https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Open Access Earth System Science Data Discussions

LamaH-Ice: LArge-SaMple Data for Hydrology and Environmental Sciences for Iceland

Hordur B. Helgason[1,2], Bart Nijssen[1]

[1]Department of Civil and Environmental Engineering, University of Washington, Seattle, 98195, USA
[2]Hydropower Division, Landsvirkjun, Reykjavík, 105 Iceland

Correspondence to: Hordur B. Helgason (helgason@uw.edu)

Abstract. Access to mountainous regions for monitoring streamflow, snow and glaciers is often difficult, and many rivers are thus not gauged and hydrological measurements are limited. Consequently, cold-region watersheds, particularly heavily glacierized ones, are poorly represented in large-sample hydrology (LSH) datasets. We present a new LSH dataset for Iceland, termed LamaH-Ice (LArge-SaMple DAta for Hydrology and Environmental Sciences for Iceland). Glaciers and ice caps cover about 10% of Iceland and while streamflow has been measured for several decades, these measurements have not been published in a consistent manner ~~before.~~ The dataset provides daily and hourly hydrometeorological timeseries and catchment characteristics for 107 river basins in Iceland, covering an area of almost 46,000 km^2 (45% of Iceland's area), with catchment sizes ranging from 4 to ~~about~~ 7,500 km^2. LamaH-Ice conforms to the structure of existing LSH datasets and includes most variables offered in these datasets, as well as additional information relevant to cold-region hydrology, e.g., timeseries of snow cover, glacier mass balance and albedo. LamaH-Ice also includes dynamic catchment characteristics to account for changes in land cover, vegetation, and glacier extent. A large majority of the watersheds in LamaH-Ice are not subject to human activities, such as diversions and flow regulations. Streamflow measurements under natural flow conditions are highly valuable to hydrologists seeking to model and comprehend the natural hydrological cycle or estimate climate change trends. The LamaH-Ice dataset (Helgason and Nijssen, 2023) is intended for the research community to improve the understanding of hydrology in cold-region environments.

1 Introduction

In cold regions, glaciers and seasonal snow act as freshwater reservoirs and provide critical water resources for downstream communities. These water resources sustain irrigation, hydropower production, water-supply and the health of ecosystems (National Research Council, 2012; Barnett et al., 2005; Mark and Seltzer, 2003). Monitoring hydrological processes in snow and glacier melt-dominated catchments is important for water resource management and to understand the impacts of climate change in cold regions, however access to mountainous regions for monitoring snow and glaciers is often difficult (Brown et al., 2014; Rohrer et al., 2013). Many rivers in these regions are thus not gauged and information about streamflow, as well as other hydrological processes, is limited. For this reason, there is currently a lack of availability of large datasets including hydrological measurements from snow and glacier melt dominated catchments.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

This paper presents a Large-sample Hydrology (LSH) dataset for Iceland, named LamaH-Ice. LamaH-Ice provides easy access to hydrometeorological timeseries, including multi-year series of observed streamflow, and catchment characteristics for 107 Icelandic catchments to the scientific community for large-sample studies. The structure of the dataset is based on existing LSH datasets. It is designed to be compatible with the Large-Sample Data for Hydrology and Environmental Sciences for Central Europe (LamaH-CE; Klingler et al., 2021) and the well-known Catchment Attributes for Large-Sample Studies (CAMELS; Addor et al., 2017; Newman et al., 2015). LamaH-Ice offers most hydrometeorological variables and catchment characteristics that are included in these datasets, as well as additional information relevant to cold-region hydrology. In particular, LamaH-Ice includes glacier characteristics as well as dynamic catchment characteristics to facilitate the study of the effects of climate change and changes in land cover on hydrology. Streamflow measurements that are uninterrupted by human activities are of great importance in the field of hydrology, e.g., in climate change studies and in evaluations of hydrological model simulations (Stahl et al., 2010; Hannah et al., 2011). A large majority of the streamflow gauges included in LamaH-Ice (79 out of 107) exhibit natural flow conditions.

The hydrological cycle is a complicated system that consists of many different processes, which are characterized by high spatio-temporal variability. To understand these processes, hydrologists rely on data. Hydrologists have long looked for ways to regionalize their understanding, i.e., to transfer knowledge about runoff response between two separate basins based on hydrological similarity (Merz et al., 2006). As this field of research has progressed, LSH has become an established branch of hydrology, developing rapidly in recent years. LSH studies, in contrast to intensive place- or regional based investigations, use data from many catchments simultaneously for hydrological investigations (Gupta et al., 2014). These studies use data from tens to thousands of catchments to learn from similarities and differences between different regions, enabling the transfer of knowledge between regions and the formulation of reliable conclusions regarding hydrological processes (Addor et al., 2020). The recent, rapid development of LSH has been strongly driven by the application of machine learning methods in hydrology. In recent years it has been shown that by leveraging LSH datasets, hydrologists can build accurate data-driven rainfall-runoff models that are generalizable and able to predict streamflow in ungauged basins (Kratzert et al., 2019; Gauch et al., 2021), which has been a longstanding problem in hydrology (Hrachowitz et al., 2013).

Several LSH datasets have now been assembled and published. The CAMELS dataset (Addor et al., 2017; Newman et al., 2015) provides catchment attributes, daily meteorological forcing data, and streamflow measurements for 671 river basins within the contiguous United States. New datasets, consistent with CAMELS, have been produced for other areas of the world, e.g., for Great-Britain (CAMELS-GB; Coxon et al., 2020), Chile (CAMELS-CL; Alvarez-Garreton et al., 2018), Australia (CAMELS-AUS; Fowler et al., 2021) and Brazil (CAMELS-BR; Chagas et al., 2020). The meteorological forcings and catchment attributes in all CAMELS datasets are aggregated over the full upstream area of streamflow gauges, thus offering no spatial variability and lacking information about upstream river networks. LamaH-CE (Klingler et al., 2021) is constructed in a similar manner as the CAMELS dataset, with the main difference that it includes a greater degree of spatially distributed forcings by encompassing intermediate catchments. Additionally, LamaH-CE enhances its temporal resolution by offering

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

hourly hydrometeorological timeseries. This is important for process understanding and modeling, especially in snow and glacier melt dominated catchments where substantial sub-daily variability exists in hydrological processes.

The consistent structure of the CAMELS and LamaH-CE datasets allows these datasets to be combined to facilitate global LSH studies. The AI4Water Python Package (Abbas et al., 2022) enables users to combine the CAMELS and LamaH datasets for modeling purposes. Kratzert et al. (2023) combined and standardized several existing LSH datasets into Caravan, a unified global LSH dataset. They also provided cloud-based tools to enable users to add additional catchments, which has led to further extensions of the dataset. The first two community additions to Caravan were from Denmark (Koch, 2022) and Israel (Efrat, 2023). We have incorporated 88 catchments from LamaH-Ice that experience no or low natural or anthropogenic influence into the Caravan dataset. The Icelandic Caravan extension is available in the LamaH-Ice HydroShare repository.

We chose the LamaH structure over the CAMELS structure for a few reasons. Streamflow observations in Iceland are available for nested catchments, and hourly streamflow observations are available for a part of the time of coverage. Further, the LamaH-CE includes meteorological forcings that are available globally (ERA-5 Land, section 4.2) and uses a pan-European land cover classification (Section 5.4) as well as a European soil database (Section 5.6). Additionally, LamaH-CE includes an important attribute for hydrology in Iceland, namely the fraction of each catchment area covered by glaciers.

The LamaH-Ice dataset contains additional variables that have not been part of previous LSH datasets. Timeseries of catchment-aggregated MODIS snow cover and glacier albedo are included, as well as in-situ glacier mass-balance measurements for glaciated catchments. Topographical characteristics that affect glacio-hydrological processes are calculated specifically for the glaciated areas within catchments. The dataset also includes dynamic attributes to account for changes in land cover, vegetation, and glacier extent. The inclusion of such dynamic attributes is important in LSH datasets given changes in the environment and their potential impact on water resources.

This paper is organized as follows: the domain of coverage is described in Section 2, the basin delineation methods are described in Section 3, the compilation of hydrometeorological timeseries, remote sensing timeseries of snow cover and glacier albedo and mass balance is described in Section 4, and the catchment attributes are described in Section 5. The available hydrometeorological timeseries and catchment attributes are listed in Appendix A, Tables A1 to A12. Information and further details about the dataset and its data sources are presented in the Supplement. The timeseries and catchment attributes are available on HydroShare at http://www.hydroshare.org/resource/86117a5f36cc4b7c90a5d54e18161c91.

2. Domain of coverage and hydroclimate of the region

Iceland is located approximately 400 km east of Greenland in the North Atlantic Ocean, with the island's northernmost coast touching the Arctic Circle. Iceland has an area of 103,100 km^2 and is the most sparsely populated country in Europe (Eurostat, 2022), with inhabited regions limited to the coastal zone and lowlands. The country has a high volcanic activity which strongly influences the landscape and hydrology of the island.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Iceland has a mild climate given the country's high latitude. The island is warmed by a branch of the Gulf Stream that flows along its southern and western coasts. Moisture-laden cyclones crossing the Atlantic frequently pass close to or over the island, especially during winter, and the weather depends highly on their tracks (Einarsson, 1984). Precipitation varies significantly within the island. The southern part of the island is generally warmer and wetter than other regions, with the highest precipitation occurring on glaciers. Winter snowfall is more common in the northern part of the island than in its southern part. With limited evaporation and high precipitation rates, annual river runoff in Iceland has been estimated to be almost 4 times the world average (Gíslason, 2008).

The island's central highlands account for 40% of the island's area, and glaciers and ice caps cover about 10% of the island. Europe's largest non-polar icecap, Vatnajökull (~7,700 km^2), is located on the southeast of the island, and two other icecaps (Langjökull; ~835 km^2, and Hofsjökull; ~810 km^2) are located in the central highlands (Figure 1). Mýrdalsjökull (~598 km^2) is located close to the central southern coast (glacier area estimates are from Hannesdóttir et al., 2020). Seven other glaciers are larger than 110 km^2, as well as more than 250 smaller glaciers (Aðalgeirsdóttir et al., 2020). Meltwater from the glaciers provides at least one-third of the country's total runoff, feeding the main rivers (Björnsson and Pálsson, 2008). Some of them are utilized for hydropower production.

Seasonal snow and groundwater also strongly influence the hydrology of Iceland. Icelandic rivers have traditionally been divided into three categories by their origin: glacial rivers, direct-runoff rivers and spring-fed rivers, with many rivers being a mixture of these categories. In a classification of rivers of Nordic countries (Petersen et al., 2006), Icelandic rivers were considered as alpine/arctic. Watersheds of Icelandic rivers have widely varying geophysical attributes, which manifest themselves in very diverse hydrological responses to atmospheric forcings (Jónsdóttir and Uvo, 2009).

3. Catchment delineation and aggregation approaches

The streamflow gauges in LamaH-Ice are distributed across Iceland with a denser gauge placement in the south-western and eastern highlands where major hydropower facilities are located. The altitude of the gauges ranges from 5 to 812 m a.s.l. (Figure 1). The geographical coordinates of the gauges were obtained from the Icelandic Meteorological Office (IMO) and Iceland's National Power Company (NPC).

Topographical catchments of the gauges were delineated using the Pysheds python package (Bartos et al., 2020). The catchment delineation is further described in the Supplement (S1). As in LamaH-CE, "basin delineation A" refers to the entire upstream area of each gauge, comparable to the catchments in the CAMELS datasets. The total area covered by LamaH-Ice watersheds is 45,945 km^2, which corresponds to 45% of Iceland's area. The glacierized area covered by LamaH-Ice watersheds in 2019 is 6,202 km^2, corresponding to 60% of the total glacier area in Iceland. The locations of the gauges in LamaH-Ice and their total drainage area are shown in Figure 1.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

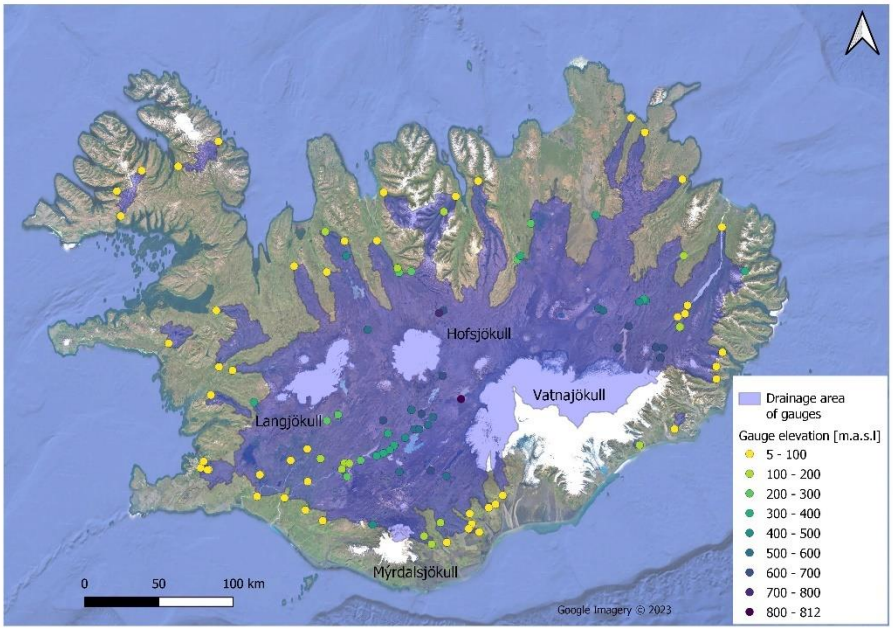

Figure 1: Locations of gauges included in LamaH-Ice and their drainage areas. The elevation of each gauge is indicated with colors. Made with QGIS. Source of background map: Google, 2023.

A total of 51 catchments in basin delineation A are located within other catchments. For a given gauge in basin delineation B, the catchment area of upstream gauges is subtracted, and thus intermediate catchments are represented. Table A1 describes the dependency among the connected catchments ("HIERARCHY", "NEXTUPID", "NEXTDOWNID"). Basin delineation C is like delineation B but excludes catchments with moderate or strong human or natural influence on streamflow (as explained in Section 5.9). The catchments in delineation C thus exhibit natural streamflow conditions. For a more detailed explanation of the difference between delineations A, B and C, see Klingler et al., 2021.

We performed aggregation of the meteorological dataset and the various spatially distributed geophysical data sources for each delineation method by calculating the area-weighted arithmetic mean. Note that for larger catchments, a mean value of meteorological variables and catchment characteristics is a great simplification and will not represent adequately the processes in the catchments. Thus, as the catchment area increases, it becomes more important to account for spatial variability.

4. Hydrometeorological time series

4.1 Runoff data

The first streamflow measurements in Iceland were conducted in the late 19th century (Helland, 1882). Systematic streamflow gauging started in the 1940's by the National Energy Authority of Iceland. Since 2009, the Icelandic Meteorological Office has been the government agency responsible for general hydrological research and streamflow measurements. The IMO operates a network of gauging stations. The National Power Company of Iceland (NPC) also operates a number of streamflow

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Open Access Earth System Science Data Discussions

gauges to study and monitor the water resource at their current and proposed hydropower facilities. The IMO provided data from 60 gauges, and the NPC provided data from 47 gauges.

All streamflow measurements are quality controlled in a consistent manner with quality codes and standard remarks. In LamaH-Ice, the streamflow measurements are provided along with the corresponding quality codes, allowing users to filter the measurements based on their reported quality. A pre-filtered version of the streamflow timeseries is also provided, which only includes data of high quality. The quality codes and filtering are further described in the Supplement (S2). Figure 2 illustrates the availability of daily streamflow measurements as a function of time.

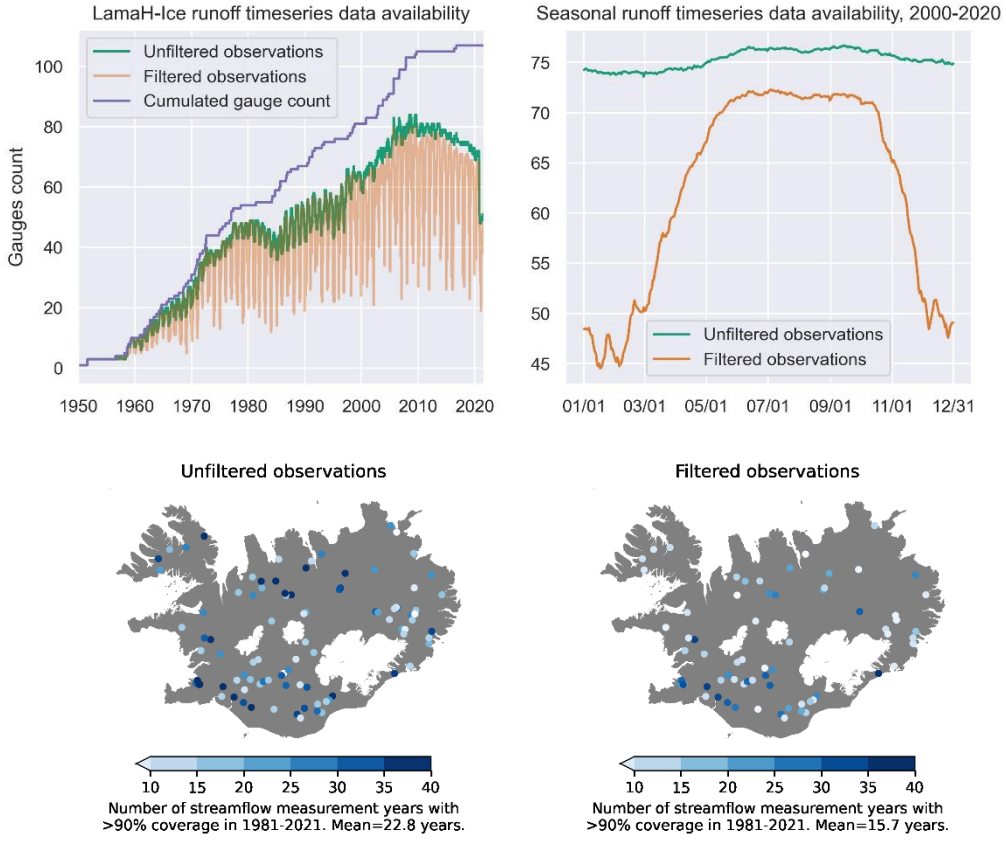

Figure 2: Gauge data availability in LamaH-Ice. Upper left figure: The availability of daily streamflow measurements as a function of time from 1950 to 2021. Filtered and unfiltered measurements are shown in different colors. Upper right figure: The average seasonal data availability for the period 2000-2020. Lower left figure: The spatial distribution of gauges with a minimum of three hydrological years with >90% data coverage between 1981-2021 for unfiltered observations (93 gauges in total). The color indicates the number of years available for each gauge. The mean temporal coverage of the 93 gauges is 22.8 years. Lower right figure: The same analysis for filtered observations. The mean temporal coverage of the 69 gauges is 15.7 years. Basemap source: Hijmans, 2015.

The number of actively (simultaneously) reporting gauges in LamaH-Ice increases with time up until 2008, when it reaches a high of 84. From 2006 to 2020 there are between 70 and 84 actively reporting gauges (Figure 2, upper left). Under winter ice conditions, streamflow measurements in Iceland are prone to interruptions, especially at gauges located at higher elevations,

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

resulting in lengthy spells of missing data during winter. Some gauges are only operated for a part of the year since year-around maintenance is not viable. The average seasonal data availability for the period 2000-2020 is shown in Figure 2 (upper right). Most of the active gauges during this period have year-round data coverage (unfiltered observations, blue line). However, after filtering out data with lower quality (estimated data due to ice interruptions or unchecked data), there is quite a reduction in data coverage during winter months, from an average of 72 gauges over the summer (June-August), to 48 gauges over the winter (December-February).

We used the 40-year period between October 1st 1981 and September 30th 2021 for water balance (section 4.2) and hydrological signature (5.3) calculations. Only hydrological years with at least 90% temporal coverage (using filtered data) during this period are considered, 69 gauges in total (gauges with less than three years of 90% coverage were omitted). The spatial distribution of these gauges and the number of valid years in each case is shown in Figure 2 (lower right figure).

4.2 Meteorological data

The meteorological data in the LamaH-Ice dataset is from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5-Land climate reanalysis dataset (Muñoz-Sabater et al., 2021). This is a re-run of the land segment of the fifth iteration of the European ReAnalysis (ERA5), with a finer spatial resolution (0.1°×0.1° on a regular latitude/longitude grid, approximately 5 × 11 km over Iceland). The dataset has global coverage, hourly time resolution and is continually updated with a latency of about 5 days. The ERA5-Land dataset contains 50 variables that describe hydro-meteorological processes over land. LamaH-Ice includes 15 of these variables at an hourly and daily resolution for the period 1950–2021 (Table A2). We obtained the ERA5-Land dataset through the C3S Climate Data Store (https://doi.org/10.24381/cds.e2161bac).

We used R scripts developed for the Lamah-CE dataset (Klingler et al., 2021) to aggregate ERA5-Land to watersheds. We calculated the area-weighted arithmetic mean of each meteorological variable for the watersheds. For watersheds that intersect multiple ERA5-Land cells, we calculated areal weights for each intersecting cell.

We also included meteorological data at a finer resolution from a reanalysis dataset (RAV-II; Rögnvaldsson, 2020). The RAV-II dataset used WRF to downscale the ERA-40 (1958-1979) and ERA-Interim (1979-2019) reanalysis onto a 2×2 km grid over Iceland. We also included precipitation from the Copernicus Arctic Regional Reanalysis (CARRA, 2.5×2.5 km grid; Schyberg et al., 2020) in LamaH-Ice. CARRA is produced with the HARMONIE-AROME weather prediction model, using ERA-5 as lateral boundary conditions, and is available from 1991 to present. The RAV-II and CARRA data were added due to an underestimation of precipitation in the rather coarsely gridded ERA5-Land dataset (Figure 3). The finer-resolution reanalysis datasets better capture the large orographic enhancement of precipitation in Iceland's complex terrain. Atmospheric reanalyses products inherently have uncertainties because they rely on numerical models and historical observational data, which both introduce sources of error and approximation in attempting to recreate past weather and climate conditions. The density of observations used for model assimilation varies across regions, resulting in varying levels of uncertainty in the model output across different areas. ERA5-Land does not directly assimilate observations but relies on indirect forcing through ERA5 which

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

uses assimilation. We do not provide quantitative uncertainty estimates of the meteorological time series in LamaH-Ice. Users interested in such estimates should consult the respective references. For example, uncertainty estimates for ERA5 and ERA5-Land are available, derived by using a 10-member ensemble of data assimilation. Uncertainty estimates are also available for the CARRA reanalysis.

We checked and verified the meteorological time series in LamaH-Ice by analyzing the components of the water balance. These were plotted (Figure 3a-c) for 69 gauges with a high temporal coverage of streamflow observations (Figure 2, lower right figure). The water balance equation on a catchment scale can be written as $P = R + ETA + \Delta S$, where P is the precipitation over the catchment, R is river discharge out of the basin, ETA is the total evapotranspiration and ΔS is the change in storage of water within the catchment. For non-glacierized catchments, over long timescales, the ΔS component can be considered negligible. Figure 3a shows $P - R$ plotted against ETA using weather variables from the ERA5-Land reanalysis. Each point on the graph corresponds to a single catchment. The points should align on the 1:1 line, but that is not the case. The modeled precipitation in both datasets seems to be insufficient, judging by the negative values of $P - R$ for a majority of catchments, with a mean of -1.43 mm/d for ERA5-Land and -0.62 mm/d for RAV-II. The precipitation from ERA5-Land is compared to the RAV-II reanalysis in Figure 3d.

A viable way to estimate actual precipitation is to use observed winter snow accumulation on glaciers, which is derived from in-situ measurements in fall and spring (section 4.4). Due to occasional winter thaw events, this should be a slight underestimation of the actual precipitation on the glaciers. A comparison is made between the observed winter accumulation for Vatnajökull (Figure 3e) and Langjökull (Figure 3f) and the simulated winter precipitation on the glaciers from ERA5-Land, CARRA and RAV-II. CARRA and RAV-II show a significantly stronger agreement with the observed data than ERA5-Land. CARRA even shows higher precipitation values than are observed, especially for Langjökull Glacier. Given that the observations themselves underestimate the actual precipitation, the CARRA precipitation product is likely the most accurate out of the three reanalyses considered. This shows that the negative $P - R$ in the water balance analysis is largely due to an underestimation of precipitation in ERA5-Land, which is improved in RAV-II and even further in CARRA. Another possible contributing factor of negative $P - R$ is a negative glacier mass balance, that is, a net loss of glacier ice, which has been the case for Icelandic glaciers since 1994 (except for 2015). However, the glaciers gained mass between 1980 and 1993 (Aðalgeirsdóttir et al., 2020). Omitting the 47 basins with >5% glaciation results in a mean $P - R$ value of -1.23 mm/day for ERA5-Land and -0.62 mm/day (no change) for RAV-II. Another possible reason for negative $P - R$ could be that streamflow measurements for up to 10% of the hydrological year are missing from many gauges used in this analysis, generally during the winter when discharge is low. This results in slightly larger average annual streamflow values than if the full year would be used. Other possible contributing reasons could be discrepancies between the topographic watersheds and groundwater watersheds of the gauges, or errors in streamflow measurements.

To maintain consistency between LamaH-Ice and other large-sample datasets using ERA5-Land, we did not attempt to correct the underestimation of precipitation in the reanalysis. For local studies in which accurate precipitation estimates are desired, we recommend using CARRA or RAV-II.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

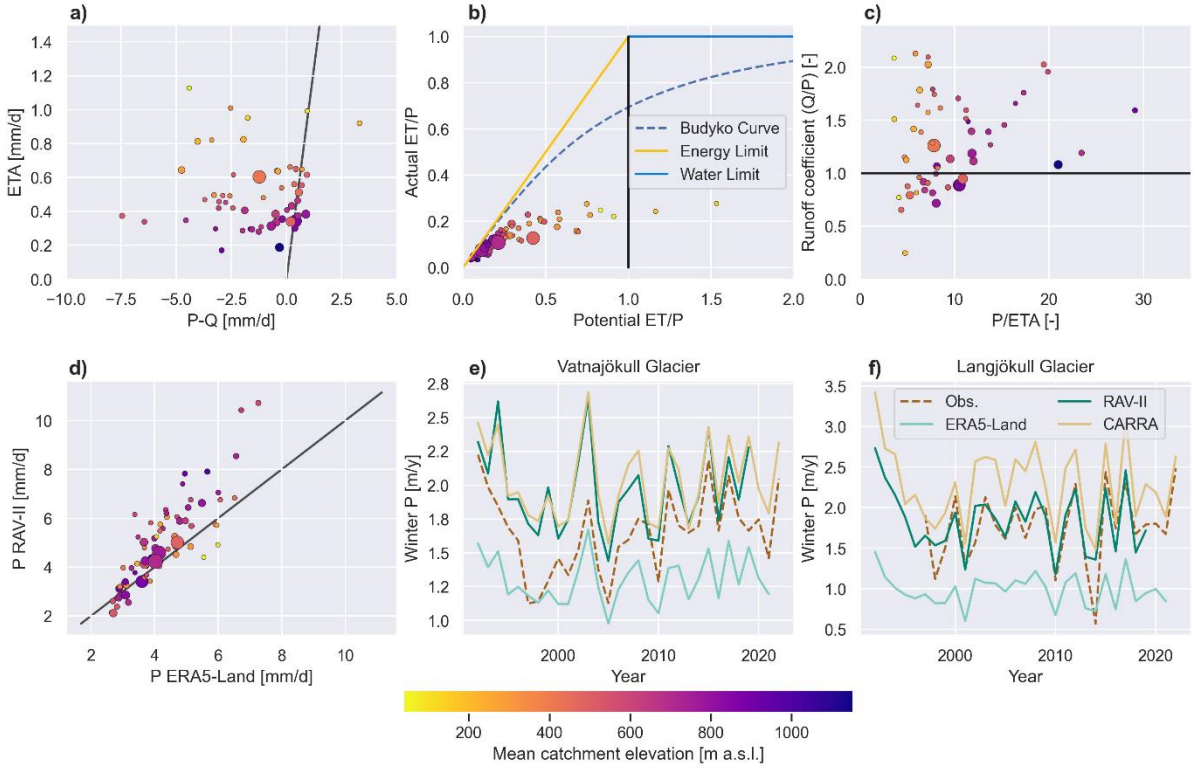

Figure 3: A water-balance analysis of the meteorological time series. Each point in panels a-d corresponds to a single catchment. Meteorological data from ERA5-Land are used in panels a-c. The size of the points indicates catchment size, and the color indicates mean catchment elevation. a) P-R plotted against actual ET (ETA) for the period 1980-2021. b) A Budyko curve analysis for all catchments in LamaH-Ice. c) The runoff coefficient plotted against P/ETA. d) A comparison between precipitation from ERA5-Land and RAV-II for the period 1991-2019. e) Vatnajökull Glacier: Winter precipitation from ERA5-Land, RAV-II and CARRA compared to measured winter accumulation. Winter precipitation is defined as precipitation falling between Oct 15 and May 1 each year. d) The same precipitation comparison for Langjökull Glacier. In panels a), b) and c) only hydrological years with >90% streamflow measurement coverage are considered, and black lines indicate physical constraints. Gauge measurements with a strong (human or natural) influence (section 5.9) are omitted.

The Budyko curve (Figure 3b) describes the relationship between actual evapotranspiration, potential evapotranspiration and precipitation in a catchment using the ERA5-Land meteorological data. Again, each point on the graph corresponds to a single catchment. The figure shows that all but two catchments (using ERA5-Land precipitation) fall on the energy limited portion of the Budyko curve (four in the case of RAV-II precipitation). The points should fall along the Budyko curve, but they fall largely to the right of the curve, which is explained by insufficient precipitation in the ERA5-Land and RAV-II reanalysis and possibly low total evapotranspiration values. However, the evapotranspiration shows a plausible dependence on elevation as in figure 3a, with high ETA in the warmer lower altitudes.

Figure 4c shows the ratio between measured runoff and modeled precipitation (runoff coefficient; Q/P) against the ratio of mean precipitation to evapotranspiration (P/ETA). The runoff coefficient is larger than one in 46 (ERA5-Land) and 38 (RAV-II) out of 69 catchments, again mostly due to the insufficient precipitation in the reanalysis.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

4.3 Remote sensing observations of snow cover and glacier albedo

Snow and glacier melt strongly influence the streamflow regime of Icelandic rivers. Timeseries of catchment fractional snow cover and average blue-sky albedo observations from the NASA MODIS sensors at 500 m resolution are included in LamaH-Ice. Iceland's 75% average annual cloud cover (with some variability depending on location), and the limited daylight hours in winter (Gunnarsson et al., 2019) limit the use of non-cloud penetrating and visible spectrum remote sensing observations. Gunnarsson et al. (2019) and Gunnarsson et al. (2021) produced temporally continuous cloud-free datasets of MODIS observations for Iceland by merging data from the Aqua and Terra satellites and temporally aggregating the observational series. The first is a gap-filled MODIS snow cover product covering all of Iceland (Gunnarsson et al., 2019). The second is a MODIS glacier albedo product (Gunnarsson et al., 2021). We used these datasets to create snow cover and glacier albedo timeseries for catchments in LamaH-Ice from 2000 to 2021.

For catchment fractional snow cover, two daily timeseries are provided, a catchment average and an average for land area outside of glaciers. For albedo, the daily average value for the glacierized portion of the watershed is provided. The snow cover timeseries are available for all catchments in LamaH-Ice, and glacier albedo timeseries are available for all 61 catchments that have some degree of glaciation (as of 2019).

While the MODIS-derived data presented here have been validated and exhibit strong agreement with in-situ observations, they are not without uncertainties. The temporal aggregation of the data smooths out short-term variations and nuances in snow cover and albedo. This can limit its suitability for certain use cases, although in remains representative for weekly to monthly applications (Gunnarsson et al., 2021). The spatial resolution of the MODIS data may not capture fine-scale variations, especially in complex terrain. This will cause greater uncertainty in the snow cover and albedo timeseries for smaller catchments as well as catchments in steep terrain. ~~Residual cloud cover and atmospheric interference may affect the data.~~ It is well known that MODIS data can suffer from cloud-snow confusion. To reduce cloud artifacts in the MODIS albedo dataset, robust statistics like median-based outlier removal were used. However, overly strict rejection criteria may lead to the loss of valid data, particularly when albedo experiences rapid fluctuations (Gunnarsson et al., 2021). Further, changes in glacier extent over the study period could introduce uncertainties in glacier albedo estimates. Users should be aware of these potential sources of error when utilizing these timeseries.

4.4 Glacier mass balance measurements

Annual in-situ mass-balance measurements have been made on the largest glaciers in Iceland for decades (Aðalgeirsdóttir et al., 2020). Winter snow accumulation is measured in spring and summer ablation is measured in fall. These measurements are conducted at multiple locations situated along ice flow lines and cover the altitudinal range of the glaciers. Digital mass balance maps have been derived by interpolating the in-situ measurements and using observed mass-balance gradients (Björnsson et al., 2002). Uncertainties in this data may arise due to the spatial interpolation as well as the representativeness of the selected measurement sites. Additionally, variations in glacier dynamics and processes not fully captured by the measurements could

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

contribute to some level of uncertainty in the derived mass-balance timeseries. In LamaH-Ice, the available mass balance maps from Vatnajökull and Langjökull were used to create annual timeseries of mass-balance changes within each catchment draining the two ice caps, 32 catchments in total.

5. Catchment attributes

The catchment characteristics in LamaH-Ice were assembled using global or pan-European datasets. The attributes and their data sources are consistent with the LamaH-CE dataset. In addition to that, local data sources with finer resolution were used to include more accurate information, as a supplement to the characteristics provided in the LamaH-CE dataset. For consistency and comparability with the LamaH-CE paper, we chose color maps for most of the plots in this section that closely match the ones used in the LamaH-CE paper (Klingler et al., 2021). Maps in this section use a basemap shapefile from Hijmans, 2015.

5.1 Topographic indices

LamaH-Ice includes ten catchment attributes related to topography (Table A3). We used the digital elevation model IslandsDEM version 1.0 (National Land Survey of Iceland, 2020) at 20 m resolution to derive 6 elevation related attributes (catchment area, mean, median, range and standard deviation of elevation within the catchments, as well as catchment mean slope).

The area of the LamaH-Ice catchments (basin delineation A) ranges from 4 to 7437 km^2 with a mean of 901 km^2 and a median of 384 km^2. The mean elevation of the catchments ranges from 38 to 1296 m a.s.l. (Figure 4b). The highest mean catchment slopes are found in catchments close to the coast, and the slope is lower in catchments located in the central plateau of the island (Figure 4d).

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

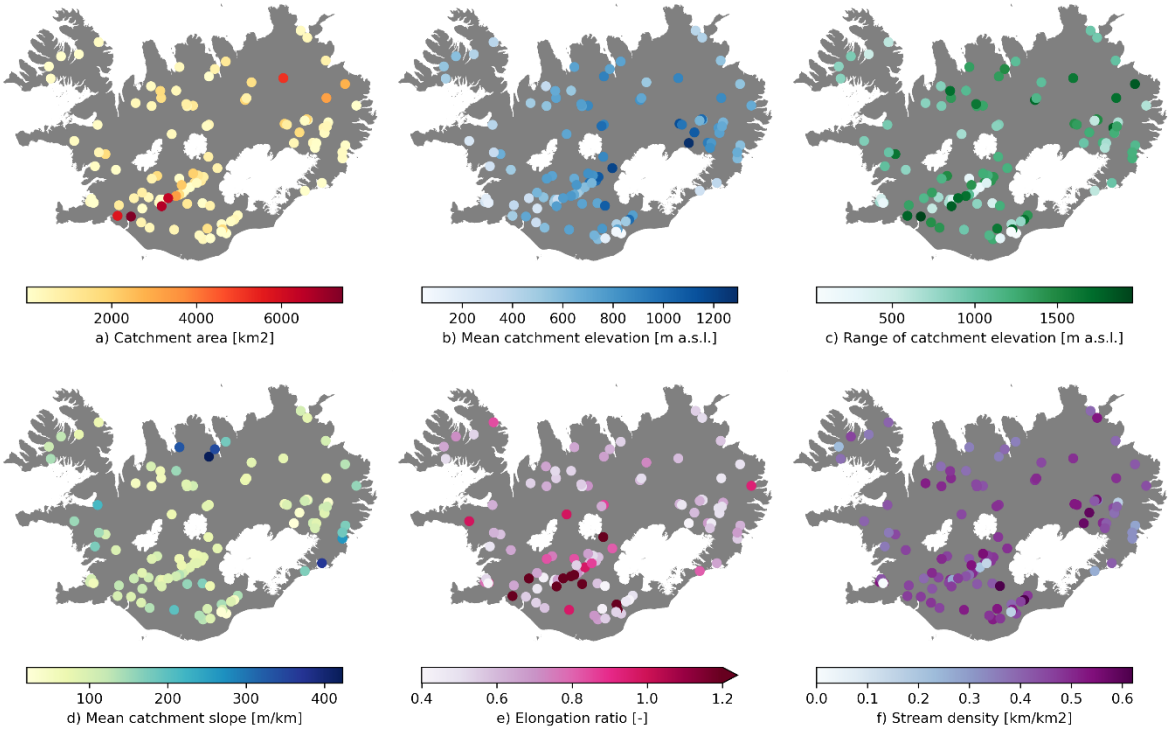

Figure 4: Spatial distribution of 6 out of 10 topographical attributes for the catchments in LamaH-Ice. Basemap source: Hijmans, 2015.

In addition to attributes related to elevation, LamaH-Ice includes four other attributes, for consistency with the original LamaH-CE dataset. These attributes describe the catchment shape and orientation, as well as the stream network within the catchment.

5.2 Climate indices

LamaH-Ice includes a total of 10 climate indices for all catchments (Table A4). These indices are computed over hydrological years 1990 to 2009 as in the existing CAMELS and LamaH datasets. The climate indices were calculated using time series of precipitation, temperature, and PET from the ERA5-Land dataset. The indices describe long-term, seasonal and short-term characteristics of the climate. Figure 5 shows the spatial distribution of the climate indices.

Indices describing long-term climatology include daily mean precipitation ("p_mean" in Table A4), potential evaporation ("peta_mean") and the aridity index ("aridity"). Watersheds in the south of Iceland experience the highest levels of precipitation, with levels gradually decreasing towards the north (Figure 5a). Evapotranspiration levels in Iceland are low due to cool temperatures and high humidity. Figure 5b shows that potential evapotranspiration levels follow a pattern similar to precipitation, with higher levels in the warmer southern region of Iceland. With low potential evaporation (median value of 0.7 mm/day) and high precipitation, the aridity index is also low, which indicates a low degree of dryness compared to other regions with similar amounts of precipitation.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

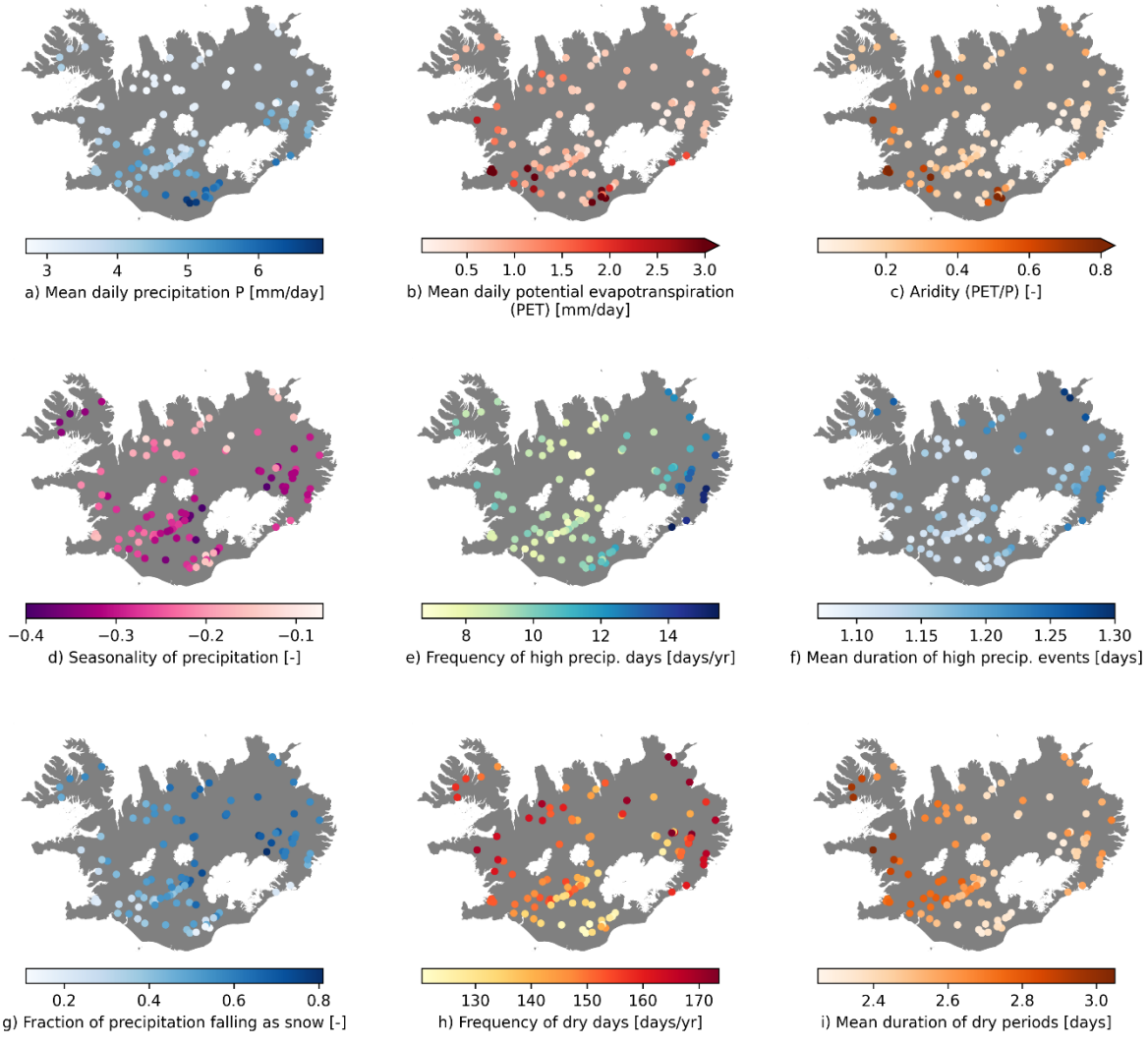

Figure 5: Spatial distribution of climate indices for LamaH-Ice watersheds. The indices are computed from ERA5-Land data for (hydrological) years 1990 to 2009. Basemap source: Hijmans, 2015.

Indices describing seasonal characteristics of the climate include the snow fraction of precipitation ("frac_snow" in Table A4) and the seasonality of precipitation index ("p_season"). Positive values of this index highlight summer as the peak precipitation season, and negative values emphasize winter. In Figure 5d ~~we see that~~ the seasonality index is negative for all regions. The largest negative values are in the highlands, and in the east and west. These are also the watersheds where the largest fraction of the annual precipitation falls as snow (Figure 5g).

Indices that provide information on short-term weather events include the frequency of high-precipitation days ("hi_prec_fr" in Table A4) and dry days ("lo_prec_fr"), the average duration of high ("hi_prec_du") and low ("lo_prec_du") precipitation days and the most common season they occur ("hi_prec_ti" and "lo_prec_ti"). The frequency of high-precipitation days is

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

highest in the eastern part of Iceland and decreases towards the west (Figure 5e). High precipitation intensity events are common in the east fjords, especially in the fall, when extratropical cyclones approach the country from the east or southeast. The heavy precipitation is caused by strong orographic uplift. The central region of Iceland experiences lower precipitation intensity as the air in this region has already shed much of its moisture due to the orographic uplift at the coast. Even though areas in the central region of the island experience few high-precipitation days, they also see few dry days (Figure 5h), which are more common in the east and the west. The duration of dry periods is longer in the west (Figure 5i).

5.3 Hydrological signatures

A total of 13 hydrological signatures were calculated to characterize the observed streamflow timeseries (Table A5). As with the climate indices, the streamflow signatures describe long-term, seasonal and short-term characteristics of the hydrological system. Figure 6 shows the spatial distribution of hydrological signatures for the 69 gauges in LamaH-Ice that have a high temporal coverage of streamflow observations (section 4.1) and exhibit natural streamflow conditions (section 5.9). The mean daily discharge (Figure 6a) is highest along the southern coast of Iceland, where precipitation is the highest. The runoff ratio is derived by dividing the mean daily streamflow by the mean daily precipitation. It thus represents the fraction of the precipitation that exits the catchment via the stream channel. However, as reported in section 4.2, the precipitation from ERA5-Land used here is biased low, and thus the computed runoff ratio is unrealistically high (Figure 6b) with a mean of 1.3. A few high outliers include glaciated basins as well as spring-fed rivers.

The relationship between changes in streamflow and changes in precipitation on annual timescales is represented via the runoff-precipitation elasticity ("stream_elas", Figure 6f). The higher the value, the more we expect runoff to increase due to increases in precipitation. The mean value for rivers in LamaH-Ice is 0.7, which means that a 10% increase in precipitation would on average lead to 7% increase in streamflow. However, we see quite a spread in the elasticity values for the catchments in LamaH-Ice.

We separated the baseflow component from the total runoff using the method by Ladson et al. (2013). The ratio of baseflow to total runoff ("baseflow_index_ladson") is shown in Figure 6e. The figure shows that the rivers draining the young and porous bedrock along the volcanic rift zone have a high baseflow index, and rivers draining the low permeability bedrock in the western and eastern regions have a low baseflow index.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

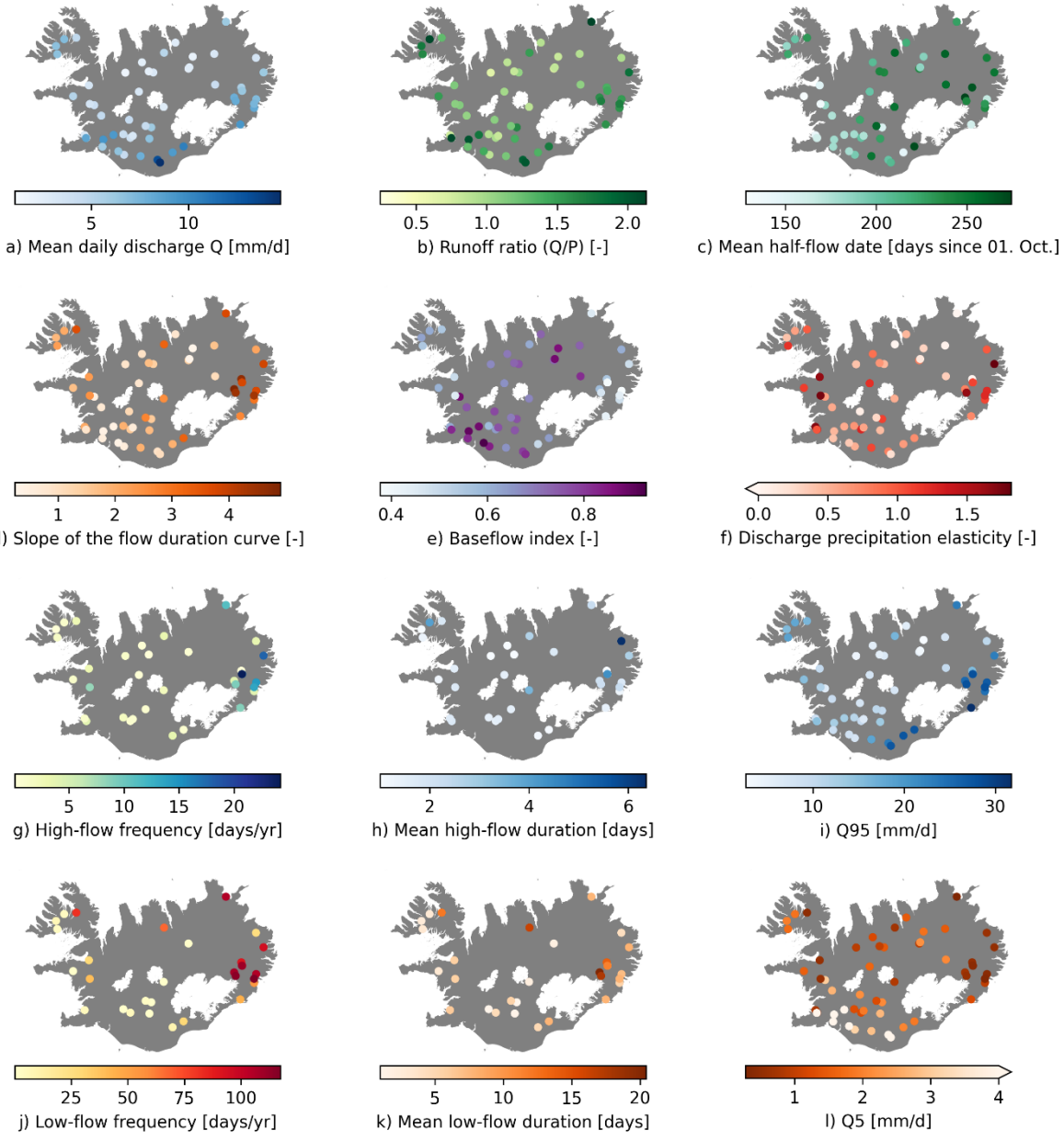

Figure 6: Spatial distribution of hydrological signatures. The signatures were calculated for the 40-year period between October 1st 1981 and September 30th 2021. Only hydrological years with at least 90% temporal coverage during this period were considered. Gauges with less than three years of 90% coverage were omitted. The total number of gauges shown is 69. Panels g, h, j, and k describe the frequency and duration of high- and low-flows. The number of gauges shown in these panels is lower since high- and low-flows, as they are defined here, do not occur in all gauges. Basemap source: Hijmans, 2015.

~~Two streamflow signatures are calculated that offer~~ insights into the seasonality of streamflow. The mean half-flow date ("hfd_mean") is shown in Figure 6c. Rivers draining the largest glaciers generally have a high mean half-flow date, as well as

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

snow-melt dominated rivers in the northern part of Iceland. The slope of the mid-section of the flow duration curve ("slope_fdc") describes the general flashiness of streamflow, with higher values indicating greater variability within the year. We see a higher flow duration curve slope for basins with a low baseflow index (R=-0.83).

To describe short-term extremes in streamflow, 7 indices are calculated that relate to high and low flows. For this purpose, high flows are defined as at least 9 times the median daily streamflow (Clausen and Biggs, 2000) and low flows are defined as flows that are less than 20% of the median daily streamflow (Olden and Poff, 2003). The frequency of these events is calculated ("high_q_freq" and "low_q_freq"), as well as their duration ("high_q_dur" and "low_q_dur"). High and low flows, as they are defined here, do not occur in all gauges in LamaH-Ice. Panels g, h, j, and k ~~thus~~ include fewer gauges than other panels in Figure 6. The magnitudes of high and low flows are described by calculating the 95th (Q_{95}) and 5th (Q_5) streamflow percentiles (Figure 6i and j). The magnitude of high flows (Figure 6i) is largest in the central south and southeast of Iceland, where precipitation is the highest. The magnitude of low flows (Figure 6l) is highest for rivers that have a high baseflow component (Figure 6e). The magnitude of low flows is low in direct runoff dominated catchments in the east and west of Iceland, and where the contribution of glacier melt is high.

5.4 Land cover characteristics

LamaH-Ice includes 7 attributes that describe the land cover (Table A6). All are based on the pan-European Coordination of Information on the Environment (CORINE) Land Cover dataset (Büttner, 2014). Two versions of CORINE land cover characteristics are provided in LamaH-Ice, a static version and a dynamic version. The static version uses the most recent CORINE update (2018). The dynamic version uses the CORINE classifications from 2000, 2006, 2012 and 2018, thereby reflecting the changes in land cover in Iceland since 2000. The dynamic version is set up as timeseries for each attribute, interpolating linearly between the 4 CORINE updates.

In LamaH-Ice, the dominant land class within each catchment ("lc dom" in Table A6) is calculated as the class with the largest area in the catchment. The areal fractions of agricultural areas ("agr_fra"), bare areas ("bare_fra"), forested areas ("forest_fra"), glaciers ("glac_fra"), water bodies ("lake_fra") and urban areas ("urban_fra") are also calculated. The static land cover attributes are shown in Figure 7. Longer timeseries of glacier fractions, that date back to 1890, are also made available (described in Section 5.8), however the CORINE glacier fractions were included for consistency with the LamaH-CE dataset.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

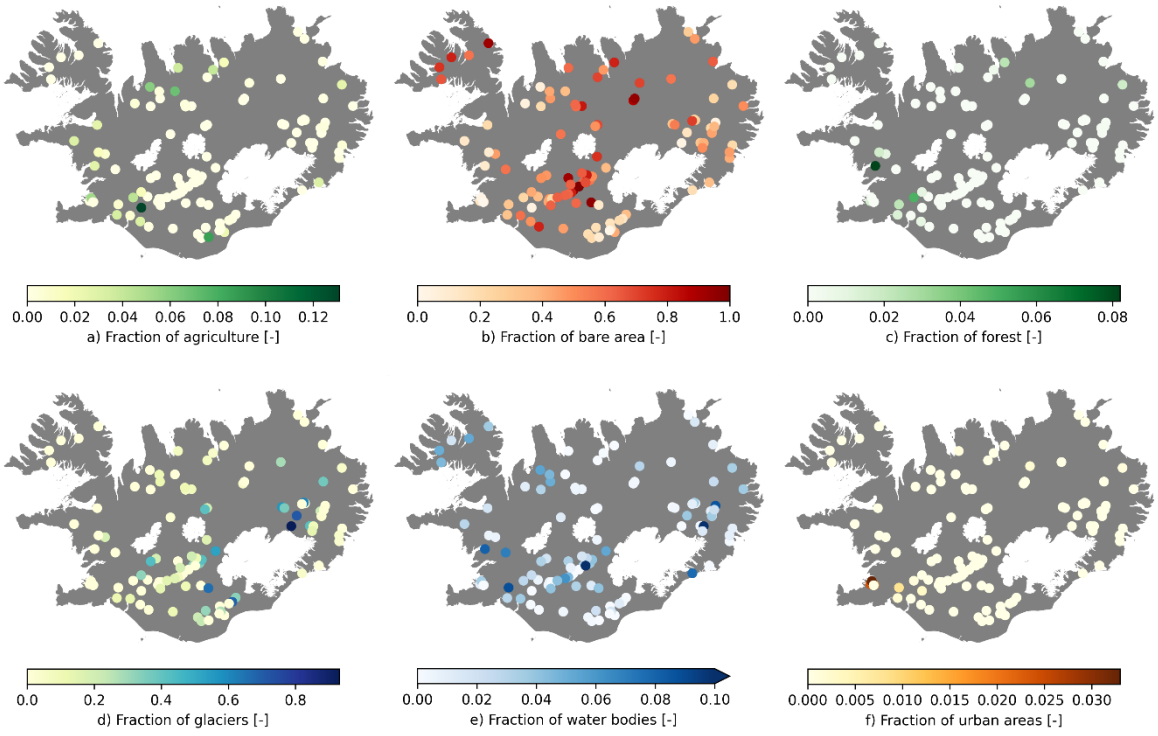

Figure 7: Land cover classification from CORINE for the LamaH-Ice catchments. Basemap source: Hijmans, 2015.

In most European countries, CORINE classifies the majority of land as forested, agricultural and urban areas. As can be seen in Figure 7a and f, this does not apply in Iceland, where urban areas are only 0.39% of the area and agricultural areas only 2.6%. In turn, the percentage of natural areas in Iceland is by far the highest in any European country. The two of the five main CORINE categories that describe natural areas, forested and semi natural areas and wetlands cover 95% of the country's area (Árnason and Matthíasson, 2020). ~~It should be noted that forests do not contribute much to this number, since forested areas are largely non-existent in Iceland (Figure 7c).~~ Currently, natural forests only cover 1% of the island (Raynolds et al., 2015). In turn, bare areas are quite common in the LamaH-Ice catchments (Figure 7b). Here, bare areas are defined as the two CORINE natural area classes "bare rock" and "sparsely vegetated areas". The fraction of glaciers is high for gauges in the vicinity of the three largest ice caps (Figure 7d).

5.5 Vegetation indices

Vegetation plays an important role in the hydrological cycle. We processed remote sensing observations in the Google Earth Engine platform (Gorelick et al., 2017) to derive six static catchment characteristics relating to vegetation (listed in Table A7), all of which are also included in LamaH-CE.

The vegetation characteristics are based on three vegetation indices, leaf area index (LAI), green vegetation fraction (GVF) and normalized difference vegetation index (NDVI). The formulation of these indices is further described in the Supplement

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

(S3.4) and in the LamaH-CE data description paper (Klingler et al., 2021). The static characteristics are based on monthly means of these indices over an extended period, ~~i.e., one mean value for each month of the year~~. For all three, we calculated the maximum monthly mean (out of 12 monthly means), as well as calculating the minimum monthly mean (for NDVI) or the difference between the lowest and highest means (for LAI and GVF). The spatial distribution of the vegetation characteristics (maximum monthly LAI, NDVI, GVF) is shown in Figure 8.

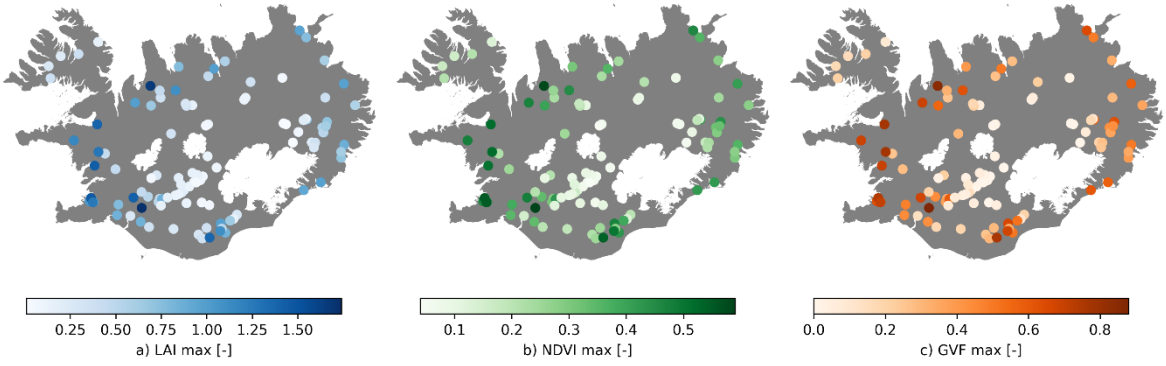

Figure 8: The spatial distribution of static vegetation indices for the catchments in LamaH-Ice. Basemap source: Hijmans, 2015.

The short growing seasons and cool temperatures in Iceland limit the growth and abundance of vegetation. This is shown by the low LAI_max and NDVI_max values for the catchments in LamaH-Ice, which reach a maximum of 1.7 and 0.6, respectively (Figure 8). Winter vegetation activity in Iceland is low and the minimum measured LAI and NDVI values for the LamaH-Ice catchments are close to zero. The maximum GVF values are quite high in lowland areas, which shows that the volcanic soils in Iceland (section 5.6) are quite fertile.

5.6 Surface deposits and soil characteristics

Many rainfall-runoff processes take place in the subsurface. It is thus important to include subsurface characteristics such as those that relate to unconsolidated surface deposits and soils in LSH datasets. At the end of the last glacial period, about 10,000 years ago, glaciers receded, and exposed erosion products that form most of the surface deposits currently found in Iceland. The primary surface deposits are glacial till, ancient deltas formed by higher sea levels after the glacial period, landslides, flood plains and sediment deposited by modern surface waters (Pétursson and Bjarnason, 2022). Soils in Iceland are mostly Andisols, which is a soil type found in active volcanic areas, characterized by having large proportions of volcanic glass (Arnalds and Óskarsson, 2009).

In Lamah-Ice, we include 9 attributes that describe the properties of the soils, regolith and sediment in Iceland (Table A8). We calculated the attributes in the same manner as in the LamaH-CE dataset. We derived the depth to bedrock attribute from the Global 1-km Gridded Thickness of Soil, Regolith, and Sedimentary Deposit Layers (GGT; Pelletier et al., 2016). According to this database, the depth to bedrock in Iceland varies widely depending on the location (Figure 9a).

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Open Access Earth System Science Data Discussions

485 Other attributes relating to soils were extracted from the European Soil Database Derived data (ESDD; Hiederer, 2013a, b). The fraction of clay, silt, sand, gravel and organic material in the soils of the catchments is shown in Figure 9c-g. Sand has the highest mean fraction over all catchments (62%), followed by silt (24%), gravel (16%) and clay (13%). The depth available to roots is generally high in the soils of Iceland (Figure 9b), which is a characteristic of Andisols soils, due to their low cohesion. In catchments with a high depth available to roots, there is a high portion of sand (R=0.95), and soil porosity is low (R=-0.89).
490 The mean organic fraction in the soils of catchments in LamaH-Ice is 1.4%.

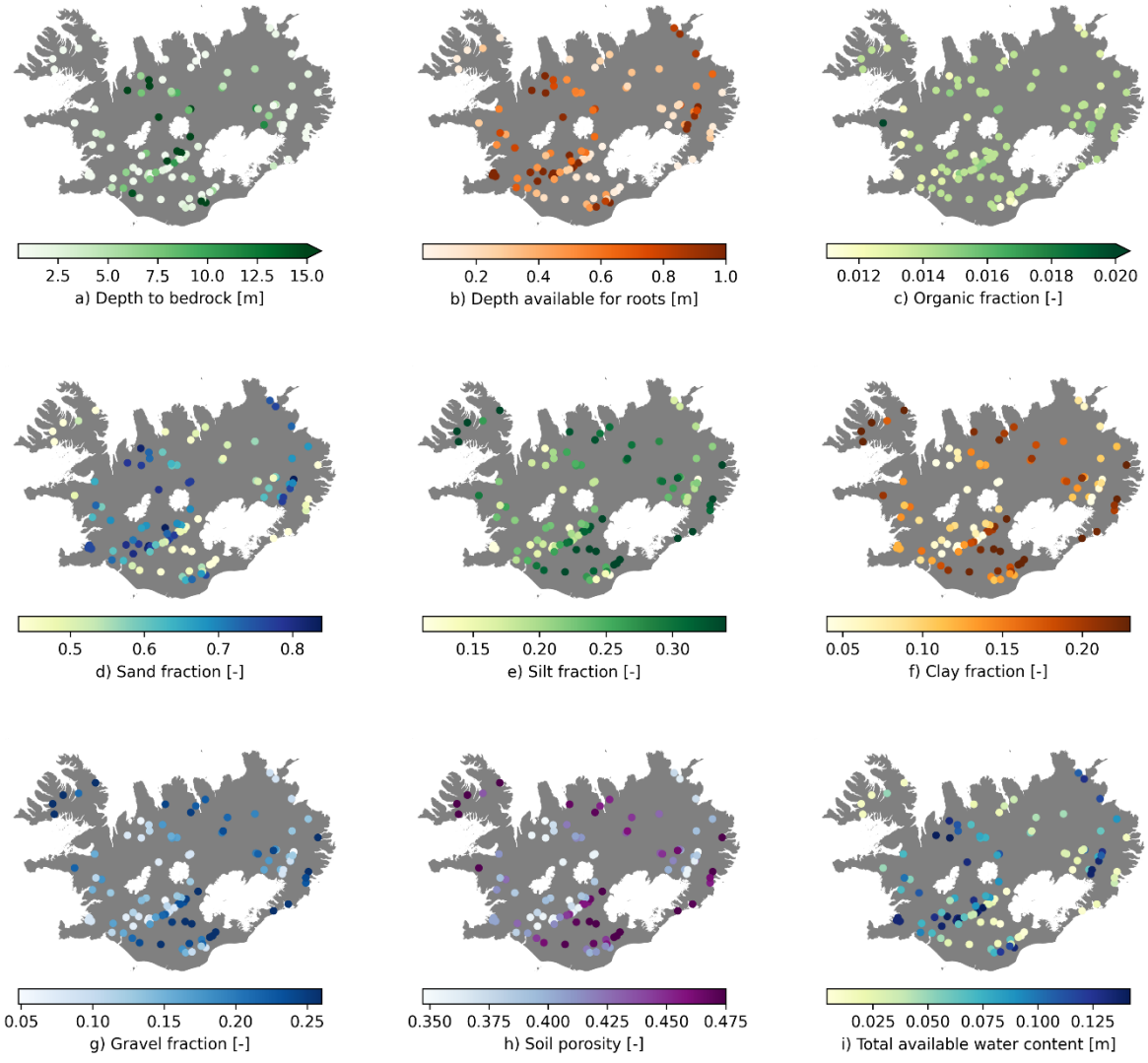

Figure 9: Spatial distribution of soil attributes for the catchments in LamaH-Ice. Basemap source: Hijmans, 2015.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

The soil attributes aggregated to catchments in LamaH-Ice are subject to various uncertainties and should be interpreted with caution. The spatial resolution of the underlying databases is quite coarse for Iceland. The ESDB is based on soil profiles that are not necessarily representative for soil conditions in all areas. Further, the soil depth is limited to 1.5m (ESDB) which is a great simplification.

5.7 Geological characteristics

Iceland is a volcanically active island crossed by a volcanic rift zone that lies from the SW corner of the island to the NNE part. The bedrock conditions strongly affect the hydrology with large contrasts within the island. The bedrock on the rift zone is young and porous and easily erodible, and permeability is high. Rivers that drain these areas generally have a high baseflow component. Conversely, rivers draining areas with older, low permeability bedrock on the eastern and western parts of the island are mostly surface-fed. Due to the variability of geological formations in Iceland, and the extent to which bedrock conditions affect hydrology, we chose to include more detailed geological information in LamaH-Ice than in the LamaH-CE or CAMELS datasets (Table A9).

To calculate geological characteristics of catchments, a global lithological map, GLiM (Hartmann et al., 2012, further explained in the Supplement, S3.6), and a geological map of Iceland at a much finer spatial resolution (Icelandic Institute of Natural History, 2014) were used. The average scale of GLiM is 1:3,750,000. The spatial distribution of the dominant rock attributes from GLiM (levels 2 and 3, Figure 10a) shows that basic volcanic rock is the most common type in Iceland, with variations depending on catchment location. The bedrock in the eastern and western regions of Iceland is characterized by subordinate sedimentary rocks, formed by the accumulation and consolidation of sediment over time, whereas the bedrock near the rift zone is characterized by the presence of pyroclastics, which are volcanic materials that are ejected during explosive volcanic eruptions.

The geological map of Iceland (Icelandic Institute of Natural History, 2014) is in the scale 1:600,000. It contains 16 classes that describe the main characteristics of Iceland's bedrock geology. Bedrock is classified by type, composition, and age. Glaciers and lakes are classified as such. The spatial distribution of the dominant geological class from this geological map is shown in Figure 10b. The eastern and western regions are characterized by extrusive rocks that are older than 3.3 million years (m.y.). Younger extrusive rocks are found towards the center of the island, along with different types of lavas.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

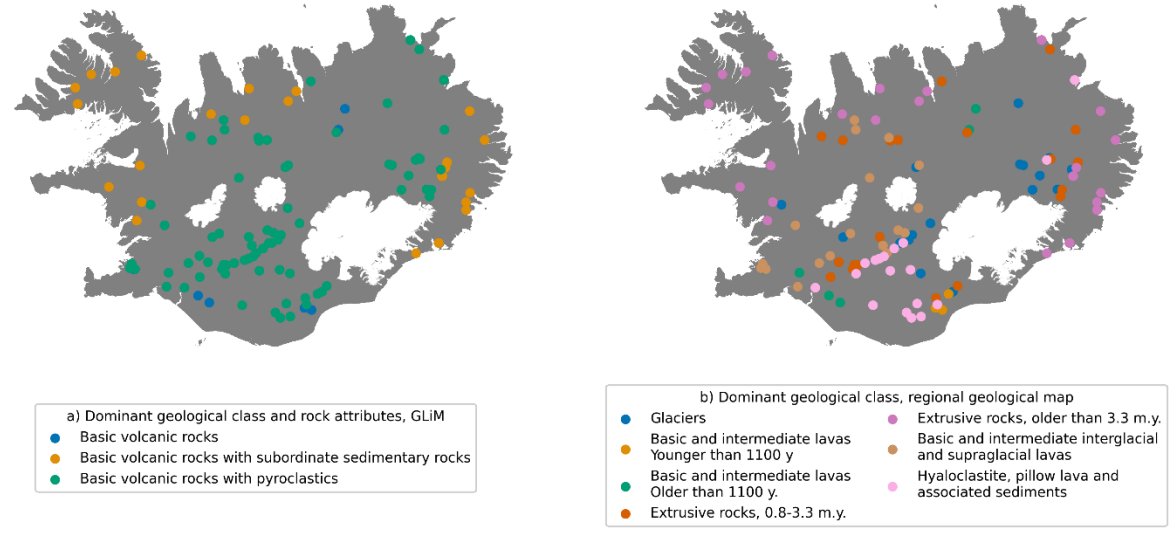

Figure 10: The dominant geological class of each catchment in LamaH-Ice, as depicted in a) the GLiM global lithological map, levels 2 and 3 (Hartmann et al., 2012)**, and b) a regional geological map of Iceland** (Icelandic Institute of Natural History, 2014)**. Basemap source: Hijmans, 2015.**

For LamaH-CE, information regarding bedrock permeability and porosity was derived from GLHYMPS (Gleeson et al., 2014). This database does not contain realistic values for Iceland and thus these parameters are not included in LamaH-Ice. In Lamah-Ice, the dominant geological class for each catchment is specified for the geological classifications described above. The fraction of each class is also specified.

5.8 Glacial characteristics

Sixty-eight out of the 107 catchments in LamaH-Ice are partly covered by glaciers. Nine glacial attributes are provided in the dataset (Table A10). The glacier covered fraction of each catchment area and the glaciated area are included both as static and dynamic attributes. The dynamic attributes are calculated from a national inventory of glacier outlines (Hannesdóttir et al., 2020). The inventory consists of historical reconstructions of glacier extents from various sources at multiple times from 1890 to 2019. The inventory is openly available as part of the Global Land Ice Measurements from Space (GLIMS) glacier database (nsidc.org/glims). The reconstructions reveal that Icelandic glaciers lost 18% of their area during this time (Hannesdóttir et al., 2020). In LamaH-Ice, we created timeseries of glacier covered fraction and glaciated area by linearly interpolating between the dates of the historical reconstructions. In 1950, the 57 catchments in LamaH-Ice with over 2% glaciation have a mean glaciation of 27.8% (372 km^2), which is reduced to 24.6% (336 km^2) in 2019. The highest decrease in percent glaciation is 8.9% (27 km^2), while the highest areal decrease is 137 km^2 (2.1%). Six catchments lost all their glaciated area between 1950 and 2019. Static glacial area and fraction attributes are based on the most recent glacial outline from the inventory (2019).

Various topographical qualities of glaciers affect their glacio-hydrological processes. Sensitivity studies performed in the Alps have shown that the lowermost 20% area slope, the mean elevation, aspect and easting and northing of the glacier centroid

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

have statistically significant relationships with changes in the glacier equilibrium line altitude (ELA) and with changes in the annual surface mass balance (Rabatel et al., 2016, 2013; Bolibar et al., 2020). These attributes, as well as mean glacier slope and maximum and minimum glacier altitude (Table A10), are included in LamaH-Ice for each glaciated catchment. We used the IslandsDEM version 1.0 DEM (National Land Survey of Iceland, 2020) to calculate the topographical attributes for the glaciated part of each catchment. Uncertainties in the DEM or watershed delineation method (described in section S1 in the Supplement) can introduce potential inaccuracies in the calculations of these topographic indices.

5.9 Natural or anthropogenic impacts on runoff

Two attributes are included in LamaH-Ice to indicate any possible anthropogenic impacts on runoff timeseries, as well as natural impacts such as the presence of an upstream lake. In LamaH-CE, storage and anthropogenic impacts on runoff are categorized into 13 classes. We adopted this classification but removed classes that were either not relevant in Iceland or for which relevant information was unavailable in the gauge metadata. The 6 remaining impact classes are shown in Table A11. The classes were assigned manually using areal imagery or publicly available information. As in LamaH-CE, the degree of impact is also provided, ranging from no influence to strong influence on runoff. Table A12 provides further information on the degree type allocation. The degree types are based on LamaH-CE with certain adaptations.

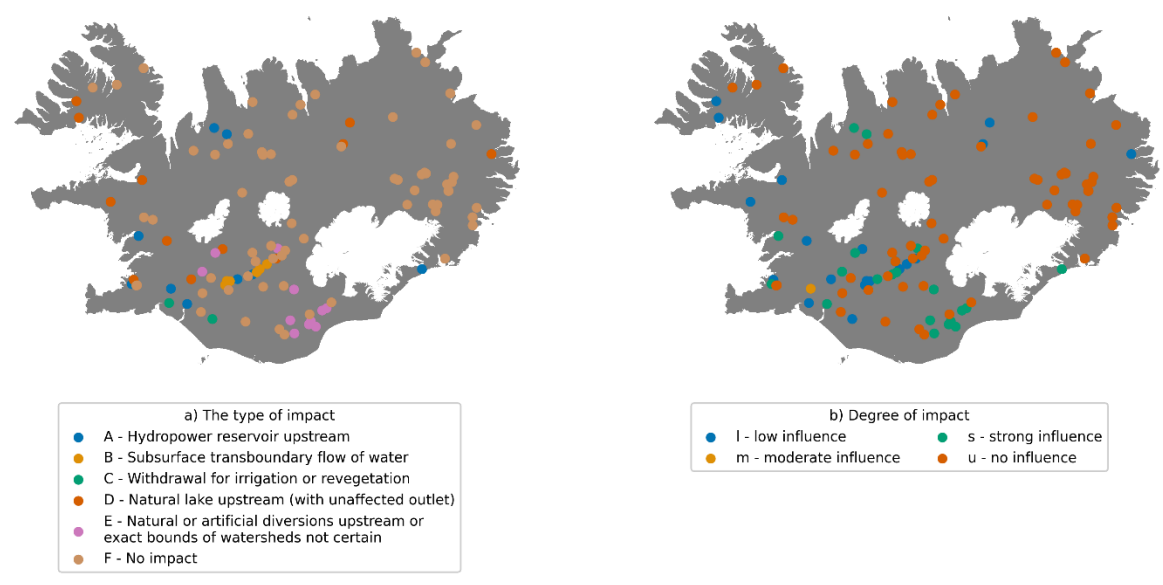

Figure 11: Spatial distribution of the type of impact (a) and degree of impact (b) on the streamflow measurements in LamaH-Ice. Basemap source: Hijmans, 2015.

Most gauges are not influenced (67 out of 107 or 63%) and thus exhibit natural and uninterrupted flow conditions (Figure 11). A total of 19 gauges (~~or~~ 18% of the total gauges) experience low natural or anthropogenic influence. Out of those 19 gauges, 12 gauges are only influenced by natural lakes upstream of the gauge. Thus, a total of 79 gauges (74%) exhibit natural flow conditions. One gauge experiences moderate influence, and 20 gauges (19%) experience strong influence.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

7. Data availability

The LamaH-Ice dataset is available for download from the HydroShare repository: http://www.hydroshare.org/resource/86117a5f36cc4b7c90a5d54e18161c91 (Helgason and Nijssen, 2023). Please consider the disclaimer and license which the data are subject to, stated on HydroShare. We offer three downloadable files: 1) "lamah_ice_hourly": The LamaH-Ice dataset containing hydrometeorological time series with both daily and hourly resolutions, 2) "lamah_ice" The dataset with hydrometeorological time series with daily resolution only and 3) "Caravan_extension_lamahice": The LamaH-Ice Caravan extension. For meteorological timeseries and catchment attributes in LamaH-Ice, three folders are supplied, one for each delineation method (A, B and C). Each folder contains three subfolders: "attributes", "timeseries" and "shapefiles". The "attributes" folder contains one .csv file with static catchment attributes. The "timeseries" folder contains meteorological timeseries and snow cover/glacier albedo timeseries as well as annual timeseries of glacier mass balance and extent and CORINE land cover change. The "shapefiles" folder contains the catchment as shapefiles and GeoPackages. A separate folder contains information about the gauges, including streamflow observations ("D_gauges"). A table showing the folder structure can be found in the Supplement (table S3).

8. Code availability

All code used to produce the LamaH-Ice dataset and plot figures in this paper is included in the GitHub repository of the dataset (https://github.com/hhelgason/LamaH-Ice), as well as in folder F in the dataset.

9. Conclusion

A new hydrological dataset for Iceland (LamaH-Ice) has been compiled, containing daily and hourly hydrometeorological timeseries and catchment characteristics for 107 river basins. The catchment characteristics describe the topographic, hydroclimatic, land cover, vegetation, soils, geological and glaciological attributes of the river catchments, as well as the human influence on streamflow in the catchments. In this paper, the dataset has been described and the distinctive attributes of Icelandic catchments have been showcased. The data in LamaH-Ice opens new possibilities to investigate the relationship between weather forcings, catchment attributes and streamflow in cold-region environments, with hourly streamflow measurements and a wealth of attributes relevant for cold regions. With the inclusion of glacier mass balance observations and dynamic catchment characteristics, LamaH-Ice provides new opportunities in evaluating the effects of changes in climate, land cover and glacier extent and mass on streamflow. This is important given the ongoing changes in these factors and their potential impact on water resources. A large majority of the gauged rivers in LamaH-Ice are unaffected by human activities, and thus the dataset can be valuable in improving our understanding of hydrological processes and our ability to manage water resources in cold regions.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Open Access Earth System Science Data Discussions

Appendix A

The tables in Appendix A describe the catchment characteristics in LamaH-Ice. With some adaptations, the tables are based on the corresponding tables in LamaH-CE (Klingler et al., 2021).

Table A1: Gauge referred attributes (Adapted from Klingler et al., 2021**)**

Attribute	Description	Unit	References
ID	Gauge ID number in LamaH-Ice (assigned alphabetically based on river name)	–	–
V_no	A numbering system of Icelandic streamflow gauges (used by the IMO). The number indicates the exact version/location of the sensor/gauge. In some cases, data from multiple sensors is combined to create longer timeseries. Where not available, the NPC WISKI database station ID number is used.	–	(IMO, 2022; NPC, 2021)
VHM_no	A numbering system of Icelandic streamflow gauges (used by the IMO), indicates the site being measured.	–	see above
name	Name of the site	–	see above
river	Name of the river	–	see above
elevation	Elevation of the gauge	m a.s.l.	(National Land Survey of Iceland, 2020)
lon	Gauge longitude (EPSG 3057)	m	(IMO, 2022; NPC, 2021)
lat	Gauge latitude (EPSG 3057)	m	see above
obsbeg_day	The year in which daily runoff time series start	year	see above
obsbeg_hr	The year in which continuous hourly runoff time series start	year	see above
obsend_day	The year in which daily runoff time series ends	year	
obsend_hr	The year in which continuous hourly runoff data time series end	year	see above
gaps_hourly	Fraction of gaps in the raw hourly runoff time series	‰	see above
country	ISO 3166 alpha-3 code for country (ISL in all cases)	–	–
degimpact	Degree of gauge impact; for classes see Table A12	–	–
typimpact	Type of gauge impact, categorized into 6 classes, see Table A11	–	–
HIERARCHY	Gauge hierarchy[a]	–	–
NEXTUPID	ID of the next upstream gauges (can be one or more); 0 indicates no upstream gauges[a]	–	–
NEXTDOWNID	ID of the next downstream gauge (can only be one); 0 indicates no downstream gauge[a]	–	–
qobs	The measured streamflow value	m^3/s	(IMO, 2022; NPC, 2021)
qc_flag	The assigned quality code of the streamflow measurements at each time step. Explained in detail in the Supplement (S2.1).	–	see above

[a] Only for basin delineations B and C.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

600 **Table A2: Meteorological variables from ERA5-Land dataset** (Adapted from Klingler et al., 2021)

Variable hourly	Daily aggregation	Description	Unit	References
DOY	unchanged	Day of year	–	
HOD	omitted	Hour of day	–	
2m_temp	max, mean, min	Air temperature at a height of 2 m above Earth surface	°C	(Muñoz-Sabater et al., 2021)
2m_dp_temp	max, mean, min	Dew point temperature at 2 m height (temperature to which the air, at 2 metres above the surface of the Earth, would have to be cooled for saturation to occur).	°C	see above
10m_wind_u	mean	Eastward component of the 10m wind	ms^{-1}	see above
10m_wind_v	mean	Northward component of the 10m wind	ms^{-1}	see above
fcst_alb	mean	Albedo (fraction of solar radiation reflected by Earth's surface)	–	see above
lai_high_veg	mean	Sum of leaf area (on one leaf side) per unit area of ground for high-vegetation type	m^2m^{-2}	see above
lai_low_veg	mean	Sum of leaf area (on one leaf side) per unit area of ground for low-vegetation type	m^2m^{-2}	see above
swe	mean	Snow water equivalent	mm	see above
surf_net_solar_rad	max, mean	Amount of solar radiation (shortwave radiation) reaching the Earth's surface (direct and diffuse) minus the amount reflected by the Earth's surface (governed by albedo). Upwards fluxes (radiation from the Earth) are positive.	Wm^{-2}	see above
surf_net_therm_rad	max, mean	Net thermal radiation at the Earth's surface. Upwards fluxes (radiation from the Earth) are positive.	Wm^{-2}	see above
surf_press	mean	Surface pressure	Pa	see above
total_et	sum	Total evaporation, including transpiration from vegetation. Positive values indicate evapotranspiration, negative values condensation	mm	see above
prec	sum	Total precipitation (liquid and frozen)	mm	see above
volsw_123	mean	Volume of water in soil layers 1, 2 and 3 (0 to 100 cm depth)	m^3m^{-3}	see above
volsw_4	mean	Volume of water in soil layer 4 (100 to 289 cm depth)	m^3m^{-3}	see above
prec_rav	sum	Total amount of precipitation (liquid and frozen) from the RAV-II reanalysis	mm	(Rögnvaldsson, 2020)
prec_carra	sum	Total amount of precipitation (liquid and frozen) from the CARRA reanalysis	mm	(Schyberg et al., 2020)

Table A3: Topographical attributes (Adapted from Klingler et al., 2021)

Attribute	Description	Unit	Data source

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

area_calc	Catchment area as delineated by the Pysheds Python package (see Section 3 and the Supplement)	km^2	(National Land Survey of Iceland, 2020)
elev_mean	Mean catchment elevation	m a.s.l.	see above
elev_med	Median catchment elevation	m a.s.l.	see above
elev_std	Standard deviation of elevation within catchment	m a.s.l.	see above
slope_mean	Mean catchment slope (Horn, 1981)	$m\ km^{-1}$	see above
mvert_dist	The length of the longitudinal axis of a catchment (horizontal distance from the point furthest away from the gauge to the gauge)	km	see above
mvert_ang	The angle between the north direction and the longitudinal axis of the catchment	degree	see above
asp_mean	The mean aspect of the catchment. The aspect values range from 0 to 360 degrees, with 0 degrees representing north-facing slopes, 90 degrees representing east-facing slopes, 180 degrees representing south-facing slopes, and 270 degrees representing west-facing slopes.	degree	see above
elon_ratio	The length elongation ratio (R_e) of a catchment (Schumm, 1956). Defined as the ratio of the diameter of a circle that has the same area as the catchment, to the catchment length (L, mvert_dist), indicating the roundness of a catchment. Calculated using the equation $R_e = \frac{1}{L}\sqrt{\frac{4*A}{pi}} = \frac{D}{L}$	–	see above
strm_dens	Stream density, i.e. the length of all channels within a catchment divided by its area.	$km\ km^{-2}$	(EU-Hydro - River Network Database, 2019)

Table A4: Climatic indices[a] (Klingler et al., 2021; Addor et al., 2017)

Attribute	**Description**	**Unit**	**Data source**
p_mean	Mean daily precipitation	mmd^{-1}	(Muñoz-Sabater et al., 2021)
eta_mean	Mean daily total evapotranspiration	mmd^{-1}	see above
aridity	Aridity, calculated as the ratio of mean daily total evapotranspiration (eta_mean) to mean daily precipitation (p_mean)	–	see above
p_season	Seasonality and timing of precipitation estimated using sine curves to represent the annual precipitation cycles. Positive values	–	see above

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Open Access Earth System Science Data Discussions

	indicate that precipitation peaks in summer. Negative values indicate that precipitation peaks in winter. Values close to 0 indicate uniform precipitation throughout the year.		
frac_snow	Fraction of precipitation falling as snow	–	see above
hi_prec_fr	Frequency of high-precipitation days (days when precipitation is ≥ 5 times mean daily precipitation)	dyr^{-1}	see above
hi_prec_du	Average duration of high-precipitation events (Number of consecutive high precipitation days as defined above)	d	see above
hi_prec_ti	The season during which most high-precipitation days (defined above) occur	season[b]	see above
lo_prec_fr	Frequency of dry days (precipitation < $1mmd^{-1}$)	dyr^{-1}	see above
lo_prec_du	Average duration of dry periods (number of consecutive dry days as defined above)	d	see above
lo_prec_ti	The season during which most dry days (defined above) occur	season[b]	see above

[a]The climate indices are calculated over the period 1 October 1989 to 30 September 2009. [b]List of List of abbreviations for seasons: djf – December–January–February, mam – March–April–May, jja –June–July–August, son – September–October–November.

Table A5: Hydrological signatures* (Klingler et al., 2021; Addor et al., 2017)

Attribute	Description	Unit	Data source
q_mean	Mean daily streamflow	mmd^{-1}	(IMO, 2022; NPC, 2021)
runoff_ratio	Runoff ratio, i.e. the ratio of mean daily streamflow to mean daily precipitation	–	(Muñoz-Sabater et al., 2021; NPC, 2021; IMO, 2022)
stream_elas	Streamflow-precipitation elasticity. The relationship between changes in streamflow and changes in precipitation on annual timescales. The higher the value, the more we expect runoff to increase due to increases in precipitation.	–	(IMO, 2022; NPC, 2021)
slope_fdc	Slope of the flow duration curve between the log-transformed 33rd and 66th streamflow percentiles.	–	see above
baseflow_index_ladson	Baseflow index (ratio of mean daily baseflow to mean daily total streamflow). Baseflow separation	–	see above

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

	performed using a digital filter method (Ladson et al., 2013) with α set as 0.925.		
hfd_mean	Mean half flow date, i.e. the date on which the cumulative discharge during the hydrological year reaches half of the annual streamflow.	Days since October 1	see above
Q5	5% flow quantile (low flows)	mmd^{-1}	see above
Q95	95% flow quantile (low flows)	mmd^{-1}	see above
high_q_freq	Frequency of high-flow days (>9 times the median daily streamflow)	dyr^{-1}	see above
high_q_dur	Mean duration of high-flow events (as defined above)	d	see above
low_q_freq	Frequency of low-flow days (<0.2 times the median daily streamflow)	dyr^{-1}	see above
low_q_dur	Mean duration of low-flow events (as defined above)	d	see above
zero_q_freq	Percentage of days with zero discharge	%	see above
year_count	Number of years included in the calculations	years	see above

*The signatures were calculated for the 40-year period between October 1st 1981 and September 30th 2021, using the filtered (high quality) version of the streamflow measurements. Only hydrological years with at least 90% temporal coverage during this period were considered. Gauges with less than three years of 90% coverage were omitted. The total number of gauges included in the calculations is 69. The signatures are also made available for unfiltered streamflow measurements, 105 gauges in total.

615

Table A6: Land cover attributes (Adapted from Klingler et al., 2021)

Attribute	**Description**	**Unit**	**Data source**
lc_dom	Dominant land cover class from the CORINE Land Cover classification (three digit code – CORINE nomenclature is listed in folder F_appendix)	–	(CORINE, 2018)
agr_fra	Fraction of agricultural areas (all CORINE classes that start with "2")	–	see above
bare_fra	Fraction of bare areas (CORINE classes 332, 333)	–	see above
forest_fra	Fraction of forested areas (CORINE classes 311, 312, 313)	–	see above
glac_fra	Fraction of glaciers (CORINE class 335)	–	see above
lake_fra	Fraction of inland water bodies (including artificial lakes with all-year water – CORINE class 512)	–	see above
urban_fra	Fraction of areas occupied by infrastructure and associated land (CORINE classes 111, 112, 121, 122, 123, 124)	–	see above

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

scrub_fra	Fraction of scrub and/or herbaceous vegetation associations (CORINE classes 321, 322, 323, 324)	–	see above
wetl_fra	Fraction of wetlands (all CORINE classes that start with "4")	–	see above

Table A7: Vegetation indices* (Unchanged from Klingler et al., 2021)

Attribute	Description	Unit	Data source
lai_max	Maximum monthly mean of one-sided leaf area index (based on 12-monthly means)	–	MODIS MCD15A3H (Myneni et al., 2015)
lai_diff	Difference between maximum and minimum monthly mean of one-sided leaf area index (based on 12-monthly means)	–	see above
ndvi_max	Maximum monthly mean of NDVI (based on 12-monthly means)	–	MODIS MOD09Q1 (Vermote, 2015)
ndvi_min	Minimum monthly mean of NDVI (based on 12-monthly means)	–	see above
gvf_max	Maximum monthly mean of the green vegetation fraction (based on 12-monthly means)	–	MODIS MOD09Q1 (Vermote, 2015), MODIS MCD12Q1 (Friedl and Sulla-Menashe, 2019)
gvf_diff	Difference between the maximum and minimum monthly mean of the green vegetation fraction (based on 12-monthly means)	–	see above

*Calculated by taking the mean of all raster cells whose centroids are located inside the watershed polygon.

Table A8: Soil characteristics (Unchanged from Klingler et al., 2021)

Attribute	Description	Unit	Data source
bedrk_dep	Depth to bedrock; maximum is 50 m[c]	m	(Pelletier et al., 2016)
root_dep	Depth available for roots; maximum is 1.5 m[a,c]	m	European Soil Database Derived data (Hiederer, 2013a, b)
soil_poros	Total soil porosity[a,b,c]	–	see above
soil_tawc	Total available water content (between field capacity and permanent wilting point)[a,b,c]	m	see above
sand_fra	Sand fraction (of soil material < 2 mm)[a,b,c]	–	see above
silt_fra	Silt fraction (of soil material < 2 mm)[a,b,c]	–	see above
clay_fra	Clay fraction (of soil material < 2 mm)[a,b,c]	–	see above
grav_fra	Fraction of gravel (of overall soil)[a,b,c]	–	see above
oc_fra	Fraction of organic material (of overall soil)[a,b,c]	–	see above

[a]Areas marked as lakes or glaciers were excluded from the calculations. [b]Depth weighted average, based on the topsoil (0-30 cm depth) and subsoil (30-150 cm) layers, using the depth available for roots as maximum depth. [c]Calculated by taking the mean of all raster cells whose centroids are located inside the watershed polygon.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Table A9: Geological attributes (Adapted from Klingler et al., 2021)

Attribute	Description	Unit	Data source
gc_dom	Dominant geological class from GLiM layer 1	–	GLiM (Hartmann et al., 2012), layer 1
gc_pa_fra	Fraction of “acid plutonic rocks” (pa)	–	see above
gc_pb_fra	Fraction of “basic plutonic rocks” (pb)	–	see above
gc_va_fra	Fraction of “acid volcanic rocks” (va)	–	see above
gc_vb_fra	Fraction of “basic volcanic rocks” (vb)	–	see above
gc_23_dom	Dominant geological class from GLiM layers 2 and 3	–	GLiM (Hartmann et al., 2012), layers 2 and 3
gc_23_pavr	Fraction of acid plutonic rocks (“pa”) with subordinate volcanics (“vr”)	–	see above
gc_23_pb	Fraction of “basic plutonic rocks” (pb)	–	see above
gc_23_vapy	Fraction of acid volcanic rocks (“va”) with pyroclastics (“py”)	–	see above
g_dom_NI	Dominant geological class from a regional geological map	–	Geological map of Iceland (Icelandic Institute of Natural History, 2014)
g621_fra	Fraction of glaciers	–	see above
g701_fra	Fraction of rivers	–	see above
g743_fra	Fraction of lakes	–	see above
g746_fra	Fraction of reservoirs	–	see above
gbinn_fra	Fraction of basic and intermediate intrusions, gabbro, dolerite and diorite.	–	see above
ggnew_fra	Fraction of basic and intermediate extrusive rocks with intercalated sediments. Upper Pliocene and Lower Pleistocene, 0.8-3.3 m.y.	–	see above
ggold_fra	Fraction of basic and intermediate extrusive rocks with intercalated sediments. Upper Tertiary, older than 3.3 m.y.	–	see above
ghraun_fra	Fraction of basic and intermediate interglacial and supraglacial lavas with intercalated sediments. Upper Pleistocene, younger than 0.8 m.y.	–	see above
gbnew_fra	Basic and intermediate lavas. Postglacial, historic, younger than 1,100 years.	–	see above
gbold_fra	Fraction of basic and intermediate lavas. Postglacial, prehistoric, older than 1,100 years.	–	see above
gmob_fra	Fraction of basic and intermediate hyaloclastite, pillow lava and associated sediments. Upper Pleistocene, younger than 0.8 m.y.	–	see above
gsgos_fra	Fraction of acid extrusives. Tertiary and Pleistocene, older than 11,000 years.	–	see above
gsinn_fra	Fraction of acid intrusions, rhyolite, granophyre and granite.	–	see above
gsn_fra	Fraction of holocene sediment layers.	–	see above

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

gsnew_fra	Fraction of acid lavas. Postglacial, historic, younger than 1,100 years.	–	see above
gsold_fra	Fraction of acid lavas. Postglacial, prehistoric, older than 1,100 years.	–	see above

Table A10: Glaciological attributes

Attribute	Description	Unit	Data source
g_frac	The fraction of the catchment that is covered by glaciers. The reference year for glacier extent is 2019.	–	(Hannesdóttir et al., 2020)
g_area	The actual glaciated area of the catchment (as of 2019)	km^2	see above
g_frac_dyn	Timeseries describing how the glaciated fractions of the catchments have changed through time	–	see above
g_area_dyn	Timeseries describing how the glaciated area of the catchments has changed through time	km^2	see above
g_lon	Longitude of the centroid of the glacier* (EPSG 3057)	m	see above
g_lat	Latitude of the centroid of the glacier* (EPSG 3057)	m	see above
g_mean_el	Mean elevation of the glacier*	m a.s.l.	IslandsDEM (National Land Survey of Iceland, 2020)
g_max_el	Maximum elevation of the glacier*	m a.s.l.	see above
g_min_el	Minimum elevation of the glacier*	m a.s.l.	see above
g_aspect	Aspect of the glacier*. The aspect values range from 0 to 360 degrees, with 0 degrees representing north-facing slopes, 90 degrees representing east-facing slopes, 180 degrees representing south-facing slopes, and 270 degrees representing west-facing slopes.	degrees	see above
g_slope	Average slope of the glacier*	$m\ km^{-1}$	see above
g_slopel20	Slope of the lowermost 20% area of the glacier*	$m\ km^{-1}$	see above

*The term "glacier" is used for the part of the catchment that is glacierized.

Table A11: Attributes for natural or anthropogenic impacts on runoff (adapted from Klingler et al., 2021**)**

Attribute	Description
typimpact	Type of gauge impact, categorized into 6 classes A – Hydropower reservoir upstream B – Subsurface transboundary flow of water (or suspicion thereof), which is affected by leakage from man-made reservoirs C – Withdrawal for irrigation or revegetation

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

	D – Natural lake (with unaffected outlet) E – Natural or artificial diversions upstream or exact bounds of watersheds not certain F – No impact
degimpact	Degree of gauge impact; for classes see Table A12

Table A12: Criteria for the different degrees of gauge impact (adapted from Klingler et al., 2021**)**

degimpact	**Criteria**
u – no influence	There is no obvious type of impact ("typimpact" in Table A11) and the gauge is located above populated areas.
l – low influence	A natural lake (with an unaffected outlet, impact type D) is located upstream of the gauge. A low amount of water is drawn from the river for irrigation or revegetation (impact type B, 2 gauges). A suspicion of (low) subsurface transboundary flow of water into the catchment, affected by leakage from man-made reservoirs (impact type B).
m – moderate influence	Hydropower reservoir located upstream, but water storage in the reservoir is limited and the reservoir does not affect the seasonality of flow (one gauge).
s – strong influence	Gauges with impact type E. Gauge with impact type A (hydropower reservoir located upstream) were assigned a strong influence if the reservoir affects the seasonality of flow (all gauges with impact type A but one). One gauge with impact type B (high effects of subsurface transboundary flow of water).

Author contributions

HBH and BN designed the study. HBH assembled the dataset, performed the analysis of the data and prepared the manuscript. BN reviewed the results and the manuscript and provided consultations and contributions throughout the work.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors gratefully acknowledge the hydrological staff at the Icelandic Meteorological Office (IMO) for collecting and providing streamflow measurements from the IMO database, in particular Njáll Fannar Reynisson, Hilmar Björn Hróðmarsson, Kristjana G. Eyþórsdóttir, Gunnar Sigurðsson and Óðinn Þórarinsson. The authors thank Egill Axelsson for providing helpful insights into the streamflow measurements database at Landsvirkjun. The authors would also like to thank all past and present hydrological surveyors in Iceland for their efforts in measuring streamflow in often difficult conditions. The authors thank the

streamflow data owners for allowing the data to be published, including the Icelandic Meteorological Office, Landsvirkjun, the City of Reykjavík, Veitur Utilities, RARIK, the National Energy Authority of Iceland, Orkusalan, Reykjavík Energy, ON Power, HS Orka and Suðurorka. The authors thank Christoph Klingler for sharing code and QGIS instructions to calculate catchment characteristics related to soils and vegetation. Lastly, the authors thank Helgi Ó. Bragason and Óli G. B. Sveinsson for providing constructive feedback during review of the manuscript.

Financial support

This research has been supported by The Valle Scholarship & Scandinavian Exchange Program at the University of Washington (UW), the UW Herbold Fellowship, the Leifur Eiríksson Foundation Fellowship Program, the American Water Resources Association (Washington Section) and Landsvirkjun. This work was partially funded by the Future Rivers program at the University of Washington as part of an NSF National Research Traineeship award (DGE 1922004) and will be uploaded to NSF's publication database.

References

Abbas, A., Boithias, L., Pachepsky, Y., Kim, K., Ahn Chun, J., Hwa Cho, K., and Ahn Chun jachun, J.: AI4Water v1.0: an open-source python package for modeling hydrological time series using data-driven methods, Geosci. Model Dev, 15, 3021–3039, https://doi.org/10.5194/gmd-15-3021-2022, 2022.

Aðalgeirsdóttir, G., Magnússon, E., Pálsson, F., Thorsteinsson, T., Belart, J. M. C., Jóhannesson, T., Hannesdóttir, H., Sigurðsson, O., Gunnarsson, A., Einarsson, B., Berthier, E., Schmidt, L. S., Haraldsson, H. H., and Björnsson, H.: Glacier Changes in Iceland From ~1890 to 2019, Front Earth Sci (Lausanne), 8, https://doi.org/10.3389/feart.2020.523646, 2020.

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.

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

Alvarez-Garreton, C., Mendoza, P. A., Pablo Boisier, J., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J., and Ayala, A.: 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, 2018.

Arnalds, Ó. and Óskarsson, H.: Íslenskt Jarðvegskort / A soil map for Iceland, Náttúrufræðingurinn, 78, 107–121, 2009.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Árnason, K. and Matthíasson, I.: CORINE land cover classification 2018 - Land cover changes in Iceland 2012-2018. (Title in Icelandic: CORINE-landflokkun 2018 - Landgerðarbreytingar á Íslandi 2012-2018). Technical report, National Land Survey of Iceland, available at https://www.lmi.is/static/files/corine/corine_2018_lokaskyrsla.pdf), 2020.

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303–309, https://doi.org/10.1038/nature04141, 2005.

Bartos, M., Debbout, R., and Huard, D.: Pysheds, Zenodo [code], https://doi.org/10.5281/ZENODO.3822495, 12 May 2020.

Björnsson, H. and Pálsson, F.: Icelandic glaciers, Jökull, 365–386, 2008.

Björnsson, H., Pálsson, F., and Haraldsson, H.: Mass balance of Vatnajökull (1991–2001) and Langjökull (1996–2001), Iceland, Jökull, 51, 75–78, 2002.

Bolibar, J., Rabatel, A., Gouttevin, I., Galiez, C., Condom, T., and Sauquet, E.: Deep learning applied to glacier evolution modelling, Cryosphere, 14, 565–584, https://doi.org/10.5194/tc-14-565-2020, 2020.

Brown, M. E., Racoviteanu, A. E., Tarboton, D. G., Gupta, A. sen, Nigro, J., Policelli, F., Habib, S., Tokay, M., Shrestha, M. S., Bajracharya, S., Hummel, P., Gray, M., Duda, P., Zaitchik, B., Mahat, V., Artan, G., and Tokar, S.: An integrated modeling system for estimating glacier and snow melt driven streamflow from remote sensing and earth system data products in the Himalayas, J Hydrol (Amst), 519, 1859–1869, https://doi.org/10.1016/j.jhydrol.2014.09.050, 2014.

Büttner, G.: CORINE land cover and land cover change products, Remote Sensing and Digital Image Processing, 18, 55–74, https://doi.org/10.1007/978-94-007-7969-3_5/FIGURES/5, 2014.

Chagas, V. B. P., L. B. Chaffe, P., Addor, N., M. Fan, F., S. Fleischmann, A., C. D. Paiva, R., and Siqueira, V. A.: 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, 2020.

Clausen, B. and Biggs, B. J. F.: Flow variables for ecological studies in temperate streams: groupings based on covariance, J Hydrol (Amst), 237, 184–197, https://doi.org/10.1016/S0022-1694(00)00306-1, 2000.

CORINE: CORINE Land Cover 2018: European Environment Agency [data set], Copenhagen, Denmark, available at https://land.copernicus.eu/pan-european/corine-land-cover, 2018.

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 Syst Sci Data, 12, 2459–2483, https://doi.org/10.5194/essd-12-2459-2020, 2020.

Efrat, M.: Caravan extension Israel - Israel dataset for large-sample hydrology [data set], https://zenodo.org/record/7758516, 2023.

Einarsson, M.: Climate of Iceland, in: Climates of the Oceans, edited by: van Loon, H., Elsevier, Amsterdam, 673–697, 1984.

EU-Hydro - River Network Database, V. 1. 2: European Environment Agency under the framework of the Copernicus program [data set], available at https://land.copernicus.eu/imagery-in-situ/eu-hydro/eu-hydro-river-network-database, last access: 30 June 2022, 2019.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Eurostat: The statistical office of the European Union. Population density. Online data code: TPS00003. Available at https://ec.europa.eu/eurostat/databrowser/view/tps00003/default/table?lang=en.: last access: 5 March 2023, 2022.

Fowler, K. J. A., Acharya, S. C., Addor, N., Chou, C., and Peel, M. C.: 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, 2021.

Friedl, M. and Sulla-Menashe, D.: MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006 [data set], https://doi.org/10.5067/MODIS/MCD12Q1.006, 2019.

Gauch, M., Mai, J., and Lin, J.: The proper care and feeding of CAMELS: How limited training data affects streamflow prediction, Environmental Modelling & Software, 135, 104926, https://doi.org/10.1016/J.ENVSOFT.2020.104926, 2021.

Gíslason, S. R.: Weathering in Iceland, Jökull, 58, 387–408, 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.

Google: Google Satellite. Retrieved from http://mt.google.com/vt/lyrs=s, last access: 20 August 2023, 2023.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote Sens Environ, 202, 18–27, https://doi.org/https://doi.org/10.1016/j.rse.2017.06.031, 2017.

Gunnarsson, A., Gardarsson, S. M., and Sveinsson, Ó. G. B.: Icelandic snow cover characteristics derived from a gap-filled MODIS daily snow cover product, Hydrol Earth Syst Sci, 23, 3021–3036, https://doi.org/10.5194/hess-23-3021-2019, 2019.

Gunnarsson, A., Gardarsson, S. M., Pálsson, F., Jóhannesson, T., and Sveinsson, Ó. G. B.: Annual and inter-annual variability and trends of albedo of Icelandic glaciers, Cryosphere, 15, 547–570, https://doi.org/10.5194/tc-15-547-2021, 2021.

Gupta, H. v., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., and Andréassian, V.: 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, 2014.

Hannah, D. M., Demuth, S., van Lanen, H. A. J., Looser, U., Prudhomme, C., Rees, G., Stahl, K., and Tallaksen, L. M.: Large-scale river flow archives: importance, current status and future needs, Hydrol Process, 25, 1191–1200, https://doi.org/10.1002/HYP.7794, 2011.

Hannesdóttir, H., Sigurðsson, O., Þrastarson, R. H., Guðmundsson, S., Belart, J. M. C., Pálsson, F., Magnússon, E., Víkingsson, S., Kaldal, I., and Jóhannesson, T.: A national glacier inventory and variations in glacier extent in Iceland from the Little Ice Age maximum to 2019, Jökull, https://doi.org/10.33799/jokull2020.70.001, 2020.

Hartmann, J., Moosdorf, N., Hartmann, J., and Moosdorf, N.: The new global lithological map database GLiM: A representation of rock properties at the Earth surface, Geochemistry, Geophysics, Geosystems, 13, 12004, https://doi.org/10.1029/2012GC004370, 2012.

Helgason, H. B. and Nijssen, B.: LamaH-Ice: LArge-SaMple Data for Hydrology and Environmental Sciences for Iceland [data set], http://www.hydroshare.org/resource/86117a5f36cc4b7c90a5d54e18161c91, 2023.

Helland, A.: Om Jakelelverne og deres Slamgehalt, Arkiv for Matematik og Naturvidenskab, VII, 213–232, 1882.

Hiederer, R.: Mapping Soil Properties for Europe - Spatial Representation of Soil Database Attributes, Luxembourg, 2013a.

Hiederer, R.: Mapping Soil Typologies - Spatial Decision Support Applied to the European Soil Database, Luxembourg, 2013b.

Hijmans, R.: Boundary, Iceland [Shapefile]. University of California, Berkeley. Museum of Vertebrate Zoology. Retrieved from https://earthworks.stanford.edu/catalog/stanford-xz811fy7881, last access: 9 september 2022, 2015.

Horn, B. K. P.: Hill Shading and the Reflectance Map, Proceedings of the IEEE, 69, 14–47, https://doi.org/10.1109/PROC.1981.11918, 1981.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. v., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB)-a review, Hydrological Sciences Journal, 58, 1198–1255, https://doi.org/10.1080/02626667.2013.803183, 2013.

Icelandic Institute of Natural History: Geological map of Iceland [data set], available at https://gatt.lmi.is/geonetwork/srv/eng/catalog.search#/metadata/%7B005FFDAD-69A1-4385-B16F-FD31B960FE33%7D, last access: 30 June 2022, 2014.

IMO: Icelandic Meteorological Office: Streamflow measurements (Data received 10.01.2022), 2022.

Jónsdóttir, J. F. and Uvo, C. B.: Long-term variability in precipitation and streamflow in Iceland and relations to atmospheric circulation, International Journal of Climatology, 29, 1369–1380, https://doi.org/10.1002/JOC.1781, 2009.

Klingler, C., Schulz, K., and Herrnegger, M.: LamaH-CE: LArge-SaMple DAta for Hydrology and Environmental Sciences for Central Europe, Earth Syst Sci Data, 13, 4529–4565, https://doi.org/10.5194/ESSD-13-4529-2021, 2021.

Koch, J.: Caravan extension Denmark - Danish dataset for large-sample hydrology [data set], https://zenodo.org/record/7962379, 2022.

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., and Nearing, G.: Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets, Hydrol Earth Syst Sci, 23, 5089–5110, https://doi.org/10.5194/hess-23-5089-2019, 2019.

Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., Shalev, G., and Matias, Y.: Caravan - A global community dataset for large-sample hydrology, Scientific Data 2023 10:1, 10, 1–11, https://doi.org/10.1038/s41597-023-01975-w, 2023.

Ladson, A. R., Brown, R., Neal, B., and Nathan, R.: A standard approach to baseflow separation using the Lyne and Hollick filter, Australian Journal of Water Resources, 17, 25–34, https://doi.org/10.7158/W12-028.2013.17.1, 2013.

Mark, B. G. and Seltzer, G. O.: Tropical glacier meltwater contribution to stream discharge: a case study in the Cordillera Blanca, Peru, Journal of Glaciology, 49, 271–285, https://doi.org/https://doi.org/10.3189/172756503781830746, 2003.

Merz, R., Blöschl, G., and Parajka, J.: Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment-MOPEX Regionalization methods in rainfall-runoff modelling using large catchment samples, 307, 2006.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Myneni, R., Knyazikhin, Y., and Park, T.: MCD15A3H MODIS/Terra+Aqua Leaf Area Index/FPAR 4-day L4 Global 500m SIN Grid V006 [data set], NASA EOSDIS Land Processes DAAC, https://doi.org/10.5067/MODIS/MCD15A3H.006, 2015.

National Land Survey of Iceland: IslandsDEM version 1.0 [data set], available at, https://gatt.lmi.is/geonetwork/srv/eng/catalog.search#/metadata/e6712430-a63c-4ae5-9158-c89d16da6361, 2020.

National Research Council: Himalayan Glaciers: Climate Change, Water Resources, and Water Security, National Academies Press, Washington, D.C., https://doi.org/10.17226/13449, 2012.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance, Hydrol Earth Syst Sci, 19, 209–223, https://doi.org/10.5194/hess-19-209-2015, 2015.

NPC: National Power Company of Iceland: Streamflow measurements (Data received 10.01.2021), 2021.

Olden, J. D. and Poff, N. L.: Redundancy and the choice of hydrologic indices for characterizing streamflow regimes, River Res Appl, 19, 101–121, https://doi.org/10.1002/RRA.700, 2003.

Pelletier, J. D., Broxton, P. D., Hazenberg, P., Zeng, X., Troch, P. A., Niu, G., Williams, Z. C., Brunke, M. A., and Gochis, D.: Global 1-km Gridded Thickness of Soil, Regolith, and Sedimentary Deposit Layers, ORNL DAAC, Oak Ridge, Tennessee, USA, https://doi.org/https://doi.org/10.3334/ORNLDAAC/1304, 2016.

Petersen, R. C., Gíslason, G. M., and Vought, L.: Rivers of the Nordic countries, in: River and stream ecosystems of the world, edited by: Cushing, C., Cummins, K., and Minshall, G., University of California Press, Berkeley, California, 295–341, 2006.

Pétursson, P. and Bjarnason, G.: Efnisgæðaritið - Viðauki 3: Jarðmyndanir. (English: The quality of earth material - Appendix 3: Regolith). Technical report, funded by the Icelandic Road and Coastal Administration Research Fund, available at https://www.vegagerdin.is/vefur2.nsf/Files/Vidauki3_jardmyndanir_2022/$file/_Vi%C3%B0auki%203%20Jar%C3%B0myndanir-loka.pdf, Reykjavík, 2022.

Rabatel, A., Letréguilly, A., Dedieu, J.-P., and Eckert, N.: The Cryosphere Changes in glacier equilibrium-line altitude in the western Alps from 1984 to 2010: evaluation by remote sensing and modeling of the morpho-topographic and climate controls, Cryosphere, 7, 1455–1471, https://doi.org/10.5194/tc-7-1455-2013, 2013.

https://doi.org/10.5194/essd-2023-349
Preprint. Discussion started: 9 October 2023

Rabatel, A., Dedieu, J. P., and Vincent, C.: Spatio-temporal changes in glacier-wide mass balance quantified by optical remote sensing on 30 glaciers in the French Alps for the period 1983–2014, Journal of Glaciology, 62, 1153–1166, https://doi.org/10.1017/JOG.2016.113, 2016.

Raynolds, M., Magnússon, B., Metúsalemsson, S., and Magnússon, S. H.: Warming, sheep and volcanoes: Land cover changes in Iceland evident in satellite NDVI trends, Remote Sens (Basel), 7, 9492–9506, https://doi.org/10.3390/rs70809492, 2015.

Rögnvaldsson, Ó.: Observed and simulated weather: Description of dynamical downscaling experiments for the water year 2014-2015 and comparison with observations, Technical report, available at http://rav.betravedur.is/LVC/ObsSim-comparison-TechReport.pdf, 2020.

Rohrer, M., Salzmann, N., Stoffel, M., and Kulkarni, A. v.: Missing (in-situ) snow cover data hampers climate change and runoff studies in the Greater Himalayas, Science of the Total Environment, 468–469, https://doi.org/10.1016/J.SCITOTENV.2013.09.056, 2013.

Schumm, S. A.: Evolution of drainage systems and slopes in Badlands at Perth Amboy, New Jersey, GSA Bulletin, 67, 597–646, https://doi.org/https://doi.org/10.1130/0016-7606, 1956.

Schyberg, H., Yang, X., Køltzow, M. A. Ø., Amstrup, B., Bakketun, Å., Bazile, E., Bojarova, J., Box, J. E., Dahlgren, P., Hagelin, S., Homleid, M., Horányi, A., Høyer, J., Johansson, Å., Killie, M. A., Körnich, H., Le Moigne, P., Lindskog, M., Manninen, T., Nielsen Englyst, P., Nielsen, K. P., Olsson, E., Palmason, B., Peralta Aros, C., Randriamampianina, R., Samuelsson, P., Stappers, R., Støylen, E., Thorsteinsson, S., Valkonen, T., and Wang, Z. Q.: Arctic regional reanalysis on single levels from 1991 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.713858f6, last access: 15 February 2023), 2020.

Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., Van Lanen, H. A. J., Sauquet, E., Demuth, S., Fendekova, M., and Jódar, J.: Streamflow trends in Europe: evidence from a dataset of near-natural catchments, Hydrol. Earth Syst. Sci, 14, 2367–2382, https://doi.org/10.5194/hess-14-2367-2010, 2010.

Vermote, E.: MOD09Q1 MODIS/Terra Surface Reflectance 8-Day L3 Global 250m SIN Grid V006 [data set], NASA EOSDIS Land Processes DAAC, https://doi.org/10.5067/MODIS/MOD09Q1.006, 2015.