Supplement of Earth Syst. Sci. Data, 16, 2741–2771, 2024
https://doi.org/10.5194/essd-16-2741-2024-supplement

Open Access Earth System Science Data

Supplement of

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

Hordur Bragi Helgason and Bart Nijssen

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

S1 Catchment delineation

We delineated topographical catchments of the gauges in LamaH-Ice using the Pysheds python package (Bartos et al., 2020), which calculates flow directions from a digital elevation model (DEM) using the D8 routing scheme. We used the digital elevation model IslandsDEM version 1.0 (National Land Survey of Iceland, 2020) as input to Pysheds, along with gauge coordinates. The IslandsDEM data is mostly based on ArcticDEM (Porter et al., 2018), derived using images from the WorldView 1-3 and GeoEye-1 satellites. The satellite images were acquired mostly from 2012 to present, with some images dating back to 2008. In addition, IslandsDEM also includes some lidar and drone data corrections. The IslandsDEM data has a 2x2 m spatial resolution. Validation efforts using lidar data suggested a positional accuracy of better than 2 meters horizontally and better than 0.5 meters vertically, which provides an indication of the data's overall quality and accuracy.

As is common with optical remote sensing products, the ArcticDEM source data may exhibit void areas or artifacts, which can arise from factors such as cloud cover, shadows, or unfrozen water bodies (PGC, 2023). Additionally, the ArcticDEM datasets comprise imagery collected over multiple years and all seasons, and therefore, the data does not necessarily represent snow-free or temporally stable conditions (PGC, 2023). Due to computational reasons, we used a down-sampled version of the IslandsDEM with a 20x20 m resolution to delineate watersheds in LamaH-Ice. Before calculating flow directions, we conditioned the DEM data by filling depressions and resolving flats.

Although precise quantitative estimates of uncertainty for the final delineated watersheds used in this study are not available, it is acknowledged that uncertainties stem from the DEM, the DEM conditioning methods used, and the flow direction scheme used in the delineation. We verified the delineated catchments by comparing them to gauge catchments available from the Icelandic Meteorological Office (IMO) and the National Power Company of Iceland (NPC) and/or hydrologic information from the Digital map database of Iceland (IS 50V, maintained by the National Land Survey of Iceland). In instances where discrepancies emerged, a thorough investigation ensued, involving a detailed analysis of aerial photographs and an examination of publicly available information regarding potential human alterations to the watersheds.

On glaciers, the topographical catchments needed adjustments since glacial drainage area cannot be fully described using only surface elevation maps. The drainage area of meltwater depends on subglacial topography and the water pressure at the glacier bed, with water flowing in the direction of the steepest descent of the hydraulic potential. The drainage areas of Icelandic glaciers have been calculated based on ice thickness and subglacial topography by glaciologists at the University of Iceland (Björnsson, 1988; Björnsson et al., 2000; Pálsson et al., 2015, 2020). We used the outlines of the glacial watersheds to override the outlines calculated by Pysheds. It should be noted that the subglacial water pressure is not constant in time, and subglacial watersheds can vary with time. The subglacial watersheds have not been calculated systematically for Icelandic glaciers as the glaciers have evolved with time, and thus glacial watersheds are represented as static in the LamaH-Ice dataset. The year of reference in most cases is 2008-2010. This introduces limitations to the accuracy of the static delineations.

S2 Runoff data

S2.1 Quality control of runoff data

Since it is generally not feasible to continuously measure streamflow, water level measurements are used to estimate the flow rate at any given time. The relationship between streamflow and water level is known as a rating curve. It is determined by manually measuring the river flow at various water levels. Errors in streamflow estimations can thus arise from uncertainties in the water level measurements as well uncertainties in establishing the rating curve (Tomkins, 2014). When a rating curve has been established, it is still necessary to regularly measure the streamflow manually to monitor whether changes in the river channel profile have altered the relationship between water level and streamflow. Streamflow measurements in Iceland follow an international standard (ISO 18320:2020) that states that if there is more than a 5% difference between the manual measurement and the corresponding outcome from the rating curve, the rating curve should be re-evaluated. The use of newer acoustic doppler current profilers has resulted in lower uncertainty in streamflow measurements compared to measurements conducted with mechanical Doppler current profilers. Consequently, uncertainty in older streamflow series is higher.

Streamflow data in LamaH-Ice does not include quantifiable uncertainty estimates, as these estimates are not systematically calculated by providers of the streamflow data. However, a total of 5 quality codes for the water level measurements, which are based on subjective judgement from the data providers, are included. These codes are used in the quality control process for streamflow measurements in Iceland. These codes are listed and described in Table S1. The first code (40) is assigned to data that is of highest quality, and the second code (80) is data with only minor interruptions, e.g., due to snow or ice. The third class (100) is data that is estimated due to instrumentation failure or missing data. The estimations are performed manually by hydrologists at the IMO and NPC, using nearby weather observations or nearby streamflow gauges. The fourth and fifth quality codes indicate lower quality or unchecked data, respectively.

A pre-filtered version of the streamflow timeseries is provided. This version only includes data of the two highest quality codes (40 and 80), along with a selection of data with the third-best quality code (100). In this version, data with the third quality code (100) is kept if the number of consecutive days is 10 or less. It is thus assumed that the manual gap filling is adequate for shorter durations (10 days or less). For longer spells without trusted observations, gap filling of streamflow measurements is more difficult, and the data is omitted.

S2.2 Runoff gauge identification

In LamaH-Ice, the gauges are sorted alphabetically by the river name, and gauges are numbered accordingly, starting from 1 (“ID“ in Table A1). Most gauges in Iceland have a so-called “VHM” number, indicating the site being measured, as well as a “V” number, indicating the exact location/version of the gauge. The “V” and “VHM” numbers are provided in the attributes table (Table A1). As in LamaH-CE, the attributes table contains a field for official identification numbers of the gauges

("off_no" in Table A1). The official identification number field in LamaH-Ice uses the VHM number, and for gauges that do not have such a number, the NPC hydrological database ID number of these gauges is provided.

S2.3 River diversions

Some rivers in Iceland are affected by diversions or hydropower dams that impact the runoff in downstream gauges. Once a diversion is built, the catchment area of downstream streamflow gauges is altered. In LamaH-Ice, streamflow observations timeseries are only included for fixed catchment areas. In the case of a diversion that changes the catchment area for a given gauge, observations are either cut short, or another series is created with an updated watershed area and an extension is given to the gauge name. Table S2 lists all gauges affected by diversions and shows which time periods are included in each series from these gauges. After a diversion is installed, the gauge (which now measures water outflow from an altered watershed) is assigned a new ID number (the original ID number multiplied by 10). Four "extra" gauges are therefore added to the 107 gauges in LamaH-Ice, resulting in 111 gauges with streamflow data from watersheds that are static over time.

S3 Meteorological, glacier mass-balance and MODIS timeseries

S3.1 Aggregation to watersheds

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. For RAV-II and CARRA data, as well as MODIS and glacier mass-balance data, we used the xESMF python package (Zhuang et al., 2021) to calculate area-weighted watershed averages.

S3.2 Precipitation and Budyko analysis

Figure S1 compares precipitation between ERA5-Land, RAV-II and CARRA reanalyses on LamaH-Ice watersheds.

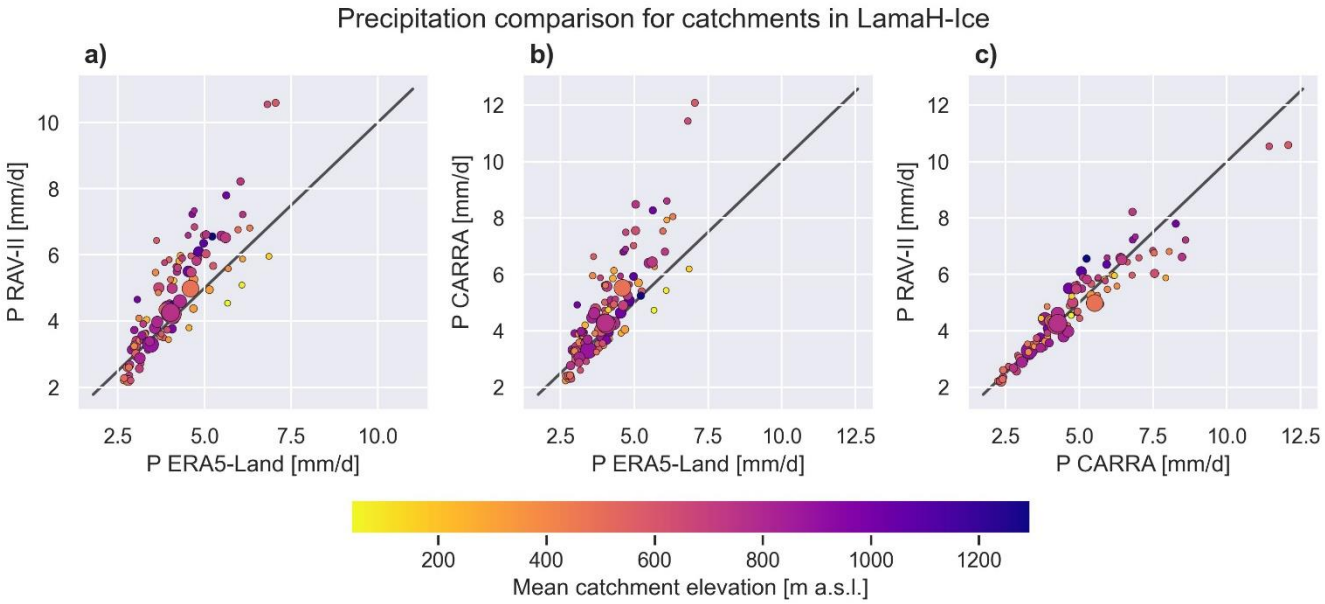

Figure S1: Mean precipitation for the period October 1, 1991 to September 30, 2018 for catchments in LamaH-Ice. Figure a) shows precipitation from RAV-II compared to ERA5-Land. Figure b) shows precipitation from CARRA compared to ERA5-Land. Figure c) shows precipitation from RAV-II compared to CARRA.

Figures S1a and S1b show that precipitation is higher in RAV-II and CARRA than in ERA5-Land. The precipitation in RAV-II and CARRA is of similar magnitude (Figure S1c).

In the Budyko curve analysis in the manuscript, only catchments with a high temporal coverage of streamflow measurements are included. Further, catchments that experience strong anthropogenic or natural impacts are excluded, resulting in a total number of 54 gauges. This was done to maintain consistency with the other subfigures in the water balance analysis. Since the Budyko curve analysis is not dependent on streamflow observations, in Figure S2 we show this analysis for all catchments in LamaH-Ice, 107 in total.

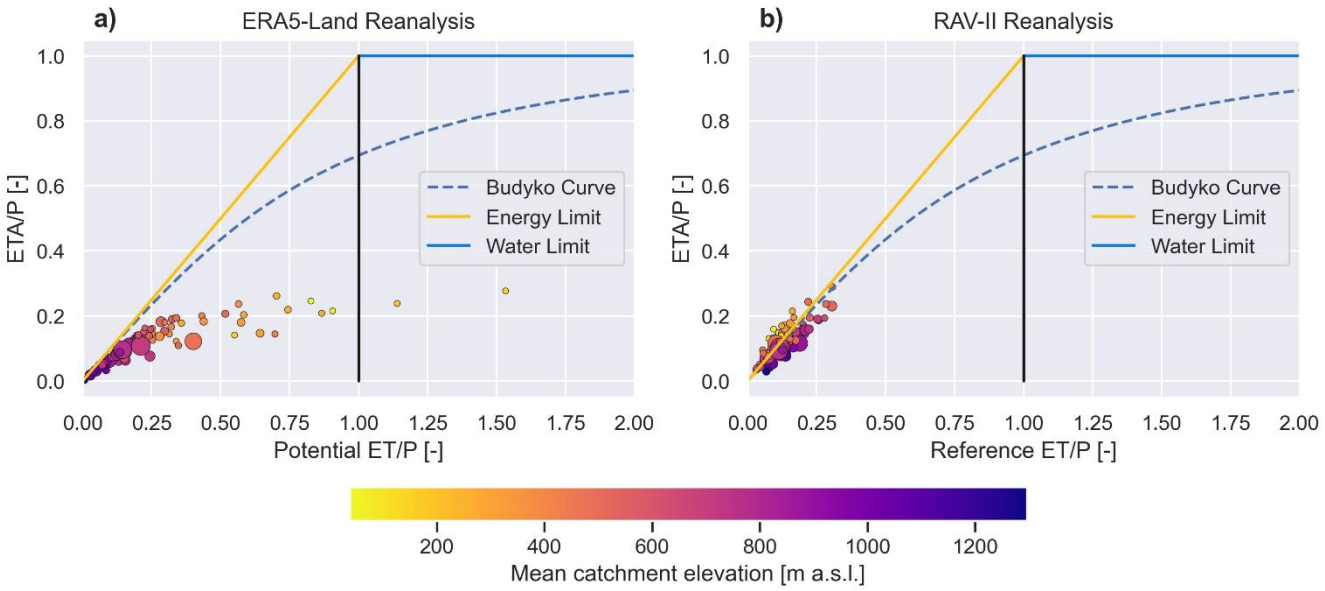

Figure S2: A Budyko curve analysis for all catchments in LamaH-Ice using a) ERA5-Land and b) RAV-II atmospheric data.

Figure S2 shows that the general pattern is the same as shown in the Budyko analysis in the manuscript for the subset of gauges. The reference ET calculated from RAV-II is lower than the PET from ERA5-Land. Evapotranspiration is strongly energy limited in the catchments in LamaH-Ice.

S3.3 Reference ET from RAV-II

We calculated reference ET using meteorological timeseries from the RAV-II reanalysis. We used the Penman-Monteith equation, as recommended by Allen et al. (1998). Input variables from RAV-II are incoming shortwave radiation, incoming longwave radiation, air temperature, dewpoint temperature, surface atmospheric pressure, ground heat flux and wind speed. Net shortwave radiation was calculated using a hypothetical crop albedo of 0.23 as recommended by Allen et al. (1998). The Python code to perform the calculations is available in the LamaH-Ice GitHub repository.

S3.4 Uncertainty in atmospheric reanalysis data

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 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.

S3.5 Glacier mass balance

The locations of the glacier mass-balance measurement stakes are shown in Figure S3. Digital mass balance maps have been derived by interpolating the in-situ measurements and using observed mass-balance gradients (Pálsson et al., 2022, 2024). In LamaH-Ice, the available mass balance maps from Vatnajökull and Langjökull were used to create annual timeseries of mass-balance changes (winter accumulation, summer melt, and annual net mass balance) within each catchment draining the two ice caps, 32 catchments in total.

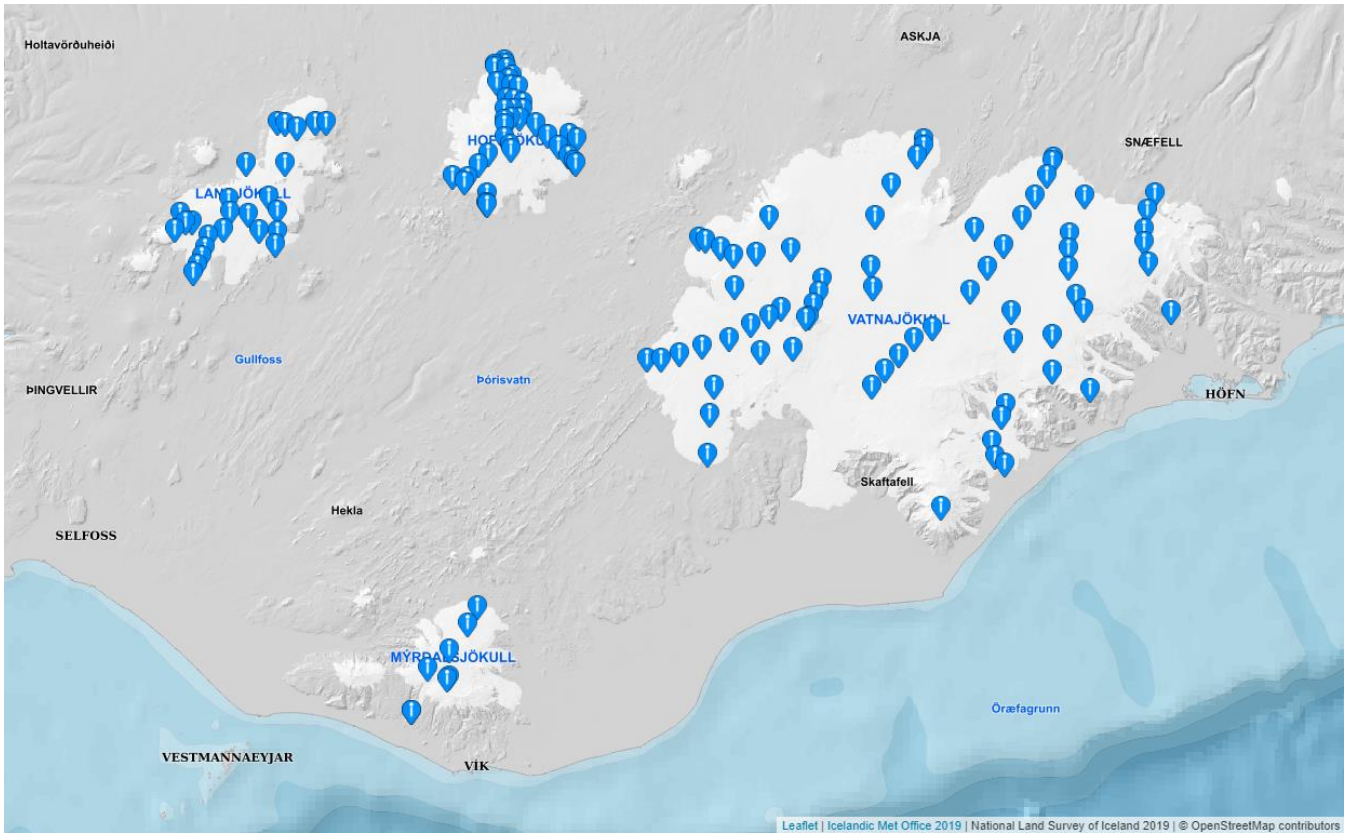

Figure S3: The locations of glacier mass-balance measurement stakes in Iceland. Image source: The Icelandic Glacier Web Portal, 2024.

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, such as wind-driven snow redistribution, could contribute to some level of uncertainty in the derived mass-balance timeseries.

Ice melting that occurs due to geothermal and volcanic activity under glaciers and energy dissipation in the flow of water and ice is known to be a non-negligible component of the mass-balance of glaciers in Iceland (Jóhannesson et al., 2020). The

timeseries provided in LamaH-Ice do not account for these non-surface mass-balance terms and are thus likely to underestimate the total melt.

S3.6 MODIS timeseries

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 it 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. It is well known that MODIS sensors have difficulty distinguishing between cloud and snow. 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.

S4 Catchment attributes

S4.1 Topographic indices

For calculations of topographic indices (other than area, which we derived using Pysheds as discussed in S1), we clipped the IslandsDEM by the catchment polygons and processed the clipped DEM with GDAL and Rasterio (python package). For slope calculation, we used the algorithm by Horn (1981). For calculating stream density, we used the EU-Hydro River Network Database (EU-Hydro - River Network Database, 2019) to represent the stream network, for consistency with the LamaH-CE dataset.

The length elongation of a catchment (Schumm, 1956) indicates the roundness of a catchment, which is an important areal property of a catchment. It is denoted as the ratio between the diameter of a circle that holds the same area as the catchment and the catchment length. The higher the elongation ratio, the more round the catchment’s shape is. We also calculated the length of the longitudinal axis of a catchment. By also including the angle formed by the northward direction and the longitudinal axis of the catchment, it becomes possible to infer the direction of precipitation (with respect to the catchment's longitudinal axis) and how precipitation systems move over catchments. We also calculated the stream density, i.e., length of all channels within a catchment divided by its area. Uncertainties in the DEM or watershed delineation method (described in section S1) can introduce potential inaccuracies in the calculations of these topographic indices.

S4.2 Climate and streamflow indices

We calculated climate indices using the code from the original CAMELS dataset (Addor et al., 2017). If the climate indices are required for different periods, the R code provided in both the supplemental files and the GitHub repository can be re-run using the raw hydro-meteorological timeseries. To calculate streamflow indices, we used a Python implementation of the R code from the original CAMELS dataset (Molin, 2021).

S4.3 Land cover characteristics

We calculated land cover characteristics based on the pan-European Coordination of Information on the Environment (CORINE) Land Cover dataset (Árnason and Matthíasson, 2020). CORINE is an inventory of land cover within Europe split into 44 categories, under 5 main classes: Artificial surfaces, agricultural areas, forested and semi natural areas (which includes glaciers), wetlands and water bodies. In total, 32 of the 44 land cover categories are found in Iceland. The inventory is produced by national agencies under supervision and quality control from the European Environment Agency. It has an update cycle of 6 years, with the most recent update in 2018.

We recognize that the CORINE dataset is subject to inherent uncertainties and limitations due to factors such as inaccuracies in the source data, limitations in remote sensing technologies, and potential misclassifications. The target thematic accuracy of CORINE is 85%. A comprehensive validation study for the 2000 version of CORINE using an independent data source showed that the target accuracy was fulfilled (Buttner and Maucha, 2006) although Iceland was not a part of the investigated area. Distinguishing between classes of similar land cover type can pose a challenge, especially for vegetation classes, many of which exhibit spectral resemblances. The geometric accuracy in CORINE has been better than 100m since 2000 (Árnason and Matthíasson, 2020). The temporal resolution of the dynamic version of the land cover characteristics in LamaH-Ice, which interpolates between the 2000, 2006, 2012, and 2018 CORINE updates, may not fully represent short-term land cover dynamics or events.

S4.4 Vegetation indices

In processing the vegetation catchment characteristics, we followed the manual from the LamaH-CE dataset (Klingler et al., 2021), supplied by Christoph Klingler. These instructions include instructions how to process the remote sensing data in Google Earth Engine as well as post-processing steps in QGIS. The instructions are available in folder F in the dataset.

Leaf area index (LAI) is described as the aggregate leaf area (on one leaf side) per unit area of ground (Watson, 1947). LAI-based indices in the LamaH datasets are calculated from the MODIS MCD15A3H LAI product (version 6.1), with a 4-day temporal resolution and 500m spatial resolution (Myneni et al., 2015). The monthly mean LAI values are calculated based on a period of 19 years between 2002-08-01 to 2021-07-31.

The NDVI (Kriegler et al., 1969) was one of the earliest multispectral remote sensing post-processing products to be introduced and is currently the most popular index for assessing vegetation at the earth's surface (Huang et al., 2021). It is

used for a variety of purposes, including vegetation monitoring, landcover type classification and environmental modeling (Li et al., 2021). The index formulation is based on how green vegetation distinctly reflects near-infrared wavelengths of incoming solar radiation, whilst absorbing a large part of light at the red end of the visible spectrum (Myneni et al., 2019). NDVI takes a value in the range of -1 to 1, with larger positive values indicating increasing vegetation and values approaching zero and below zero suggesting other surface types, such as water bodies, snow, ice, clouds, rock and soil (Saravanan et al., 2019). We calculated maximum and minimum monthly NDVI values for each catchment based on the MODIS MOD09Q1 dataset version 6.1 (Vermote, 2015). The spatial resolution of the dataset is 250m and temporal resolution is 8 days. We used a period of 22 years, from 2000-04-01 to 2022-03-31. We include the full range of NDVI values as obtained from the MODIS observations and leave the user to decide the minimum acceptable NDVI value to be used. A suggestion is 0.1 or 0.2, as values lower than that are usually considered non-vegetated areas (Gandhi et al., 2015).

The GVF index is the fraction of green vegetation on soils and is widely used to provide information on vegetation in atmospheric and hydrologic models (Broxton et al., 2014; Jiang et al., 2010). The GVF index can be calculated from the MODIS NDVI values, using Eq. (1):

$$GVF = \frac{NDVI_{max/min} - NDVI_s}{NDVI_{c,v} - NDVI_s}, \quad (1)$$

Where $NDVI_s$ is the yearly peak NDVI value of bare soil (set as 0.09) and $NDVI_{c,v}$ is the annual peak NDVI of the green surface (Broxton et al., 2014). The annual peak value is derived from the IGBP land category (depicted in table 1 in Broxton et al. (2014)), and the IGBP land categories were acquired from the MODIS MCD12Q1 dataset (Friedl and Sulla-Menashe, 2019). We calculated the maximum monthly mean GVF index for each catchment, as well as the difference between the maximum and minimum values.

The vegetation indices in LamaH-Ice are derived from monthly mean or monthly maximum/minimum values of remote sensing data. While this approach provides valuable insights into the overall vegetation characteristics, it may either not capture or over-emphasize short-term, rapid fluctuations or specific events. The accuracy and reliability of the vegetation indices are contingent on the quality of the source data, i.e. the MODIS products. Limitations in remote sensing technology, cloud cover, and atmospheric conditions may introduce some degree of uncertainty.

S4.5 Surface deposits and soil characteristics

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). Andisols generally have a high soil fertility (Castillo et al., 2022), water holding capacity, total porosity, and hydraulic conductivity (Fontes et al., 2004). Globally, Andisols are the least extensive soil type out of the 12 soil orders defined by the US Department of Agriculture (USDA), occupying approximately 0.7% of the earth's ice-free surface (USDA, 1999).

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). We used the European Soil Database Derived data (ESDD; Hiederer, 2013a, b)

to calculate the other 8 attributes. The ESDD has a spatial resolution of 1 km and is based on the European Soil Database (Panagos, 2006), the Harmonized World Soil Database (Nachtergaele et al., 2008) and the Soil-Terrain Database (van Engelen and Dijkshoorn, 2013). The ESDD provides the soil depth available for roots and then provides 8 soil properties for two soil layers, a topsoil layer (0-30 cm depth) and a subsoil layer (30-150 cm depth). To calculate soil attributes based on the ESDD, we combined the two soil layers by weighting the values from each layer based their depth, with the available root depth as the maximum depth. For the attribute describing total available water content, we summed the values from the two layers.

The data sources described here have very coarse spatial resolution compared to many of the other input data used to calculate catchment characteristics in LamaH-Ice. Uncertainties in these data sources are considerable and the catchment characteristics should be interpreted very carefully. The GGT data set was calibrated and validated using independent soil thickness measurements in the U.S. and Europe, and observations from groundwater wells in the U.S. The calibration and validation mainly used data from midlatitude regions. It is thus acknowledged that the data has limited accuracy in high-latitude regions. Further, the treatment of valley bottoms in the data set (assumed to be V-shaped, rather than U-shaped) does not work well for glacial valleys (Pelletier et al., 2016a) such as commonly found in Iceland. The ESD data set has a maximum soil depth of 1.5 and relies heavily on extrapolated soil profile observations and expert estimations. The accuracy of the data can thus be especially low in situations involving deep soils, heterogeneous soil conditions or substantial distances between soil profiles.

In LamaH-CE, saturated hydraulic conductivity was calculated from the 3D Soil Hydraulic Database of Europe (3DSHD; Tóth et al., 2017). Information regarding bedrock permeability and porosity was derived from GLHYMPS (Gleeson et al., 2014). These databases do not contain realistic values for Iceland and we did thus not include these parameters in LamaH-Ice.

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.

S4.6 Geological characteristics

The GLiM dataset describes the rock types of the earth's surface. The global map containing a total of 1,235,400 polygons was made from joining 92 existing regional geological maps. The average scale of GLiM is 1:3,750,000. The regional map used in GLiM for Northern Europe (including Iceland) was a geological map from the Geological Survey of Norway (Sigmond, 2002) at the scale 1:4,000,000.

GLiM contains lithological classifications on three levels. The first level describes the most dominant lithological class, and is the only level included in LamaH-CE. The classification is comparable to previously defined global lithological maps. The first level contains 16 classes, of which only four are contained in Iceland; "basic volcanic rocks" (which covers most of the

land area), “acid volcanic rocks”, “basic plutonic rocks” and “acid plutonic rocks”. The dominant geological class for all catchments in LamaH-Ice is “basic volcanic rocks”.

GLiM contains two additional layers that describe more specific rock attributes. Since the first level does not adequately describe the lithology of Iceland, we include information from the additional two layers. While the GLiM map has a classification for both water bodies and ice and glaciers, it was only assigned if the regional map represented water or ice areas. In Iceland’s case, the regional map used in GLiM (Sigmond, 2002) did not represent water or glaciers. To supplement the GLiM map, we included catchment characteristics derived from a geological map of Iceland in a much finer resolution (Icelandic Institute of Natural History, 2014).

S5 Dataset structure

For meteorological timeseries and catchment attributes, 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”). The folder structure is shown in Table S3.

Table S1: Quality codes assigned to water level observations.

Quality code	Class	Description/Criteria
40	Good (of highest quality)	Data is good. Water level is recorded without any interruptions.
80	Fair (second class)	Water level data has minor interruptions, e.g. due to ice.
100	Estimated	Data is estimated due to instrumentation failure, ice interruptions or missing data. Estimations use nearby weather observations and/or nearby streamflow gauges.
120	Suspect	Suspected data. Low quality. Two example reasons: A) The water level recording shows spikes but the main line is correct. B) The sensor experiences

		fluctuations, and there are no manual measurements available to confirm the accuracy of the data.
200	Unchecked	Unchecked data.
250	Missing	Missing data.

Table S2: Rivers affected by diversions.

River	Gauge ID	Period	Area [km^2]	Description
Tungufljót, Faxi, V68	89	1951-09-01 to 1986-09-30	829	Before 1986, the Tungufljót river ran from the lake Sandvatn. To prevent wind erosion from mud-flats on the banks of the glacier-fed Sandvatn, the water level in the lake was raised in 1986, and the lake outflow was diverted to Sandá. Since 1986, the lake outlow has occasionally run back to Tungufljót, for example in large flood events. After a flood in 2006, the flow to Tungufljót increased (Jónsson et al., 2010). The type of impact for gauge IDs 89 and 890 is set as "E" and the degree of impact is set as "s", because of uncertainty of how much water flows from Sandvatn to Tungufljót and to Sandá.
	890	1986-10-01 to 2019-09-30	201	
Þjórsá, Norðlingaalda, V100	99	1970-01-01 to 1980-09-30	2044	The Kvíslaveita Diversion (for hydropower) was built in 5 stages between 1980 and 1997. Water from the upper reaches of Þjórsá was diverted into Lake Þórisvatn. During the first period of measurements (gauge ID 99), the river's watershed is natural. No measurements are provided while the
	990	1985-10-01 to 1996-09-30	1620	
	9900	1997-10-01 to 2021-09-30	839	

				first 4 stages of the diversion were being built (1980 to 1985). The watershed did not change during 1985-1996. The 5th stage of the Kvíslaveita diversion was built in 1996-1997. After 1997, this water is occasionally allowed to flow back into Þjórsá. For this reason, after 1997 (for gauge ID 9900), the type of impact is set as "E" and the degree of impact is set as "s".
Þjórsá, Dynkur, V112	101 1010	1988-10-10 to 1996-09-30 1997-10-01 to 2021-09-30	2238 1457	See above. This gauge is located roughly 20 km below the V100 gauge.

Table S3: Dataset structure.

Base folder	Subfolder	Subfolder	Data
A_basins_ total_upstrm **and** B_basins_intermediate_all **and** C_basins_intermediate_lowimp	1_attributes		Catchment attributes and water balance table
	2_timeseries	Annual	Corine land cover and glacier extent and mass balance time series
		Daily	Meteorological time series and modis fractional snow cover and glacier albedo time series
		Hourly	Meteorological time series
	3_shapefiles		Basins and glacier shapefiles and GeoPackages
D_gauges	1_attributes		Gauge attributes and hydrological indices
	2_timeseries	Daily	Streamflow measurements
	3_shapefiles		Gauge shape file and

			GeoPackage file
E_stream_network			A shape file (and GeoPackage) of the EU-Hydro network and information file
F_appendix	1_codes		Code used to create the LamaH-Ice dataset (also available on Github)
	2_shapefiles		The ERA5-Land grid on shape file and geopackage format
	3_CORINE_ landcover		Source information about the CORINE land cover data
	4_QGIS_and_ GEE_processing _instructions		Instructions to re-create vegetation and soil indices in Google Earth Engine and QGIS
G_Information			Information about the folder structure of the dataset, the references of the source data and the license of the data

References

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

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper no. 56, FAO, Rome, 1998.

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

Árnason, K. and Matthíasson, I.: CORINE-landflokkun 2018 - Landgerðarbreytingar á Íslandi 2012-2018 (English: CORINE land cover classification 2018 - Land cover changes in Iceland 2012-2018), National Land Survey of Iceland, Reykjavík, Technical report, 2020.

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

Björnsson, H.: Hydrology of Ice Caps in Volcanic Regions, Reykjavik Societas Scientarium Islandica, University of Iceland, Polar Record, 26, 132–132, https://doi.org/https://doi.org/10.1017/S0032247400011293, 1988.

Björnsson, H., Pálsson, F., and Guðmundsson, M. T.: Surface and bedrock topography of the Mýrdalsjökull ice cap, Iceland: The Katla caldera, eruption sites and routes of jökulhlaups, Jökull, 49, 29–46, https://doi.org/10.33799/JOKULL2000.49.029, 2000.

Broxton, P. D., Zeng, X., Scheftic, W., and Troch, P. A.: A MODIS-Based Global 1-km Maximum Green Vegetation Fraction Dataset, J Appl Meteorol Climatol, 53, 1996–2004, https://doi.org/10.1175/JAMC-D-13-0356.1, 2014.

Buttner, G. and Maucha, G.: The thematic accuracy of Corine land cover 2000 Assessment using LUCAS (land use/cover area frame statistical survey). Technical report no. 7/2006, Copenhagen, 2006.

Castillo, X., Materna, J., Jannoura, R., and Joergensen, R. G.: Peanut monoculture-induced decline in fertility of Andosols in Nicaragua, Journal of Plant Nutrition and Soil Science, 185, 677–684, https://doi.org/10.1002/JPLN.202200112, 2022.

van Engelen, V. W. P. and Dijkshoorn, J. A.: Global and National Soils and Terrain Digital Databases (SOTER) Procedures Manual Version 2.0, Wagenigen, Netherlands, 2013.

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.

Fontes, J. C., Gonçalves, M. C., and Pereira, L. S.: Andosols of Terceira, Azores: Measurement and significance of soil hydraulic properties, Catena (Amst), 56, 145–154, https://doi.org/10.1016/J.CATENA.2003.10.008, 2004.

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.

Gandhi, G. M., Parthiban, S., Thummalu, N., and Christy, A.: Ndvi: Vegetation Change Detection Using Remote Sensing and Gis – A Case Study of Vellore District, Procedia Comput Sci, 57, 1199–1210, https://doi.org/10.1016/J.PROCS.2015.07.415, 2015.

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.

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.

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.

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.

Huang, S., Tang, L., Hupy, J. P., Wang, Y., and Shao, G.: A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing, J For Res (Harbin), 32, 1–6, https://doi.org/https://doi.org/10.1007/s11676-020-01155-1, 2021.

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.

ISO 18320:2020: Hydrometry - Measurement of liquid flow in open channels - Determination of the stage-discharge relationship, International Organization for Standardization, 2020.

Jiang, L., Kogan, F. N., Guo, W., Dan Tarpley, J., Mitchell, K. E., Ek, M. B., Tian, Y., Zheng, W., Zou, C.-Z., Ramsay, B. H., Jiang, C. :, Kogan, F. N., Guo, W., Tarpley, J. D., Mitchell, K. E., Ek, M. B., Tian, Y., Zheng, W., Zou, C.-Z., and Ramsay, B. H.: Real-time weekly global green vegetation fraction derived from advanced very high resolution radiometer-based NOAA operational global vegetation index (GVI) system, Journal of Geophysical Research: Atmospheres, 115, 11114, https://doi.org/10.1029/2009JD013204, 2010.

Jóhannesson, T., Pálmason, B., Hjartarson, Á., Jarosch, A. H., Magnússon, E., Belart, J. M. C., and Gudmundsson, M. T.: Non-surface mass balance of glaciers in Iceland, Journal of Glaciology, 66, 685–697, https://doi.org/10.1017/JOG.2020.37, 2020.

Jónsson, S. A., Þórarinsdóttir, T., and Guðlaugsson, E.: Sandá (vhm 408, V408) - Rennslislíkan Lykilsmíði með HEC-RAS straumlíkaninu. The Icelandic Meteorological Office, Reykjavík, Technical report, ISSN 1670-8261, Reykjavík, 2010.

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.

Kriegler, F., Malila, W., Nalepka, R., and Richardson, W.: Preprocessing transformations and their effect on multispectral recognition, Remote Sens Environ, VI, 97–132, 1969.

Li, S., Xu, L., Jing, Y., Yin, H., Li, X., and Guan, X.: High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques, International Journal of Applied Earth Observation and Geoinformation, 105, 102640, https://doi.org/10.1016/J.JAG.2021.102640, 2021.

Molin, M. D.: HydroAnalysis, [code], available at, https://github.com/dalmo1991/HydroAnalysis, last access: 17 February 2023, 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.

Myneni, R. B., Hall, F. G., Sellers, P. J., and Marshak, A. L.: The interpretation of spectral vegetation indexes, IEEE Transactions on Geoscience and Remote Sensing, 33, 481–486, https://doi.org/10.1109/TGRS.1995.8746029, 2019.

Nachtergaele, F. O. +, van Velthuizen, H., Verelst, L., Batjes, N. H., Dijkshoorn, J. A., van Engelen, V. W. P., Fischer, G., Jones, A., Montanarella, L., Petri, M., Prieler, S., Teixeira, E., Wilberg, D., and Shi, X.: Harmonized World Soil Database (version 1.0)., Food and Agric Organization of the UN (FAO); International Inst. for Applied Systems Analysis (IIASA); ISRIC - World Soil Information; Inst of Soil Science-Chinese Acad of Sciences (ISS-CAS); EC-Joint Research Centre (JRC), 2008.

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.

Pálsson, F. and Gunnarsson, A.: Afkomu- og hraðamælingar á Langjökli jökulárið 2012–2013, Landsvirkjun, Reykjavík, Technical report, LV-2015-076, 2015.

Pálsson, F., Gunnarsson, A., Jónsson, G., Pálsson, H. S., and Steinþórsson, S.: Vatnajökull: Mass balance, meltwater drainage and surface velocity of the glacial year 2018–19, Landsvirkjun, Reykjavík, Technical report, RH-01-20 – LV-2020-016, 56, 2020.

Pálsson, F., Gunnarsson, A., Steinþórsson, S., Eiríksson, K., and Sigurbjörnsdóttir, Þ. A.: Afkomu- og hraðamælingar á Langjökli jökulárið 2021-2022, Landsvirkjun, Reykjavík, Technical report, LV-2022-053, 2022.

Pálsson, F., Gunnarsson, A., Magnússon, E., Steinþórsson, S., Pálsson, H. S., Björnsson, A., and Helgadóttir, S.: Vatnajökull: Mass balance, meltwater drainage and surface velocity of the glacial year 2022-2023, Landsvirkjun, Reykjavík, Technical report, LV-2024-010, 2024.

Panagos, P.: The European Soil Database, GEO: connexion, 5, 32–33, 2006.

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

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, 2016b.

PGC: The Polar Geospatial Center: PGC DEM Products – ArcticDEM, REMA, and EarthDEM: https://www.pgc.umn.edu/guides/stereo-derived-elevation-models/pgc-dem-products-arcticdem-rema-and-earthdem/, last access: 27 September 2023, 2023.

Porter, C., Morin, P., Howat, I., Noh, M.-J., Bates, B., Peterman, K., Keesey, S., Schlenk, M., Gardiner, J., Tomko, K., Willis, M., Kelleher, C., Cloutier, M., Husby, E., Foga, S., Nakamura, H., Platson, M., Wethington Michael, Jr., Williamson, C., Bauer, G., Enos, J., Arnold, G., Kramer, W., Becker, P., Doshi, A., D'Souza, C., Cummens, P., Laurier, F., and Bojesen, M.: ArcticDEM, Version 3, https://doi.org/10.7910/DVN/OHHUKH, 2018.

Saravanan, S., Jegankumar, R., Selvaraj, A., Jacinth Jennifer, J., and Parthasarathy, K. S. S.: Utility of Landsat Data for Assessing Mangrove Degradation in Muthupet Lagoon, South India, Coastal Zone Management: Global Perspectives, Regional Processes, Local Issues, 471–484, https://doi.org/10.1016/B978-0-12-814350-6.00020-3, 2019.

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)67[597:EODSAS]2.0.CO;2, 1956.

Sigmond, E. M. O.: Geological map of land and sea areas of Northern Europe, Geological Survey of Norway, Map, 2002.

The Icelandic Glacier Web Portal: Available at https://islenskirjoklar.is/#, last access: January 30, 2024, 2024.

Tomkins, K. M.: Uncertainty in streamflow rating curves: Methods, controls and consequences, Hydrol Process, 28, 464–481, https://doi.org/10.1002/HYP.9567, 2014.

Tóth, B., Weynants, M., Pásztor, L., and Hengl, T.: 3D soil hydraulic database of Europe at 250 m resolution, Hydrol Process, 31, 2662–2666, https://doi.org/10.1002/HYP.11203, 2017.

USDA (US Department of Agriculture): Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys, U.S. Department of Agriculture Handbook 436, 2nd ed., Natural Resources Conservation Service, 1999.

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.

Watson, D. J.: Comparative physiological studies in the growth of field crops. I. Variation in net assimilation rate and leaf area between species and varieties, and within and between years, Ann Bot, 11, 41–76, 1947.

Zhuang, J., Dussin, R., Huard, D., Bourgault, P., Banihirwe, A., and Hamman, J.: Pangeo-data/xesmf: v0.5.3. Zenodo [code], https://doi.org/10.5281/zenodo.4681767, 2021.