Detailed description on data sources and processing will be made available as a data publication.

CAMELS-FI contains regulated and unregulated catchments. Look at CAMELS_FI_humaninfluence_attributes.csv for regulation_level to see how strongly the catchment is regulated. 1 = Unregulated, 2 = Minor regulation, 3 = Major regulation

Code used for creating the dataset is accessible at https://github.com/iiroseppa/CAMELS-FI. Please note that the work included some manual work, so the code alone isn't enough for full reproducibility.

We follow CAMELS-CH and CAMELS-GB for selecting most of the attributes and their names, and CAMELS-SE for soil attributes and lack of hydrolithological data (Coxon et al., 2020b; Höge et al., 2023; Teutschbein, 2024). The file names and directory structure of the data follow convention of CAMELS-GB in order to facilitate easier integration of CAMELS-FI to existing software that interfaces with CAMELS data.

1 Table of abbreviations

Full name	Abbreviation
Catchment Attributes and MEteorology for Large-sample Studies	CAMELS
Finnish Environment Institute	SYKE
Application Programming Interface	API
Centre for Economic Development, Transport and the Environment	ELY
Geological survey of Finland	GTK
National Land Survey of Finland	NLS
Corine Land Cover	CLC

2 Directory structure and file contents

You can click a linked name to go to the section.

CAMELS-FI

- **data**
  - CAMELS_FI_catchment_boundaries.zip
  - CAMELS_FI_catchment_boundaries.gpkg
  - CAMELS_FI_climatic_attributes.csv
  - CAMELS_FI_humaninfluence_attributes.csv
  - CAMELS_FI_hydrologic_attributes.csv
  - CAMELS_FI_landcover_attributes.csv
  - CAMELS_FI_meta_attributes.csv
  - CAMELS_FI_soil_attributes.csv
  - CAMELS_FI_topographic_attributes.csv
  - **annual-timeseries**
    - CAMELS_FI_landcover_timeseries_1000_2000-2018.csv
    - CAMELS_FI_landcover_timeseries_1003_2000-2018.csv
    - ...
  - **artificial-cross-catchment-bifurcations**
    - artificial_bifurcations.gpkg
    - cross_catchment_flow_19610101-20231231.csv
  - **timeseries**
    - CAMELS_FI_hydromet_timeseries_1000_19610101-20231231.csv
    - CAMELS_FI_hydromet_timeseries_1003_19610101-20231231.csv
    - ...
  - **timeseries-by-attribute**
    - discharge_spec.csv
    - discharge_vol.csv
    - humidity_rel.csv
    - pe_era5_land.csv
    - pet.csv
    - pet_fmi.csv
    - precipitation.csv
    - radiation_global.csv
    - snow_depth.csv
    - snow_evaporation.csv
    - swe_cci3-1.csv
    - swe.csv
    - temperature_gmin.csv
    - temperature_max.csv
    - temperature_mean.csv
    - temperature_min.csv
- **supporting-documents**
  - support_document.pdf

All csv files have comma as a value separator, point as a decimal separator and text is unquoted. Missing values are empty.

2.1 data

Contains three timeseries directories, one directory containing data on some artificial bifurcations, static attributes as seven csv-files and catchment boundaries as well as flow gauge locations in a geopackage, as well as catchment boundaries as a zipped shp.

2.1.1 CAMELS_FI_catchments.gpkg

A geopackage containing two layers: gauges and catchments, and also default styling (works with QGIS and helps visualize nested nature of the catchment). Gauges show the location of the gauge, snapped to the closest suitable location in hydrologically preprocessed DEM. This location was used as pour point when creating the catchments. Due to this, gauge locations are guaranteed to be at the edge of the corresponding catchment, but still inside. Both contain gauge id, and catchments also contain area as an attribute. Other attributes are not provided here to allow users to join only the attributes they need, which avoids visual clutter.

2.1.2 CAMELS_FI_catchment_boundaries.zip

Zipped shapefile containing the catchment geometries as shapefile. Included for compatibility reasons, contains the same information as catchments layer of CAMELS_FI_catchments.gpkg.

2.1.3 Static attributes

Six csv files with naming convention CAMELS_FI_{type}_attributes.csv. The first column contains catchment identifier that is present on all data files and allows combining different data types together. There are a total of 320 catchments. Please note that some of the gauges have been created by combining values of two or more gauges (quite often a bifurcation where the branches have separate gauges). These “virtual gauges” can be recognized by their id, which is in format xxxx-xxxx instead of typical xxx or xxxx.

Metadata attributes contain general information about the gauge or catchment.

Table 1: contents of CAMELS_FI_metadata_attributes.csv

Attribute name	Description	Unit	Data source
gauge_id	catchment identifier (corresponds to SYKE’s discharge station id)	-	SYKE’s hydrology API
gauge_name	gauging station name (river or lake name, often followed by more precise location)	-	
owner_id	Identifier of the maintainer of the gauge (SYKE, one of the ELY-centres or other)	-	
owner_name	Name of the maintainer of the gauge	-	
gauge_lat	gauge longitude (EPSG:4326)	◦	
gauge_lon	gauge station latitude (EPSG:4326)	◦	
gauge_easting	ETRS-TM35FIN coordinates (EPSG:3067), easting	m	
gauge_northing	ETRS-TM35FIN coordinates (EPSG:3067), northing	m	

basin_id	id of the drainage basin assigned by SYKE	-	
basin_name	name of the drainage basin	-	
area	area of the watershed	km^2	Derived by authors from watersheds and gauge locations. Water regions based on Korhonen and Kuusisto, (2010).
nestedness	how many catchments contain this catchment as subcatchment	-	
water_region_code	water region code	-	
water_region_name	water region name	-	
cross_border_perc	percentage outside of Finland.	%	
reference_gauge	is the gauge among SYKE's reference hydrometric network. Close to natural and high quality observations.	yes/no	Turner et al., 2025
ice_correction	have there been manual corrections to the flow values due to ice dams	yes/no	Internal document from SYKE

Hydrologic attributes contain different hydrologic indices calculated for each catchment from 1990-09-01 to 2020-08-31. This corresponds to water years of climatic reference period 1991-2020. If there were less than 1826 days of observations during that period, all available observations were used.

Table 2: contents of CAMELS_FI_hydrologic_attributes.csv

Attribute name	Description	Unit	Data source
gauge_id	catchment identifier (corresponds to SYKE's discharge station id)	-	SYKE's hydrology API
timeseries_number_of_years	total duration of timeseries in years, without gaps	y	SYKE's hydrology API
sign_start_date	Start date for signature evaluation	date	
sign_end_date	End date for signature evaluation	date	
sign_number_of_years	Number of years for signature evaluation, without gaps	y	
sign_number_of_obs	Total number of observations	-	
q_mean	Mean daily specific discharge	$mm\ d^{-1}$	
runoff_ratio	Runoff ratio (ratio of mean daily discharge to mean daily precipitation)	-	
stream_elas	Stream flow precipitation elasticity (sensitivity of stream flow to changes during a year)	-	
slope_fdc	Slope of the flow duration curve (between the log-transformed 33rd and 66th stream flow percentiles) (Yadav et al., 2007). Value is NaN if over a third of the observations are zero.	-	
baseflow_index_landson	base flow index (ratio of mean daily base flow to daily discharge, using the Ladson et al. (2013) digital filter	-	

baseflow_index_lfstat	base flow index (ratio of mean daily base flow to daily discharge, using the lfstat implementation of Tallaksen and Van Lanen (2004)	-
hfd_mean	Mean half-flow date (number of days since 1 September at which the cumulative discharge reaches half of the annual discahrge)	d
Q5	5 % flow quantile (low flow)	mm d^{-1}
Q95	95 % flow quantile (high flow)	mm d^{-1}
high_q_freq	frequency of high-flow days (> 9 times the median daily flow)	d y^{-1}
high_q_dur	average duration of low-flow events (number of consecutive days > 9 times the median daily flow)	d
low_q_freq	frequency of low-flow days (< 0.2 times the mean daily flow)	d y^{-1}
low_q_dur	average duration of low-flow events (number of consecutive days < 0.2 times the mean daily flow)	d
zero_q_freq	fraction of days with Q = 0	-

Climatic attributes contain climatic indices calculated for each catchment from 1990-09-01 to 2020-08-31. This corresponds to water years of climatic reference period 1991-2020.

Table 3: contents of CAMELS_FI_climatic_attributes.csv

Attribute name	Description	Unit	Data source
gauge_id	catchment identifier (corresponds to SYKE's discharge station id)	-	SYKE's hydrology API
p_mean	Long-term mean daily precipitation	mm d^{-1}	FMI gridded climatology
temperature_mean	Mean annual temperature	℃	
pet_mean	long-term mean daily PET	mm d^{-1}	FMI gridded climatology, ERA5 & ERA5-Land
aridity	aridity, calculated as the ratio of mean daily potential evapotranspiration to mean daily precipitation	-	
p_seasonality	seasonality and timing of precipitation (estimated using sine curves to represent the annual temperature and precipitation cycles, positive (negative) values indicate that precipitation peaks in summer (winter), and values close to zero indicate uniform precipitation throughout the year). See Eq. (14) in Woods (2009)	-	FMI gridded climatology
frac_snow	fraction of precipitation falling as snow (T< 0 ℃)	-	
high_prec_freq	frequency of high-precipitation days (≥ 5 times mean daily precipitation)	d y^{-1}	
high_prec_dur	average duration of high-precipitation events (number of consecutive days ≥ 5 times mean daily precipitation)	d	
high_prec_timing	Season during which most high precipitation days occur,If two seasons register the same number of events, a value of NaN is given	season	
low_prec_freq	frequency of dry days (< 1 mm d−1)	d y^{-1}	

low_prec_dur	Average duration of dry periods (number of consecutive days < 1 mm d−1 mean daily precipitation)	d	
low_prec_timing	season during which most dry days (< 1 mm d−1) occur. If two seasons register the same number of events, a value of NaN is given	season	

Topographic attributes give an overview on the topography of the catchment.

Table 4: contents of CAMELS_FI_topographic_attributes.csv

Attribute name	Description	Unit	Data source
gauge_id	catchment identifier (corresponds to SYKE's discharge station id)	-	SYKE's hydrology API
slope	catchment mean slope	%	10 m x 10 m DEM, NLS
elev_gauge	gauge elevation. Not always equal to minimum due to mines.	m a.s.l.	
elev_min	minimum elevation	m a.s.l.	
elev_10	10th elevation percentile	m a.s.l.	
elev_50	median elevation	m a.s.l.	
elev_90	90th elevation percentile	m a.s.l.	
elev_max	maximum elevation	m a.s.l.	
elev_range	gauge elevation substracted from maximum	m	

Soil attributes give areal percentages of the dominant soil class, following the classification by GTK. Please note that that this differs from most previous CAMELS, but similar data to previous CAMELS is not easily available for Finland in high enough resolution. Soil depth is also included.

Table 5: Contents of CAMELS_FI_soil_attributes.csv

Attribute name	Description	Unit	Data source
gauge_id	catchment identifier (corresponds to SYKE's discharge station id)	-	SYKE's hydrology API
bedrock_perc	Percentage of rocky outcrops of total area	%	Superficial deposits of Finland 1:200 000 (GTK)
coarse_perc	Percentage of coarse grained (glaciofluvial) deposits of total area	%	
silt_perc	Percentage of silt dominated soils of total area	%	
till_perc	Percentage of till dominated soil of total area	%	
clay_perc	Percentage of clay dominated soil of total area	%	
peat_perc	Percentage of peat dominated soil of total area	%	
soil_depth	Mean soil depth to bedrock	m	Superficial deposits thickness 1:1000 000 (GTK)

Table 6: GTK's soil classes that were combined for CAMELS-FI

	bedrock	coarse	silt	till	clay	peat

Original class	bedrock outcrop	coarse grained	fine grained	till	clay	peat < 0.3 m
	bedrock with less than 1 m of till on top				gyttja	peat > 0.3 m, < 0.6 m
	fragmented rock				gyttja and fine particles	peat > 0.6 m

Landcover attributes are available for four different years based directly on data, present in CAMELS_FI_landcover_attributes.csv. Interpolated values between different years are also available at **annual-timeseries** directory.

Table 7: Contents of CAMELS_FI_landcover_attributes.csv

Attribute name	Description	Unit	Data source
gauge_id	catchment identifier (corresponds to SYKE's discharge station id)	-	SYKE's hydrology API
crop_perc_2000	Percentage of agricultural land of total area, 2000	%	CLC Finland 2000
grass_perc_2000	Percentage of grassland of total area, 2000	%	
shrub_perc_2000	Percentage of shrubland of total area, 2000	%	
dwood_perc_2000	Percentage of deciduous forests of total area, 2000	%	
ewood_perc_2000	Percentage of evergreen forests of total area, 2000	%	
urban_perc_2000	Percentage of impermeable or poorly permeable human made surfaces and built areas of total area, 2000	%	
inwater_perc_2000	Percentage of water areas of total area, 2000	%	
bares_perc_2000	Percentage of bare land of total area, 2000	%	
wetland_perc_2000	Percentage of wetlands of total area, 2000	%	
crop_perc_2006	Percentage of agricultural land of total area, 2006	%	CLC Finland 2006
grass_perc_2006	Percentage of grassland of total area, 2006	%	
shrub_perc_2006	Percentage of shrubland of total area, 2006	%	
dwood_perc_2006	Percentage of deciduous forests of total area, 2006	%	
ewood_perc_2006	Percentage of evergreen forests of total area, 2006	%	
urban_perc_2006	Percentage of impermeable or poorly permeable human made surfaces and built areas of total area, 2006	%	
inwater_perc_2006	Percentage of water areas of total area, 2006	%	
bares_perc_2006	Percentage of bare land of total area, 2006	%	
wetland_perc_2006	Percentage of wetlands of total area, 2006	%	
crop_perc_2012	Percentage of agricultural land of total area, 2012	%	CLC Finland 2012
grass_perc_2012	Percentage of grassland of total area, 2012	%	
shrub_perc_2012	Percentage of shrubland of total area, 2012	%	
dwood_perc_2012	Percentage of deciduous forests of total area, 2012	%	
ewood_perc_2012	Percentage of evergreen forests of total area, 2012	%	
urban_perc_2012	Percentage of impermeable or poorly permeable human made surfaces and built areas of total area, 2012	%	
inwater_perc_2012	Percentage of water areas of total area, 2012	%	
bares_perc_2012	Percentage of bare land of total area, 2012	%	

wetland_perc_2012	Percentage of wetlands of total area, 2012	%	
crop_perc_2018	Percentage of agricultural land of total area, 2018	%	CLC Finland 2018
grass_perc_2018	Percentage of grassland of total area, 2018	%	
shrub_perc_2018	Percentage of shrubland of total area, 2018	%	
dwood_perc_2018	Percentage of deciduous forests of total area, 2018	%	
ewood_perc_2018	Percentage of evergreen forests of total area, 2018	%	
urban_perc_2018	Percentage of impermeable or poorly permeable human made surfaces and built areas of total area, 2018	%	
inwater_perc_2018	Percentage of water areas of total area, 2018	%	
bares_perc_2018	Percentage of bare land of total area, 2018	%	
wetland_perc_2018	Percentage of wetlands of total area, 2018	%	

Table 8: CLC lvl 3 classes used for each land cover class in CAMELS-FI. Please note that some of the classes are not present in all editions.

	crops	grass	shrub	dwood	ewood	urban	inwater	bares	wetland
CLC	211	141	324	311	312	111	511	131	411
codes	222	142		311	312	112	512	132	412
	243	231			312	121		331	
		244			313	123		332	
		321			313	124			
		322			313	133			
						142			

Human influence attributes describe the degree of human influence on the catchment. We define reservoir as a lake that has known volume range.

Attribute name	Description	Unit	Data source
gauge_id	catchment identifier (corresponds to SYKE's discharge station id)	-	SYKE's hydrology API
num_inhabitants	population count		Statistics Finland, 2024
dens_inhabitants	population density	km^{-2}	
num_dam	dam count, includes active and passive dams	-	NLS's Topographic database 2024
num_reservoir	reservoir count	-	SYKE's lake API
reservoir_cap	reservoir capacity	1000 m^3	
num_regulation_other	Count of other water regulation permits than reservoirs	-	
regulation_level	how strongly the catchment is regulated. 1 = Unregulated, 2 = Minor regulation, 3 = Major regulation	-	

2.1.4 timeseries

Contains 320 files, one for each catchment, in format

CAMELS_FI_hydromet_timeseries_{gauge_id}_19610101-20231231.csv. Each file contains daily catchment averaged meteorology from first of January 1961 to last of December 2023 and all

available daily discharge data from that period for one catchment. Minimum temperature near ground has some missing periods after 15th October 2021, due to data missing in the original dataset.

Table 9: Summary of timeseries variables in CAMELS-FI

Type	Attribute	Description	Unit	Data source
Hydrological time series	discharge_vol	catchment discharge, calculated from water level and channel geometry at gauge	$m^3 s^{-1}$	SYKE's hydrology API
	discharge_spec	catchment-spesific discharge converted to millimeters using catchment area	mm d^{-1}	
Meteorological time series	precipitation	catchment daily averaged precipitation	mm d^{-1}	FMI gridded climatology
	pet	If there is snow depth is larger than zero: ERA5 snow evaporation. Otherwise months 4 to 9: FMI; months 10 to 3: potential evaporation from ERA5-Land	mm d^{-1}	FMI gridded climatology, ERA5, ERA5-Land
	pe_era5_land	catchment daily averaged potential evaporation, calculated as pan evaporation	mm d^{-1}	ERA5-Land
	snow_evaporation	catchment daily averaged evaporation from snow	mm d^{-1}	ERA5
	pet_fmi	catchment daily averaged potential evapotranspiration, calculated for well watered 12 cm grass.	mm d^{-1}	FMI gridded climatology
	temperature_gmin	catchment daily averaged minimum temperature near ground	℃	
	temperature_min	catchment daily averaged minimum temperature at 2m	℃	
	temperature_mean	catchment daily averaged air temperature at 2m	℃	
	temperature_max	catchment daily averaged maximum air temperature at 2m	℃	
	radiation_global	catchment daily averaged global radiation sum	kJ m^{-2}	
	humidity_rel	catchment daily averaged relative humidity	%	
	snow_depth	catchment daily averaged snow depth at 6:00 UTC	cm	
	swe	catchment daily averaged snow water equivalent, ERA5-Land	mm	ERA5-Land
	swe_cci3-1	catchment daily averaged snow water equivalent, ESA Snow Climate Change Initiative, version 3.1	mm	Luojus et al., 2024

2.1.5 timeseries-by-attribute

Contains the same information as timeseries folder, just organized by variable for convenience, so that one file contains timeseries data of one variable for all catchments. Files are named with convention {attribute_name}.csv. gauge_id is included to identify catchments.

2.1.6 annual-timeseries

Contains 320 landcover timeseries files, one for all catchments, in format CAMELS_FI_landcover_timeseries_{gauge_id}_2000-2018.csv. Land cover values for 2000, 2006, 2012 and 2018 equal those present in CAMELS_FI_landcover_attributes.csv, other values are linear interpolations between them. One row per year, one column per variable.

2.1.7 artificial-cross-catchment-bifurcations

Contains supporting information about locations where water is being pumped from one watershed to another. For these gauges, only timeseries data (in cross_catchment_discharge_19610101-20231231.csv) is discharge ($m^3\ s^{-1}$) and only static attributes (in cross_catchment_discharge_19610101-20231231.csv) provided are limited gauge metadata and mean discharge ($m^3\ s^{-1}$, different from proper catchments which are in mm d^{-1}). No watershed is provided. The gauges are: Päijänne Water Tunnel (1092 & 1093); Paimio river to Aura river (1115); Lokka reservoir to Luiro river (1352) and Saimaa channel to Rakkola river (3683). Overall, the effects for both source and recipient catchments are usually relatively minor, altough the effect might be noticable during low flow. These artificial bifurcations are fully separated from the rest of the data, and are not counted as catchments in CAMELS-FI.

2.2 supporting-documents

Contains this file.

3 References

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., and Woods, R.: CAMELS-GB: hydrometeorological time series and landscape attributes for 671 catchments in Great Britain, Earth System Science Data, 12, 2459–2483, https://doi.org/10.5194/essd-12-2459-2020.

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., and Fenicia, F.: CAMELS-CH: hydro-meteorological time series and landscape attributes for 331 catchments in hydrologic Switzerland, Earth System Science Data, 15, 5755–5784, https://doi.org/10.5194/essd-15-5755-2023, 2023.

Korhonen, J. and Kuusisto, E.: Long-term changes in the discharge regime in Finland, Hydrology Research, 41, 253–268, https://doi.org/10.2166/nh.2010.112, 2010.

Luojus, K., Venäläinen, P., Moisander, M., Pulliainen, J., Takala, M., Lemmetyinen, J., Derksen, C., Mortimer, C., Mudryk, L., Scwaizer, G., and Nagler, T.: ESA Snow Climate Change Initiative (Snow_cci): Snow Water Equivalent (SWE) level 3C daily global climate research data package (CRDP) (1979 - 2022), version 3.1., https://dx.doi.org/10.5285/9d9bfc488ec54b1297eca2c9662f9c81, 2024.

Tallaksen, L. M. and Van Lanen, H. A. J.: Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater, Developments in Water Science, 48, 2004.

Teutschbein, C.: CAMELS-SE: Long-term hydroclimatic observations (1961–2020) across 50 catchments in Sweden as a resource for modelling, education, and collaboration, Geoscience Data Journal, 11, 655–668, https://doi.org/10.1002/gdj3.239, 2024.

Turner, S., Hannaford, J., Barker, L. J., Suman, G., Killeen, A., Armitage, R., Chan, W., Davies, H., Griffin, A., Kumar, A., Dixon, H., Albuquerque, M. T. D., Almeida Ribeiro, N., Alvarez-Garreton, C., Amoussou, E., Arheimer, B., Asano, Y., Berezowski, T., Bodian, A., Boutaghane, H., Capell, R., Dakhaoui, H., Daňhelka, J., Do, H. X., Ekkawatpanit, C., El Khalki, E. M., Fleig, A. K., Fonseca, R., Giraldo-Osorio, J. D., Goula, A. B. T., Hanel, M., Horton, S., Kan, C., Kingston, D. G., Laaha, G., Laugesen, R., Lopes, W., Mager, S., Rachdane, M., Markonis, Y., Medeiro, L., Midgley, G., Murphy, C., O'Connor, P., Pedersen, A. I., Pham, H. T., Piniewski, M., Renard, B., Saidi, M. E., Schmocker-Fackel, P., Stahl, K., Thyer, M., Toucher, M., Tramblay, Y., Uusikivi, J., Venegas-Cordero, N., Visessri, S., Watson, A., Westra, S., and Whitfield, P. H.: ROBIN: Reference observatory of basins for international hydrological climate change detection, Sci Data, 12, 654, https://doi.org/10.1038/s41597-025-04907-y, 2025.

Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, Advances in Water Resources, 30, 1756–1774, https://doi.org/10.1016/j.advwatres.2007.01.005, 2007.