flow	mean flow volume of autumn/retreating monsoon season (Oct, Nov) in the catchment (computed for 1950 to 2020)	MCM/season	
	mean_wint_flow	mean flow volume of the winter season (Dec, Jan, Feb) in the catchment (computed for 1950 to 2020)	MCM/season	
	mean_sumr_flow	mean flow volume of the summer season (Mar, Apr, May) in the catchment (computed for 1950 to 2020)	MCM/season	
	q_mean_swmn	mean daily streamflow of the southwest monsoon season (June, July, Aug, Sept) in a catchment	mm/day	IndiaWRIS

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
	q_5_swmn	5th percentile of daily streamflow in southwest monsoon season (June, July, Aug, Sept)	mm/day	IndiaWRIS
	q_25_swmn	first quartile of daily streamflow in southwest monsoon season (June, July, Aug, Sept)	mm/day	
	q_50_swmn	median of daily streamflow in southwest monsoon season (June, July, Aug, Sept)	mm/day	
	q_75_swmn	third quartile of daily streamflow in southwest monsoon season (June, July, Aug, Sept)	mm/day	
	q_95_swmn	95th percentile of daily streamflow in southwest monsoon season (June, July, Aug, Sept)	mm/day	
	rise_rate_mean	mean of all positive differences between consecutive daily flows	m^3/s	
	rise_rate_median	median of all positive differences between consecutive daily flows	m^3/s	
	rise_days	mean number of days in a year with positive differences between consecutive daily flows	days/year	
	fall_rate_mean	mean of all negative differences between consecutive daily flows	m^3/s	
	fall_rate_median	median of all negative differences between consecutive daily flows	m^3/s	
	fall_days	mean number of days in a year with negative differences between consecutive daily flows	days/year	
	num_hyd_alt	mean number of hydrologic reversals in a year (change from rise to fall)	-	
Land cover	water_frac	water cover fraction (2017 - 2022)	-	ESRI land cover (Karra et al., 2021)
	trees_frac	trees cover fraction (2017 - 2022)	-	
	flooded_veg_frac	flooded vegetation fraction (2017 - 2022)	-	
	crops_frac	crop cover fraction (2017 - 2022)	-	
	built_area_frac	urban cover fraction (2017 - 2022)	-	
	bare_frac	bare cover fraction (2017 - 2022)	-	
	range_frac	range cover fraction (2017 - 2022)	-	
	dom_land_cover	dominant land cover type (2017 - 2022)	-	
	dom_land_cover_frac	dominant land cover fraction (2017 - 2022)	-	
	lai_mean	catchment mean leaf area index (2001-2020)	-	MODIS MCD15A2H (Myneni et al., 2015)
	lai_min	minimum leaf area index (2001-2020)	-	
	lai_max	maximum leaf area index (2001 - 2020)	-	

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
	lai_diff	difference between maximum and minimum leaf area index (2001 - 2020)	-	MODIS MCD15A2H (Myneni et al., 2015)
Soil	soil_depth	mean soil and sedimentary-deposit thickness	m	(Pelletier et al., 2016)
	soil_conductivity_top	mean saturated hydraulic conductivity of topsoil (30 - 200 cm)	cm/day	HiHydroSoil v2.0 (Simons et al., 2020)
	soil_conductivity_sub	mean saturated hydraulic conductivity of subsoil (0 - 30 cm)	cm/day	
	soil_awc_top	mean available water content of topsoil (30 - 200 cm)	m^3/m^3	
	soil_awc_sub	mean available water content of subsoil (0 - 30 cm)	m^3/m^3	
	soil_awsc_min	minimum available water storage capacity of the soil	mm/m	FAO Soil Data (Fischer et al., 2008)
	soil_awsc_max	maximum available water storage capacity of the soil	mm/m	
	soil_awsc_major	available water storage capacity of the soil for the majority part of the catchment	mm/m	
	sand_frac_top	fraction of sand in topsoil (0 - 30 cm) for the majority of the catchment area	% wt	HWSD v2.0 (FAO and IISA, 2023)
	sand_frac_sub	fraction of sand in subsoil (30 - 100 cm) for the majority of the catchment area	% wt	
	silt_frac_top	fraction of silt in topsoil (0 - 30 cm) for the majority of the catchment area	% wt	
	silt_frac_sub	fraction of silt in subsoil (30 - 100 cm) for the majority of the catchment area	% wt	
	clay_frac_top	fraction of clay in topsoil (0 - 30 cm) for the majority of the catchment area	% wt	
	clay_frac_sub	fraction of clay in subsoil (30 - 100 cm) for the majority of the catchment area	% wt	
	gravel_frac_top	fraction of gravel in topsoil (0 - 30 cm) for the majority of the catchment area	% vol	
	gravel_frac_sub	fraction of gravel in subsoil (30 - 100 cm) for the majority of the catchment area	% vol	
	bulkdens_top_major	bulk density of topsoil (0 - 30 cm) for the majority of the catchment area	kg/dm^3	
	bulkdens_top_mean	mean reference bulk density of topsoil (0 - 30 cm)	kg/dm^3	

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
	bulkdens_sub_major	reference bulk density of subsoil (30 - 100 cm) for the majority of the catchment area	kg/dm^3	HWSD v2.0 (FAO and IISA, 2023)
	bulkdens_sub_mean	mean reference bulk density of subsoil (30 - 100 cm)	kg/dm^3	
	org_carb_top_major	organic carbon content in topsoil (0 - 30 cm) for the majority of the catchment area	% wt	
	org_carb_top_mean	mean organic carbon content in topsoil (0 - 30 cm)	% wt	
	org_carb_sub_major	organic carbon content in subsoil (30 - 100 cm) for the majority of the catchment area	% wt	
	org_carb_sub_mean	mean organic carbon content in subsoil (30 - 100 cm)	% wt	
	organic_frac_top	mean fraction of organic matter content in topsoil (30 - 200 cm)	-	HiHydroSoil v2.0 (Simons et al., 2020)
	organic_frac_sub	mean fraction of organic matter content in subsoil (0 - 30 cm)	-	
	hsg_major	hydrological soil group for the majority of the catchment area	-	
	wtd	catchment mean water table depth	m	(Fan et al., 2013)
Geological	geol_porosity	mean subsurface porosity	-	GLHYMPS (Gleeson et al., 2014)
	geol_permeability	mean subsurface permeability	m^2	
	geol_class_1st	most common geological class in a catchment	-	GLiM (Hartmann and Moosdorf, 2012)
	geol_class_1st_frac	fraction of catchment area associated with its most common geological class	-	
	geol_class_2nd	second most common geological class in catchment	-	
	geol_class_2nd_frac	fraction of catchment area associated with its second most common geological class	-	
	carb_rocks_frac	fraction of catchment area characterized as “carbonated sedimentary rocks”	-	
Anthropogenic influences	num_dams	total number of large and medium dams in catchments	-	IndiaWRIS

Table 3 – continued from the previous page

Attribute class	Attribute name	Description	Unit	Data Source / Reference
	res_store_sum	sum of total volume content of dams within the catchment	10^3 m^3	IndiaWRIS
	n_dams	total number of dams in a catchment	-	GRaND (Lehner et al., 2011)
	first_dam_year	year of construction of the first dam	-	
	latest_dam_year	year of construction of the recent dam	-	
	total_storage	total storage of the reservoirs	m^3	
	reservoir_index	ratio of total storage to multi-year annual streamflow	-	
	irrigation_frac	percentage of dams used for irrigation	-	
	hydroelec_frac	percentage of dams used for hydroelectric generation	-	
	drinking_frac	percentage of dams used for drinking	-	
	flood_frac	percentage of dams used for flood storage	-	
	overflow_frac	percentage of dams used for overflow control	-	
	navigation_frac	percentage of dams used for navigation	-	
	tailings_frac	percentage of dams used for tailings (storing by-products of mining operations)	-	
	pop_density_2000	averaged population density of the catchment in 2000	people/km^2	data.humdata.org (WorldPop and CIESIN, 2018)
	pop_density_2005	averaged population density of the catchment in 2005		
	pop_density_2010	averaged population density of the catchment in 2010		
	pop_density_2015	averaged population density of the catchment in 2015		
	pop_density_2020	averaged population density of the catchment in 2020		
	urban_frac_1985	fraction of urban land cover in a catchment in 1985	-	(Roy et al., 2015)
	urban_frac_1995	fraction of urban land cover in a catchment in 1995	-	
	urban_frac_2005	fraction of urban land cover in a catchment in 2005	-	
	crops_frac_1985	fraction of cropland land cover in a catchment in 1985	-	
	crops_frac_1995	fraction of cropland land cover in a catchment in 1995	-	
	crops_frac_2005	fraction of cropland land cover in a catchment in 2005	-	

5. Filter catchments based on available streamflow observations

The Python script file "filter_catchment.py" allows users to quickly filter catchments based on the percentage duration of available observed streamflow data from 1980 to 2020. We have filtered 228 catchments with more than 30% streamflow data availability over this period, provided in the file "*CAMELS_IND_Catchments_Streamflow_Sufficient-.zip*". However, users can filter specific catchment data according to their requirements using the provided script.

6. Change logs and disclaimer

The file "*change_logs_and_disclaimer.txt*" contains the CAMELS-IND dataset's change logs and a disclaimer regarding data usage.

Acknowledgments

The authors gratefully acknowledge the Central Water Commission (CWC), the National Water Informatics Centre (NWIC), and the Ministry of Jal Shakti (MoJS) for providing the streamflow dataset through the online portal, India – Water Resources Information System (India-WRIS; https://indiawris.gov.in/wris/#/). The authors also extend their gratitude to the India Meteorological Department (IMD), Ministry of Earth Sciences, Government of India, for providing the gridded rainfall and temperature datasets through their respective websites. Additionally, the authors gratefully acknowledge the National Centre for Medium Range Weather Forecasting (NCMRWF), Ministry of Earth Sciences, Government of India, for the Indian Monsoon Data Assimilation and Analysis (IMDAA) reanalysis. The IMDAA reanalysis was produced under the collaboration between UK Met Office, NCMRWF, and IMD, with financial support from the Ministry of Earth Sciences under the National Monsoon Mission programme. The authors utilized numerous publicly available datasets for compiling catchment attributes and meteorological forcing time series, duly acknowledging and citing them where applicable. The authors extend their gratitude to all the researchers and contributing authors of these open-source datasets.

References

Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: 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, 2017.

Fan, Y., Li, H., and Miguez-Macho, G.: Global Patterns of Groundwater Table Depth, Science (80-.)., 339, 940–943, https://doi.org/10.1126/science.1229881, 2013.

FAO and IISA: Harmonized World Soil Database version 2.0, FAO; International Institute for Applied Systems Analysis (IIASA);, https://doi.org/10.4060/cc3823en, 2023.

Feng, X., Thompson, S. E., Woods, R., and Porporato, A.: Quantifying Asynchronicity of Precipitation and Potential Evapotranspiration in Mediterranean Climates, Geophys. Res. Lett., 46, 14692–14701, https://doi.org/10.1029/2019GL085653, 2019.

Fischer, G., Nachtergaele, F., Prieler, S., van Velthuizen, H. T., Verelst, L., and Wiberg, D.: Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008), IIASA, Laxenburg, Austria and FAO, Rome, Italy, 2008.

Gleeson, T., Moosdorf, N., Hartmann, J., and van Beek, L. P. H.: A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity, Geophys. Res. Lett., 41, 3891–3898, https://doi.org/10.1002/2014GL059856, 2014.

Goteti, G.: Geospatial dataset for hydrologic analyses in India (GHI): a quality-controlled dataset on river gauges, catchment boundaries and hydrometeorological time series, Earth Syst. Sci. Data, 15, 4389–4415, https://doi.org/10.5194/essd-15-4389-2023, 2023.

Hartmann, J. and Moosdorf, N.: The new global lithological map database GLiM: A representation of rock properties at the Earth surface, Geochemistry, Geophys. Geosystems, 13, 1–26, https://doi.org/10.1029/2012GC004370, 2012.

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., and Brumby, S. P.: Global land use / land cover with Sentinel 2 and deep learning, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 4704–4707, https://doi.org/10.1109/IGARSS47720.2021.9553499, 2021.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., and Wisser, D.: High-resolution mapping of the world's reservoirs and dams for

sustainable river-flow management, Front. Ecol. Environ., 9, 494–502, https://doi.org/10.1890/100125, 2011.

Mangukiya, N. K., Sharma, A., and Shen, C.: How to enhance hydrological predictions in hydrologically distinct watersheds of the Indian subcontinent?, Hydrol. Process., 37, e14936, https://doi.org/10.1002/hyp.14936, 2023.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci., 15, 453–469, https://doi.org/10.5194/hess-15-453-2011, 2011.

Myneni, R.; Knyazikhin, Y.; Park, T.: MCD15A2H MODIS/Terra+Aqua Leaf Area Index/FPAR 8-day L4 Global 500m SIN Grid V006 [Data set], NASA EOSDIS L. Process. DAAC, https://doi.org/10.5067/MODIS/MCD15A2H.006, 2015.

Pai, D. ., Rajeevan, M., Sreejith, O. ., Mukhopadhyay, B., and Satbha, N. .: Development of a new high spatial resolution (0.25° × 0.25°) long period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region, MAUSAM, 65, 1–18, https://doi.org/10.54302/mausam.v65i1.851, 2014.

Pelletier, J. D., Broxton, P. D., Hazenberg, P., Zeng, X., Troch, P. A., Niu, G., Williams, Z., Brunke, M. A., and Gochis, D.: A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling, J. Adv. Model. Earth Syst., 8, 41–65, https://doi.org/10.1002/2015MS000526, 2016.

Rani, S. I., T, A., George, J. P., Rajagopal, E. N., Renshaw, R., Maycock, A., Barker, D. M., and Rajeevan, M.: IMDAA: High Resolution Satellite-era Reanalysis for the Indian Monsoon Region, J. Clim., 34, 5109–5133, https://doi.org/10.1175/JCLI-D-20-0412.1, 2021.

Roy, P., Roy, A., Joshi, P., Kale, M., Srivastava, V., Srivastava, S., Dwevidi, R., Joshi, C., Behera, M., Meiyappan, P., Sharma, Y., Jain, A., Singh, J., Palchowdhuri, Y., Ramachandran, R., Pinjarla, B., Chakravarthi, V., Babu, N., Gowsalya, M., Thiruvengadam, P., Kotteeswaran, M., Priya, V., Yelishetty, K., Maithani, S., Talukdar, G., Mondal, I., Rajan, K., Narendra, P., Biswal, S., Chakraborty, A., Padalia, H., Chavan, M., Pardeshi, S., Chaudhari, S., Anand, A., Vyas, A., Reddy, M., Ramalingam, M., Manonmani, R., Behera, P., Das, P., Tripathi, P., Matin, S., Khan, M., Tripathi, O., Deka, J., Kumar, P., and Kushwaha, D.: Development of Decadal (1985–1995–2005)

Land Use and Land Cover Database for India, Remote Sens., 7, 2401–2430, https://doi.org/10.3390/rs70302401, 2015.

Sankarasubramanian, A., Vogel, R. M., and Limbrunner, J. F.: Climate elasticity of streamflow in the United States, Water Resour. Res., 37, 1771–1781, https://doi.org/10.1029/2000WR900330, 2001.

Simons, G. ., Koster, R. ., and Droogers, P.: HiHydroSoil v2.0 - A high resolution soil map of global hydraulic properties, Wageningen, The Netherlands, 2020.

Singer, M. B., Asfaw, D. T., Rosolem, R., Cuthbert, M. O., Miralles, D. G., MacLeod, D., Quichimbo, E. A., and Michaelides, K.: Hourly potential evapotranspiration at 0.1° resolution for the global land surface from 1981-present, Sci. Data, 8, 224, https://doi.org/10.1038/s41597-021-01003-9, 2021.

Srivastava, A. K., Rajeevan, M., and Kshirsagar, S. R.: Development of a high resolution daily gridded temperature data set (1969-2005) for the Indian region, Atmos. Sci. Lett., 10, 249–254, https://doi.org/10.1002/asl.232, 2009.

Trabucco, A. and Zomer, R. J.: Global Aridity Index and Potential Evapo-Transpiration (ET0) Climate Database v2, CGIAR Consortium for Spatial Information (CGIAR-CSI), 2018.

WorldPop and CIESIN: Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076), WorldPop (www.worldpop.org - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Sc, https://doi.org/10.5258/SOTON/WP00674, 2018.