https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

Thermo-hydrological observatory in a permafrost river valley landscape in Syrdakh, Central Yakutia

Eric Pohl[1], Christophe Grenier[2,✞], Antoine Séjourné[3], Frédéric Bouchard[4], Emmanuel Léger[3], Albane Saintenoy[3], Pavel Konstantinov[5], Amélie Cuynet[2], Catherine Ottlé[2], Christine Hatté[6,2], Aurélie Noret[3], Kensheri Danilov[5], Kirill Bazhin[5], Ivan Khristoforov[5], Daniel Fortier[7], Alexander Fedorov[5], and Emmanuel Mouche[2]

[1]Department of Geosciences, University of Fribourg, 1700 Fribourg, Switzerland
[2]Laboratoire des Sciences du Climat et de l'Environnement (LSCE/IPSL), UMR 8212 CEA CNRS UVSQ, Université Paris-Saclay, 91190 Gif-sur-Yvette, France
[3]Laboratoire Geosciences Paris-Saclay, Université Paris-Saclay, CNRS, GEOPS, 91405 Orsay, France.
[4]Department of Applied Geomatics, Université de Sherbrooke, Sherbrooke, Canada
[5]Melnikov Permafrost Institute, Siberian Branch Russian Academy of Science, Yakutsk, Russia
[6]Institute of Physics, Silesian University of Technology, 44-100 Gliwice, Poland
[7]Department of Geography, Université de Montréal, Montréal, Canada
✞deceased

Correspondence: Eric Pohl (eric.pohl@unifr.ch)

Abstract. Permafrost thaw affects the global carbon cycle and can significantly alter landscape morphology and associated processes of mass and energy transfer. An understudied aspect of affected permafrost landscapes are ubiquitous rivers connecting thermokarst lakes. These ubiquitous features of Arctic landscapes exhibit particularly high variability in water and energy transfer, and thus provide an excellent field laboratory for analyzing how expected changes in meteorological forcing under climate change affect permafrost dynamics and carbon exchange within the land- and limnoscape. This paper presents a database from 2012 through 2022 for one such small stream connecting two thermokarst lakes. First, two main stream cross sections were instrumented with multiple thermistor chains to record temperature evolution from the surface to different soil depths. The cross sections covered different topography and vegetation cover. One was located near the upper, and one in between the two thermokarst lakes. The main focus was set on the cross section midway between the two lakes due to the absence of a thermal imprint from the lake. Air, water, and ground temperatures, as well as river water parameters, and soil properties of the surrounding environment were measured as time series or single tests during annual field campaigns. The data are organized in three main categories: atmosphere, water and ground, and are complemented by a GIS including a digital surface model and an ortho-mosaic photo of the entire river valley to facilitate the search for measurements of interest. The database comes with a complete set of scripts to process any of the data, which are provided in CSV or other easily accessible standard file formats. Ultimately, the data can be used to develop models and validate numerical codes for improving the representation of permafrost processes in land surface and climate models where climate change induces significant changes in heat and mass transfer. All data and processing scripts are available through an online repository (https://doi.org/10.5281/zenodo.14619854; Pohl et al. (2025)).

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

1 Introduction

Recent increases in air temperature in the Arctic and sub-Arctic regions are well above global average trends (Meredith et al., 2019). Consequently, permafrost in these regions is experiencing significant warming (Biskaborn et al., 2019). Thawing of ice-rich permafrost can trigger landscape changes such as ground subsidence, mass movements, and hydrological changes. In addition to disrupting water resources and threatening the stability of buildings, thawing typically creates thermokarst lakes (French, 2017) that form in the resulting depressions. These lakes and their connecting streams alternate the dominant heat transfer processes and drive a climate feedback cycle including the activation and mobilization of large carbon stocks (Hugelius et al., 2014; Mishra et al., 2021; Miner et al., 2022) with global socio-economic impacts (Hope and Schaefer, 2016; Schuur et al., 2015).

The ground thermal state of permafrost can show high spatial variability. Sparse observation networks for air, and particularly for soil and deeper ground temperatures (e.g. Boike et al., 2019), as well as a large range of processes involved in the modes of heat and water transfer render past and future predictions of permafrost evolution difficult (Koven et al., 2013). Additionally, a multitude of important small scale processes related to heat and water fluxes in permafrost landscapes (Walvoord and Kurylyk, 2016; Westermann et al., 2016, 2017; Aas et al., 2019; Song et al., 2020; Steinert et al., 2021; Tananaev and Lotsari, 2022) are barely represented in regional to global scale land surface and climate models operating at coarse spatial scales. This is despite the need for such models to produce reliable estimates of ground-atmosphere interactions and represent the interactions and feedback a warming climate has with permafrost landscapes. Improving small scale process representations and how these can be accounted for at larger scale is therefore of high importance (Grenier et al., 2018; Jan et al., 2020).

One particular emerging landscape feature that has received rather limited attention are small rivers and creeks that connect or drain newly developing thermokarst lakes (e.g. Liu et al., 2022; Tananaev and Lotsari, 2022). In general, Arctic rivers are characterized by high seasonal flow variability (e.g. Gautier et al., 2018) inducing a likewise high variability in heat transport at the water-land interface, subsurface saturation and flow, and as a consequence, also regarding solute and particle fluxes. Compared to perennially unfrozen water bodies of lakes and large rivers, small rivers and creeks might experience freeze-through and do not provide heat to maintain a talik, i.e. an unfrozen zone underneath the river or lake bed that can form for large enough water bodies (Kurylyk and Walvoord, 2021; Léger et al., 2023). A warming climate implies the formation of more water bodies, in particular small rivers connecting newly formed thermokarst lakes (e.g. Morgenstern et al., 2011; Tananaev and Lotsari, 2022). Continued observations, not only of the thermal state of permafrost but also of the multiple other types of variables required to understand the changes to permafrost, are therefore of great importance (e.g. Boike et al., 2019).

This paper presents a database of an observatory of a small non-perennial river or stream (namely the Syrdakh River) in the Syrdakh valley in Central Yakutia, Eastern Siberia (Pohl et al., 2025), which we assume to be representative for large parts of the Yedoma Permafrost Domain (Strauss et al., 2021), prone to permafrost thaw and consequent landscape changes. The data aim particularly at serving the validation and calibration of numerical modelling code for heat and water transfer processes in the ground and at the water-ground interface. Therefore, monitoring of ground temperatures and the determination of physical ground parameters for relevant processes were the focus of the instrumentation and the annually performed tests. The two main

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

observation sites in the form of cross sections are located between two large thermokarst lakes: one site about 50 m downstream of the "upper lake" with assumed thermal influence of the lake water body, and one site a few hundred meters away (further downstream) without such thermal influence.

2 Site description

The Syrdakh River is located about 100 km northeast of Yakutsk in Central Yakutia (Eastern Siberia), Russia (Fig. 1), within the continuous permafrost zone (300 m in thickness, Soloviev (1973)). The climate is sub-Arctic and strongly continental with long, cold and dry winters (January mean temperature ≈ -40 °C) as well as warm summers (July mean temperature ≈ +20 °C). This results in a notably strong seasonal variability, with annual extremes of nearly 100 °C between minimum (-60 °C) and maximum (+35 °C) temperatures (Gorokhov and Fedorov, 2018). Total precipitation is low (< 250 mm) and mostly confined to the summer, resulting in a relatively low winter snow cover (< 30 cm) (Gorokhov and Fedorov, 2018). The evapotranspiration rate is high, up to ten times the precipitation in summer during dry years (about 150-200 mm on average). The region is characterized by taiga vegetation dominated by larch, pine and birch forests with dense shrub cover in forested areas and steppe grasses or bog plant communities (salt-tolerant species) otherwise. Active layer thickness varies between ≈ 1 m below forested areas to > 2 m in exposed grassland areas (Fedorov and Konstantinov, 2008). The region is underlain by the Yedoma ice complex (30-50 m), defined as ice-rich permafrost of late-Pleistocene age (about 70–80% ice per volume), composed of frozen silts and silty sands penetrated by several meter-thick syngenetic ice wedges (Ivanov, 1984). Central Yakutia underwent a strong warming trend of ≈+0.05 °C yr^{-1} for the period 1966-2016 (Gorokhov and Fedorov, 2018; Fedorov et al., 2014). Numerous thermokarst lakes of different origin are observed in the region: early Holocene thermokarst led to the formation of thermokarst basins called alas (Soloviev, 1973), while during the early 1990s an intense thermokarst formation period created many small thermokarst lakes (Fedorov and Konstantinov, 2008; Iijima et al., 2010).

Zones of unfrozen ground, called taliks, exist underneath major rivers and lakes deeper than the winter ice-cover thickness (to a maximum of 1.5 m to 2 m) (Lütjen et al., 2024). The presence of numerous small and young (< 50 yr), fast-developing thermokarst lakes (> 3-4 m deep) and retrogressive thaw slumps along lake shores attest to the recent warming in the area (Morgenstern et al., 2011; Desyatkin et al., 2015; Séjourné et al., 2015). These younger thermokarst lakes are observable alongside older (early Holocene) and shallower (≈ 1 m deep) alas lakes.

The small non-perennial Syrdakh River connects two thermokarst lakes from east to west over a distance of around 3 km. In 2012 the first thermistor chains were installed at various depths, following common ground temperature measurement protocols in the region. Over the following years this was extended with soil measurements (e.g. granulometric soil composition, soil infiltration, thermal soil parameters), measurements on water chemistry (e.g. stable water isotopes, conductivity, pH), as well as geophysical measurements (e.g. Ground-Penetrating Radar (GPR), Electrical Resistivity Tomography (ERT)) at other cross-sections (Fig. 1). The installation of instruments was conducted at two main sites, one (Site 1) near the upstream thermokarst lake and one (Site 2) roughly half way between Site 1 and the downstream thermokarst lake, where the village of Syrdakh is located. Multiple instrument failures, particularly of instruments installed in the river and riverbed, have led to the discontinu-

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

ation of instrument reinstallation and maintenance at Site 1. In addition to that, the focus of the database revolved around the thermal imprint of rivers. As a consequence, Site 2 is the prioritized site where most measurements were conducted. In 2019, a new row of piezometer tubes were installed at Site 2 (CS9) to modernize the instruments.

3 Database

The database comprises files of different data and information types. All data are provided in non-proprietary file formats to achieve easy accessibility across platforms. All times are provided in UTC; the local time zone of Syrdakh is UTC+9. All data files follow a naming convention outlined in Section 3.3. All variable names, relevant instruments and their characteristics, including accuracies, resolution, and measurement intervals are listed in Table 1.

3.1 GIS database

The database comes with a QuantumGIS (QGIS) (QGIS Development Team, 2024) project file that displays all sampling and measurement locations (Fig. 1). In several cases, multiple instruments and/or measurements at various depth were employed/taken at a single point. In this case, the feature (point, lines, and polygons) name will reflect the common measurement location, e.g. "pit" (soil pit), "CS-*i*" (cross section *i*), or "piezo" (piezometer tube). The location where each measurement was taken is part of the filenames and facilitates identifying the measurement location. Data files that contain data from multiple measurement locations use the exact location names in the column "point" (e.g., isotopes, water conductivity measurements).

3.2 Data location

The compilation of measurement points over multiple years made use of a differential GPS (dGPS), a theodolite, and handheld GPS instruments. This resulted in some inaccuracies regarding the measuring positions. Coordinates were in several cases adjusted manually in the feature vector files (Section 4.4.4 to match the relative positions in the imagery obtained with an Uncrewed Aerial Vehicle (UAV). For this reason, all position information should be taken from a relational geographical location document "auxiliary/gps" that holds the individual UTM zone 52N Easting (x) and Northing (y), as well as latitude and longitude information for each measurement/instrument position. This file is generated from running a Python script (Section 4.4.3) that extracts all these locations from the feature vector files accessible through the QGIS project that is included in the database. Making use of this, the user can adjust the locations for experiments by editing the feature shape files and create their own coordinate lookup table. Observations like river cross section water depths, and velocities are reported as relative positions from one shoreline to the other one, or with respect to the river position and might not match the imagery as the width of the river might have changed between years. In each of the individual data files the location name is mentioned so that users can adjust such a profile according to their application needs with relevant information of the surrounding landscape available in the GIS.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

Table 1. List of variables, sensor or instrument names, type of measurement (time-series <x>ts, tests <x>test, for atmosphere (x=A), water (x=W), and ground (x=G)), and instrument characteristics.

Variable	Sensor/Instrument	Type	Long name	Unit	Measuring interval	Accuracy (±)	Sensor resolution	Range
aclt[a]		Gtest	Active layer thickness / Thaw depth	cm	1 yr			
	PIT					NA[h]		0 cm to 300 cm
	PVC-TUBE					NA[h]		0 cm to 200 cm
	HANDBAR					NA[h]		0 cm to 200 cm or 0 to 300 cm (two versions)
	DRILL					NA[h]		0 cm to 500 cm
	ERT - SibER-64					RMSD[c] < 4%		
	GPR - OKO 150 MHz						35 cm	0 cm to 1600 cm
gwd		Gtest	Ground water depth	cm	1 yr			
	Electric measuring tape					2 cm	1 cm	0 cm to 500 cm
gps		Gtest	GPS coordinates	° lat/° lon	1 yr			
	Leica Viva Uno 10					dGPS (GPS+GNSS): 0.5 m		
	Ural Optics and Mechanical Plant 3T5KP					2 arcseconds (horizontal), 5 arcseconds (vertical)		
tsl<cccc>[g]		Gts/Gtest	Soil temperature time-series or tests	°C	2 hr, 3 hr, 4 hr, 6 hr, 1 yr			
	Proges Plus Thermo Buttons 22L[b]					0.5 °C	0.1 °C (0.5 °C from 2017 onward)	-40 °C to +85 °C
	Onset HOBO TMCx-HD[b]					0.25 °C	0.03 °C (at 20 °C)	-40 °C to +100 °C
ta		Ats	Air temperature time-series	°C	1 hr, 2 hr, 3 hr			
	Proges Plus Thermo Buttons 22L[b]					0.5 °C	0.1 °C	-40 °C to +85 °C
	T&D RTR-502[b]					0.3 °C (-20 °C to +80 °C), 0.5 °C (-40 °C to -20 °C), 1.0 °C (-60 °C to -40 °C)	0.1 °C	-60 °C to +155 °C
tw		Wts/Wtest	Water temperature time-series or tests	°C	1 hr, 2 hr, 3 hr			
	vanEssen Micro-Diver DI601 (D)[b]					0.1 °C	0.01 °C	-20 °C to +80 °C
	WTW multi-parameter meter					0.1 °C		0 °C to +35 °C
	YSI Pro DSS multi-parameter meter					0.2 °C	0.1 °C	-5 °C to +70 °C
wvl		Wtest	Water velocity (horizontal)	$m\,s^{-1}$	1 yr			
	Hydreka BFM 801					0.5 % of measured value		-5 $m\,s^{-1}$ to 5 $m\,s^{-1}$
wl		Wts/Wtest	Water level time-series or tests	cmH2O	1 hr, 2 hr, 3 hr, 1 yr			
	vanEssen Micro-Diver DI601 (D)[b]					1.0 cm	0.2 cm	0 cm to 1000 cm
	Measuring rod					5.0 cm		0 cm to 200 cm
swc<cccc>[f,g]		Gts/Gtest	Volumetric soil water content time-series or tests	$m^3\,m^{-3}$	1 hr			
	Decagon EC-5[b]					3 %	0.1 % (mineral soils), 0.25% (high organic soils)	0 $m^3\,m^{-3}$ to 1 $m^3\,m^{-3}$
grnl		Gtest	Granulometry (gravimetric class abundance)	Vol. %	1 yr			
	Malvern Mastersizer 2000 (high res.)					clay 1.1 %(SD=1.3 %); silt 2.8 %(SD=3.1 %); sand 3.5 %(SD=4.0 %)[d]		0.01 μm to 10'000 μm
	Sieves (Kachinsky, 1965 [in Russian])							40 μm to 125'000 μm
thermal[f]		Gtest	Thermal properties incl. volumetric soil water content	various	1 yr			
	KD2 - T1		Temperature (KD2-Pro)	°C			0.001 °C	NA
	KD2 - K		Thermal conductivity	$W\,m^{-1}\,K^{-1}$		10 % (0 $W\,m^{-1}\,K^{-1}$ to 2 $W\,m^{-1}\,K^{-1}$), 0.01 $W\,m^{-1}\,K^{-1}$ (0.02 $W\,m^{-1}\,K^{-1}$ to 0.20 $W\,m^{-1}\,K^{-1}$)		0.02 $W\,m^{-1}\,K^{-1}$ to 2.00 $W\,m^{-1}\,K^{-1}$
	KD2 - C		Volumetric specific heat	$MJ\,m^{-3}\,K^{-1}$		10 % (K > 0.1 $W\,m^{-1}\,K^{-1}$)		0.5 $MJ\,m^{-3}\,K^{-1}$ to 4.0 $MJ\,m^{-3}\,K^{-1}$
	KD2 - D		Thermal diffusivity	$mm^2\,s^{-1}$		10 % (K > 0.1 $W\,m^{-1}\,K^{-1}$)		0.1 $mm^2\,s^{-1}$ to 1.0 $mm^2\,s^{-1}$
	KD2 - R		Thermal resistivity	$K\,cm\,W^{-1}$		10 % (K > 0.1 $W\,m^{-1}\,K^{-1}$)		50 $K\,cm\,W^{-1}$ to 5'000 $K\,cm\,W^{-1}$
	Anritsu - T2		Temperature (Anritsu)	°C		0.5 °C	0.1 °C	-200 °C to 800 °C
	HydroSense (I and II) VWC		Volumetric soil water content (HydroSense)	Vol. %		3 %		dry to saturation[i]
	EC-5 VWC		Volumetric soil water content (EC-5)	Vol. %		3 %		0 Vol. % to 100 Vol. %
	Laboratory VWC		Volumetric soil water content (Laboratory)	Vol. %				0 Vol. % to 100 Vol. %
	Laboratory Porosity		Soil porosity (Laboratory)	Vol. %				0 Vol. % to 100 Vol. %
ksat		Gtest	Saturated hydraulic conductivity	$m\,s^{-1}$	1 yr			
	Infiltration rings							10^{-3} $m\,s^{-1}$ to 10^{-6} $m\,s^{-1}$
cond		Wtest	Water electrical conductivity	$\mu S\,cm^{-1}$	1 yr			
	WTW multi-parameter meter					0.5 %		10 $\mu S\,cm^{-1}$ to 20'000 $\mu S\,cm^{-1}$
	YSI Pro DSS multi-parameter meter					0.5 % of reading or 1 $\mu S\,cm^{-1}$, whichever is greater	1 $\mu S\,cm^{-1}$ to 10 $\mu S\,cm^{-1}$	0 $\mu S\,cm^{-1}$ to 100'000 $\mu S\,cm^{-1}$
DO		Wtest	Dissolved oxygen saturation	%	1 yr			
	YSI Pro DSS multi-parameter meter					1% of reading or 1% air saturation, whichever is greater (0 % to 200 % dissolved oxygen)	0.1 %	0 % to 500 %
DOC		Wtest	Dissolved organic carbon	$mg\,L^{-1}$	1 yr			
	Shimadzu TOC-5000A					0.1 $mg\,L^{-1}$ (above quantification limit of 1 $mg\,L^{-1}$)		1 $mg\,L^{-1}$ to 30'000 $mg\,L^{-1}$
pH		Wtest	pH	/	1 yr			
	YSI Pro DSS multi-parameter meter					0.2 pH units	0.01 pH units	0 pH units to 14 pH units
	WTW multi-parameter meter					0.2 pH units		0 pH units to 14 pH units
age		Gtest	Radiocarbon dating of soil organic carbon	$F^{14}C$	1 yr			
	MICADAS (Synal et al., 2007)					1.3 to 28%according to the carbon mass[j]		
istp		Wtest	Isotope concentration	‰[e]	1 yr			
	LWIA DLT-100 Los Gatos					1.5‰ (2H); 0.2‰ (18O)		

[a] aclt represents thaw depth to frozen layer and differs from actual active layer thickness (see Section 4.1.3). Interpretation using electrical resistivity tomography (ERT) and ground penetrating RADAR (GPR) is subjective as it depends on relative signal strength changes.

[b] Time-series.

[c] RMSD - R mean squared difference between observation and interpreted aclt as presented in Léger et al. (2023).

[d] Estimates from Miller and Schaetzl (2012). SD - standard deviation.

[e] δ notation with respect to Vienna Standard Mean Ocean Water (VSMOW).

[f] Note that one-time *swc* tests/measurements are included in the *thermal* files.

[g] <cccc> indicates the depth of time-series measurements in cm, using a 4 digit number with leading zeros, e.g. "0400" for a measurement depth of 400 cm.

[h] Method accuracies and resolutions are subjective and should be seen as rough estimates and might have varied based on individual operators.

[i] For standard soil types in the range of 0% to 50% but can be higher for high porosity organic soils as e.g. in the soil pits near the river.

[j] Accuracy refers to the entire 14C measurement process, including statistical counting, blank subtraction, and normalization using Oxalic Acid I. It does not reflect the intrinsic accuracy of the atomic mass spectrometer instrument only.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

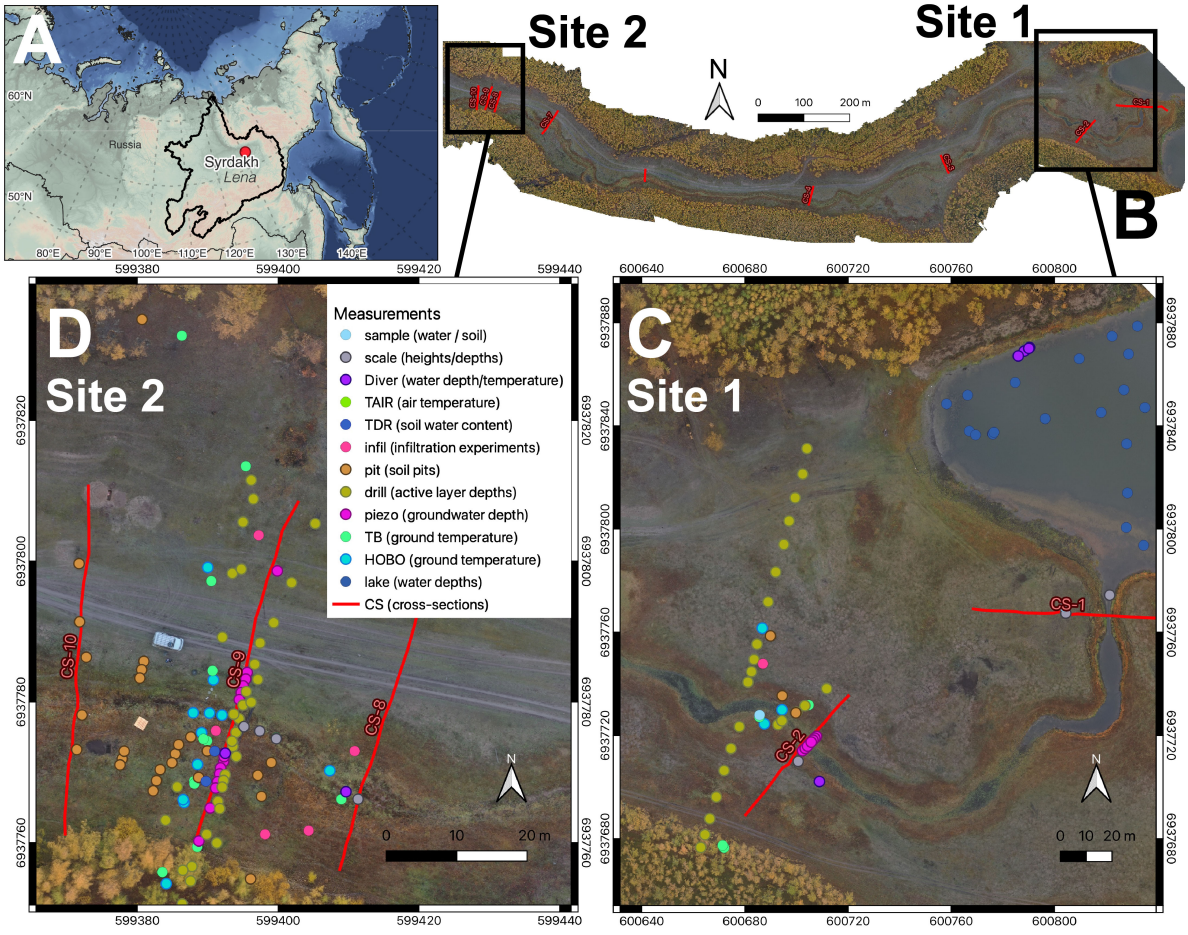

Figure 1. (A) General location of Syrdakh in Siberia within the Lena River catchment (black outline), and (B) UAV orthomosaic aerial view of the Syrdakh River field site with cross sections from the upstream thermokarst lake (CS-1) to the mid-section of the river (CS-10). Zoomed in views show highly instrumented cross section CS-9 or "Site 1"(C) and CS-2 referred to as "Site 2"(D). Coordinates in the zoomed in views are in UTM zone 52N. Data source for (A) is the Earth Topography (ETOPO1) digital elevation model in 1 arc-minute resolution (NOAA National Geophysical Data Center, 2009; Amante and Eakins, 2009).

3.3 Database structure

The database is provided as a folder structure with most files in a simple CSV format in UTF-8 text encoding, imagery is provided as GeoTiff, vector files as GeoPackage, Ground Penetrating RADAR (GPR) data in SEGY format, and electrical resistivity tomography (ERT) data as inversion output from using the BERT software (Günther and Rücker, 2012), and programming and processing code in the form of R and Python scripts. The simple file format provides a way to present sufficiently and

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

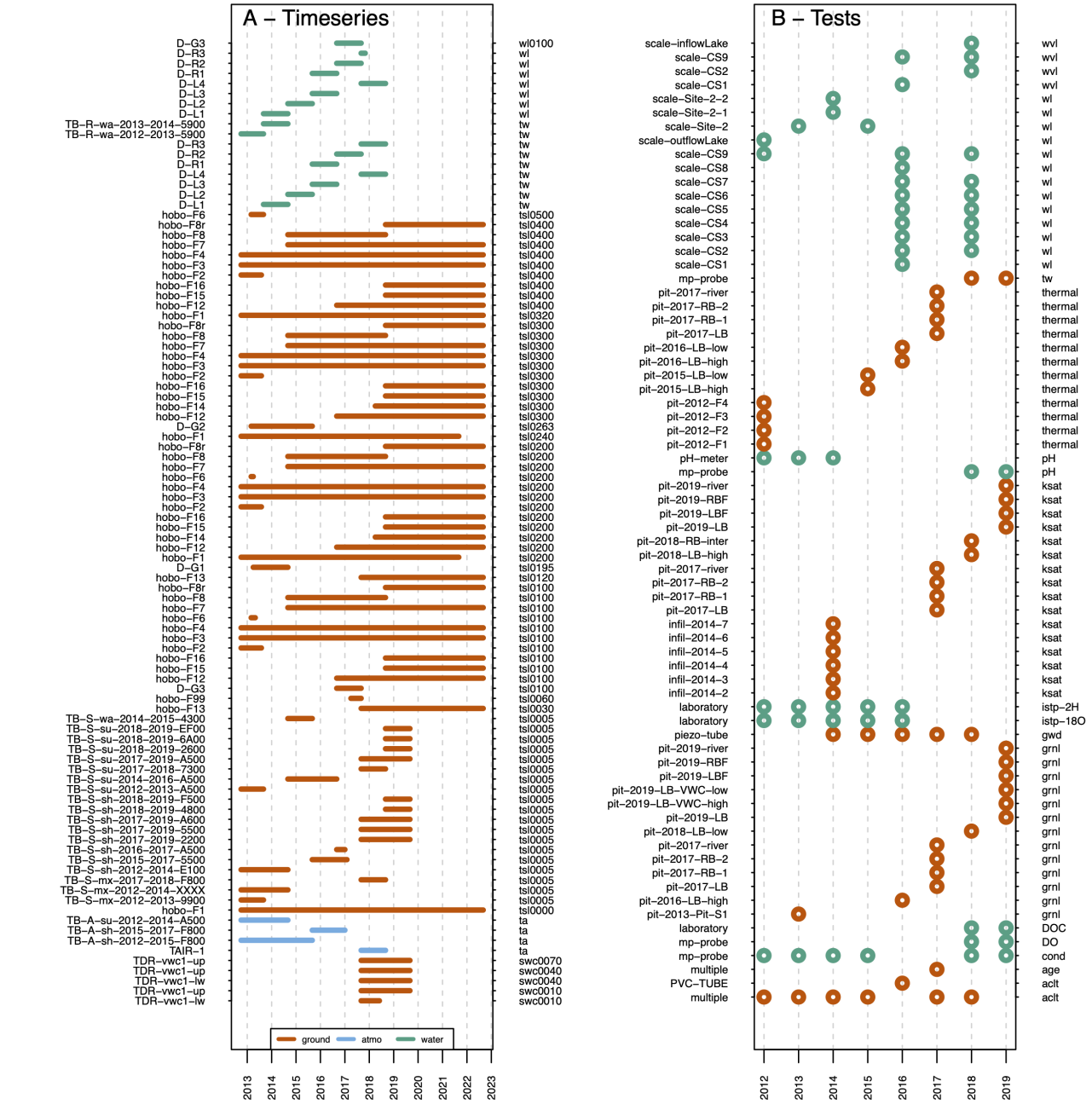

Figure 2. Temporal and variable type coverage of the dataset for (A) time series and (B) individual tests. Left-hand side indicates the location and right-hand side (color-coded) the variable. See Table 1 for complete variable names.

120 without proprietary software the different types of measurements (e.g. time-series vs. snapshot experiments in soil pits). The

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

main subdivision of the database follows the different domains in which measurements were taken according to the following 4 main categories: ground, atmo(sphere), water, and auxiliary.

Filenames are constructed to provide the most important information about the measurements according to equation (1):

<variable_id> _ <table_id> _ <experiment_id> _ <location_id> _ <process_level> _ <time_range> .csv (1)

, where variable_id is the short variable name (see Tab. 1), table_id is the type of measurement (time-series <x>ts, or tests <x>test, for atmosphere (x=A), water (x=W), and ground (x=G)), experiment_id is the site descriptor ("syrdakh"), location_id is the name of a measurement point or static instrument if a single instrument is installed, process_level indicates if data has been post-processed (L2), is provided in raw format (L0), or if instrument output is used as is without post-processing (L1), and time_range provides a time range (time-series), or a year with the addition "-mean" to indicate a single measurement as is the case for most of the tests. The incorporated dataset by Hughes-Allen et al. (2020) is indicated with "sample-HA" as location_id and provides the seasonal information when the sampling took place, as an additional column in the data file. If precise time information was available, these are provided within the meta-data inside the file header.

3.4 Data files

In addition to the information included in the filename (Eq. 1), an extended header of 22 lines includes meta-data, some of which are extracted from Table 1. An example of a header including corresponding data for a time-series as well as multi-point groundwater measurements is provided in Table B1. For time series, line 19 is used for a short description of the quality flags (see Section 3.5). The main data body varies depending on the type of data. Data not categorized as auxiliary are either time-series or tests (Tab. 1). All time-series are three column files with date and time in the first, data values in the second, and a quality flag (Section 3.5) in the third column. The structure for tests varies depending on the type of data and are explained in the respective subsections for individual variables.

3.5 Data quality assessment

Each time-series was checked automatically to flag data outside the valid instrument ranges and missing data, as well as manually to identify spurious data (Table 2). Spurious data are data points for temperature time-series that show a deviation from 0°C during the freeze-through period exceeding the instrument accuracy, or when temperatures show monotonous temperature shifts over a certain time period (n=1). Quality flags are provided in the time-series data files as a separate column ("QF"). Additional information on flagging data as spurious is provided in the relevant section.

For the tests, a quality assessment is given in the relevant sections. The assessment is based on field observations from independent datasets, experiences in the field, and assumptions on expected value ranges that are explained in the relevant sections.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Table 2. Data quality flags "QF" of time-series data.

Flag	Quality	Description
0	Good data	No identified issues other than general quality assessment issues (see respective data sections)
1	Missing data	All missing data points due to maintenance, and data cleaning
2	Instrument range	All data outside the specified instruments' valid value ranges
3	Spurious data	All temperature data where freeze-through period shows offset from 0°C outside instruments' accuracy ranges

4 Data description

4.1 Ground measurements

Ground measurements refer to any experiments and monitored time-series in the ground or at the ground-atmosphere, or ground-water interface and are stored in the main category "ground".

4.1.1 Ground temperature - "temp"

Ground temperatures at various depths and at the water-ground interface were measured continuously with three different devices (Fig. 3). These are 1) Proges Plus Thermo Buttons 22L temperature loggers (TB) for measurements at the surface in around 5 cm depth, 2) van Essen Micro-Diver DI601 (D) installed at the water-ground interface to simultaneously measure water level as well as water-ground interface temperature (or surface air temperature in case of drying out of the rivers), and 3) Onset HOBO TMCx-HD Water/Soil Temperature Sensors as thermistor chains in combination with Onset HOBO U12 4-External Channel loggers (HOBO) for measurements at various depths, ranging from 5 cm to 500 cm. All temperature measurements are stored in the category "ground/temp".

The TB loggers are installed at various positions to provide temperature boundary conditions for numerical modeling. They cover a large range of land surface and vegetation covers, shadings, and terrain expositions. The instrument names together with the two element location_id reflect the main position (R-river, S-soil, A-air) and dominant exposure (su-sun, sh-shade, wa-water/wet, mx-mixed or unclear). TBs were installed on around 20 cm long pegs made of wooden branches with the TB attached to the pegs (Fig. D1).

HOBO loggers were installed at the two main cross sections (CS) CS-2 and CS-9 (Fig. 1). The main purpose of the HOBO time series is the monitoring of active layer and permafrost temperatures. HOBOs were installed in boreholes, drilled with a gas-powered auger (Fig. D2). Some attempts were made to install thermistor chains in the river bed but the instruments always broke, possibly as a result of cable shear stress when the river was frozen. All Divers (D) were located at the bottom of the river or lake and the temperature thus represents the interface temperature. While the lake never encountered freeze-through, the river measurements showed prolonged zero-curtain periods, followed by negative temperatures (Section 4.3.1).

Uncertainty and errors: All temperature time-series were manually checked for plausibility by comparing the temperature records during the freeze-through period with expected 0°C values. None of the TB sensors show a deviation higher than their

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

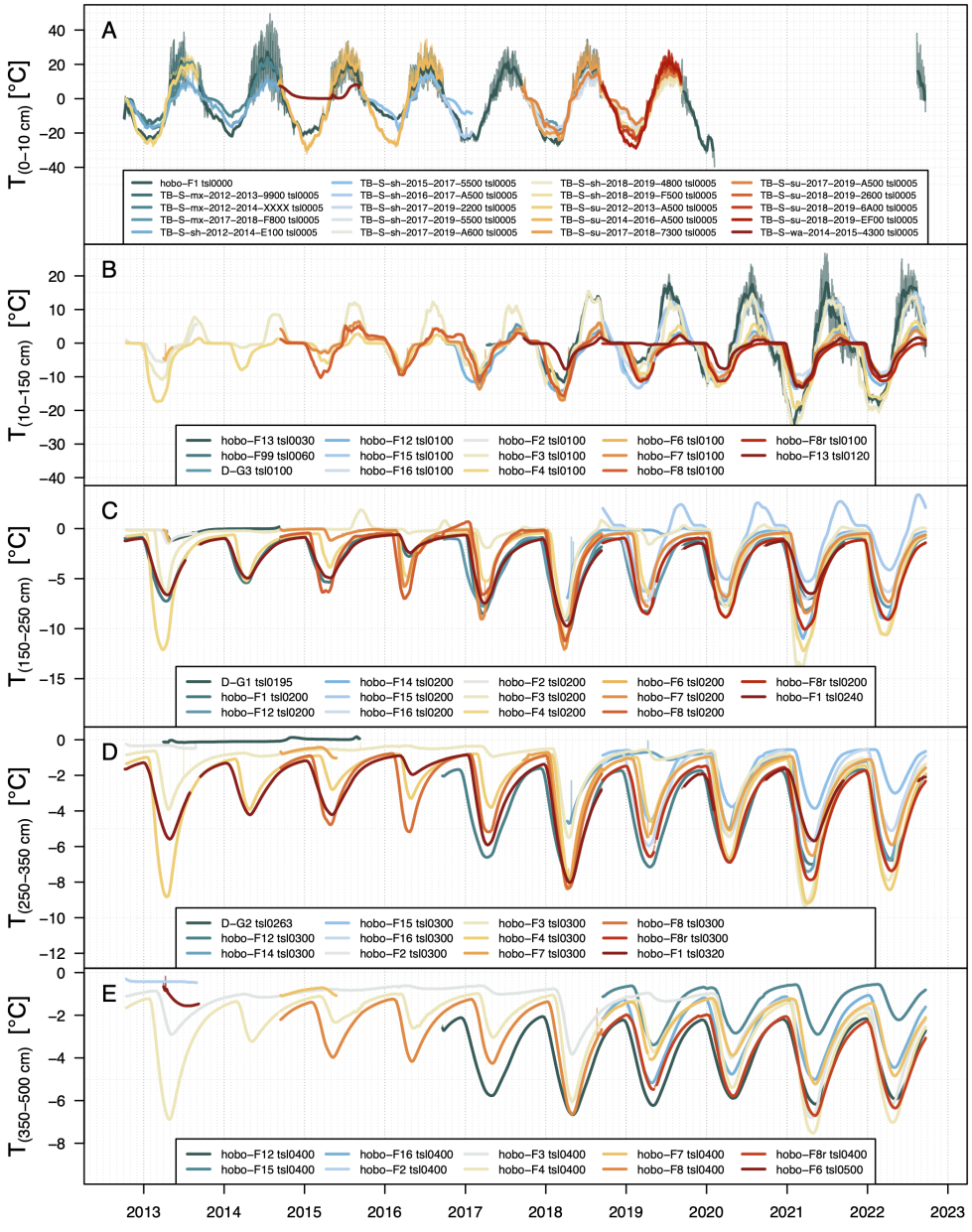

Figure 3. Ground temperature time series for different depths (A-E). Ground surface temperatures mainly TBs and only available until 2019 (A). Identified data issues are mentioned in the main text.

175 accuracy. Issues are apparent for the HOBO data for "tsl0200_hobo-F8" in late 2016, and "tsl0200_hobo-F15" in autumn and winter. These data have been flagged as spurious.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

The TB data were analyzed regarding the minimum and maximum temperatures during the freeze-through period. They show a total value range between -0.2°C and 0.2°C for the maximum temperatures with a standard deviation of 0.07°C, and a range between -0.5°C and 0.1°C for the minimum temperatures with a standard deviation of 0.23°C (n=21).

The HOBO data show a total value range between -0.12°C and 0.6°C for the maximum temperatures with a standard deviation of 0.21°C, and a range between -0.6°C and 0.51°C for the minimum temperatures with a standard deviation of 0.3°C (n=16).

The temperature signal of several HOBO time-series spike in spring and are visible through multiple soil layers. Temperatures return to their original background level or trajectory after these occurrences. Such features have been associated with infiltrating melt water from snow melt as pointed out in other studies (Vonder Mühll et al., 2004). Here, such occurrences were not flagged as spurious.

4.1.2 Thermal properties - "thermal"

Thermal properties were determined using a KD2 Pro Thermal Properties Analyzer from Decagon Devices with a 3 cm dual-needle (SH-1) sensor. Properties were measured *in situ* in soil pits (Fig. 1, Fig. D3), as well as from soil samples taken and analyzed in a laboratory at the Melnikov Permafrost Institute (MPI), Yakutsk, or in a laboratory of the Laboratoire des Sciences du Climat et de l'Environnement (LSCE), Paris. Temperatures at the measurement locations were additionally measured using an Anritsu HD-1200K thermometer (Anritsu Meter Co.) (Tab. 1). The thermal files include *in situ* volumetric soil water content (SWC) measurements (Section 4.1.4), providing basic information for the parameterization of ground properties for hydrothermal modelling. In case SWC was determined from soil samples in the laboratory, soil porosity estimates are provided as well (Section 4.1.4). The file structure provides the results for each variable within "thermal" (Tab. 1) as separate column for the individual depths (rows) of a soil pit.

Uncertainty and errors: The calculation of thermal properties of the KD2 Pro device might have high uncertainties. This is because significant differences in some cases between the also utilized Anritsu thermometer are apparent (Fig. 4). How potentially biased temperature measurements affect the accuracy of other derived parameters cannot be assessed as the relevant information is not publicly available (Tab. 1). The obtained value ranges for the different soil grain size compositions and SWC contents are similar to values obtained in other studies in the region (Zhirkov et al., 2021).

4.1.3 Thaw depth / active layer thickness - "aclt"

The ground temperatures (Fig. 3) indicate a downward heat transfer in the active layer and its thawing that lasts until September or October, with differences among sites (Fig. 5). The field campaigns were usually conducted late September for logistic reasons. As a consequence, what we report here are thaw depths in the active layer rather than absolute active layer thicknesses. The thaw depth does not necessarily represent the maximum thaw depth but only the depth to the frozen layer at the time of the measurement. The database uses the abbreviation aclt.

The depth was determined using different methods, including probing with metal rods, and drilling (Fig. D2), and as part of an independent study (Léger et al., 2023), by means of electrical resistivity tomography (ERT), and ground-penetrating

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

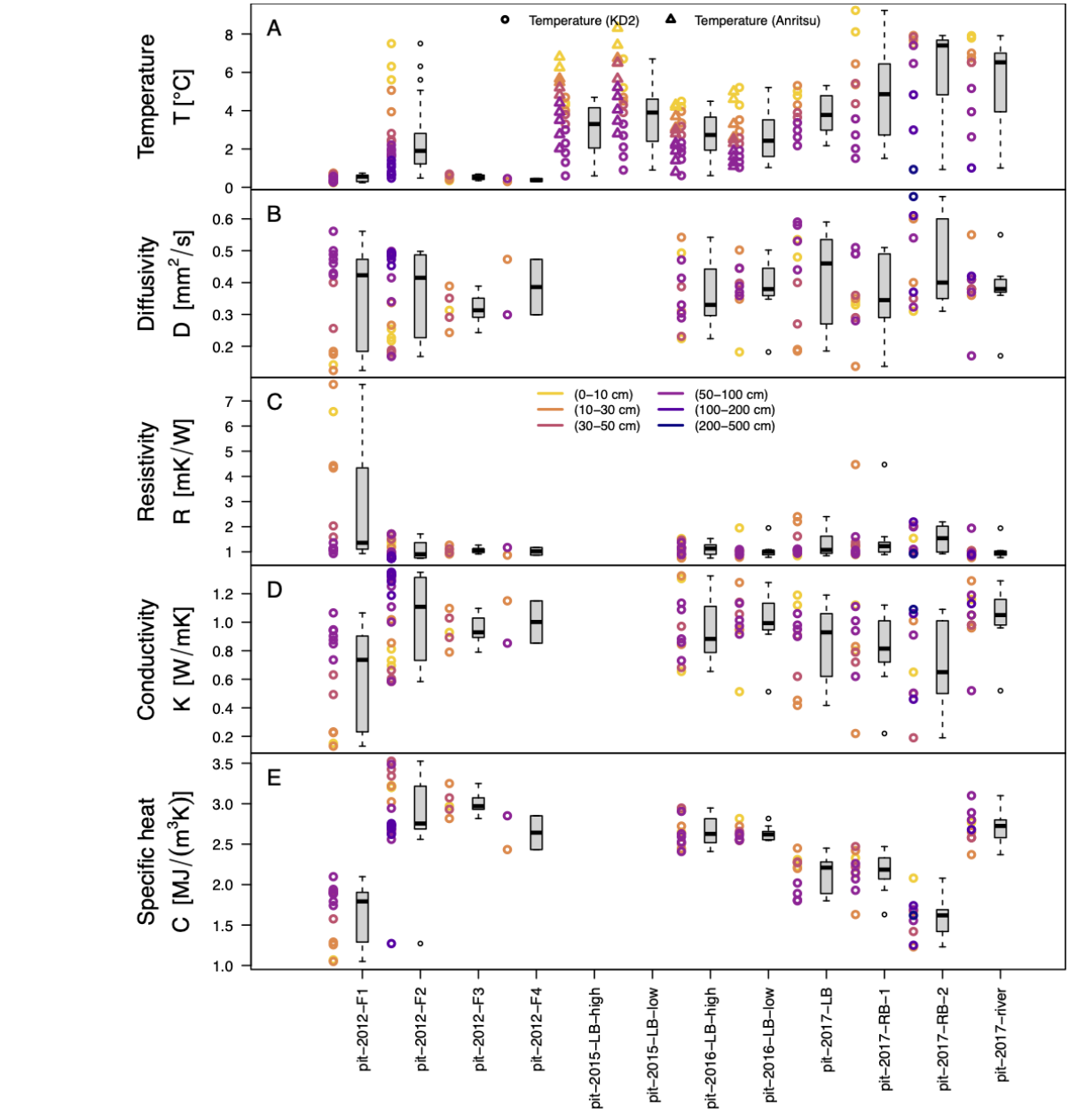

Figure 4. Various thermal properties (A-E) of soil samples analyzed *in situ* in various soil pits.

radar (GPR) (Fig. C1). The data of Léger et al. (2023) are provided in the "auxiliary/external/" folder (see Section 4.1.3 and Section 4.1.3). The regular aclt data files have five columns that provide, in addition to the location and the aclt determined,

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

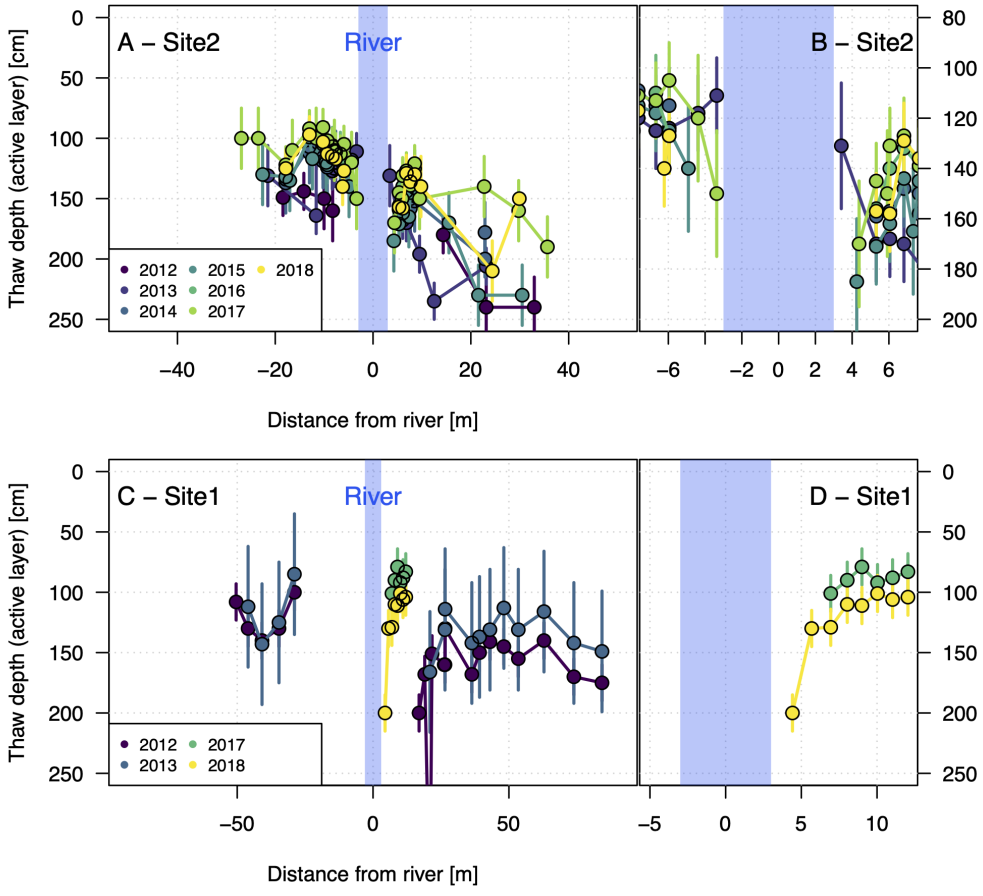

Figure 5. Thaw depths in the active layer at two cross-sections at the two main locations Site 2 (A) and Site 1 (C) with zoomed-in detail views (B and D, respectively). Error bars represent unverified assessment of measurement errors of different techniques as explained in the text.

information on whether the maximum probing depth was exceeded ("max_exceed"), and the method used to determine the aclt (see Tab. 1).

Boreholes and piezometer tubes

The depth to the frozen layer was determined by pushing a metal rod into the soft soils or through the piezometer tubes which extend through the active layer. Probing in the compacted soils on the right-hand side river bank was often not possible and a drill was used instead (Fig. D2).

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Ground Penetrating Radar (GPR)

Two ground-penetrating radar (GPR) prospecting campaigns were conducted in October 2017 and October 2018, when the upper thawed layer depths were at their annual maximums, just before the onset of the freezing period. GPR data were acquired in the time domain using the Russian OKO system, comprising one set of antennas centered on 150 MHz. In October 2017, the river had mostly dried out in its narrower part, allowing for the use of the GPR with facilitated access to the riverbed in the cross-section CS1. In contrast, on CS2, the data were acquired on each side of the water pond (approximately 50 cm at its deepest point). For all GPR data, the spatial sampling interval was set to 0.02 m, the time sampling interval was set to 0.39 ns, and the time window was adjusted to 100 ns. Further details can be found in Léger et al. (2023).

Electrical Resistivity Tomography (ERT)

Electrical Resistivity Tomography (ERT) data were acquired using a 16-channel SibER-64 system with 64 electrodes and a 0.5-m spacing between electrodes, using Dipole-Dipole, Schlumberger and Wenner configurations. Using a roll-over procedure, we were able to obtain a 63.5 m and a 71.5 long transect for CS1 and CS2, respectively. The transect for CS2 was collected in 2018 and for CS1 in 2017 and 2018. Data processing was performed prior to the inversion consisting mainly of removing extreme contact resistance values due to bad contact with the ground. Data were inverted using the finite-element inversion program BERT Günther and Rücker (2012) to obtain the spatial distribution of soil electrical resistivity, and including topography. We used a robust inversion (L1-normalization), giving a higher probability to obtain blocky models with sharp boundaries. Further details can be found in Léger et al. (2023).

Uncertainty and errors: The determination of thaw depths with metal rods in soft soils is feasible as the frozen ground provides a significant change in hardness that can easily be determined when inserting the metal rod into the ground or piezometer tubes. On the right bank near CS-9, very compacted soils prevented using the rod. The determination of the frozen boundary from drilling was subject to the experience of the drill operator. All mechanical measurements have been somewhat validated in two years when soil pits were dug next to the aclt measurements. We assume maximal uncertainties in the range of ± 2 cm for soil pits, ± 15 cm for piezo tubes and metal rods pushed by hand, and ± 25 cm for the mechanical drill. GPR- and ERT-derived estimates are in agreement with the manual estimates within the uncertainties in interpreting where the frozen layer starts from the non-invasive geophysical methods (Léger et al., 2023).

4.1.4 Soil water content - "swc"

In situ soil water content was measured with a permanently installed system for obtaining time-series since 2017, and in soil pits in different years. For soil pit measurements, the main instrument was a HydroSense I and later a HydroSense II (Campbell Scientific) time domain reflectometry (TDR) device. Regarding the performance, the only difference is an improved accuracy for HydroSense II for very high electrical conductivity (electrical conductivity > 2000 μS cm^{-1}). High volumetric water contents outside the regular measuring range (0 % to 50 %) were obtained in high organic, and porous soils near the river and below the river bed, independently calculated also from soil samples in the laboratory (Fig. 6). The time-series were

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

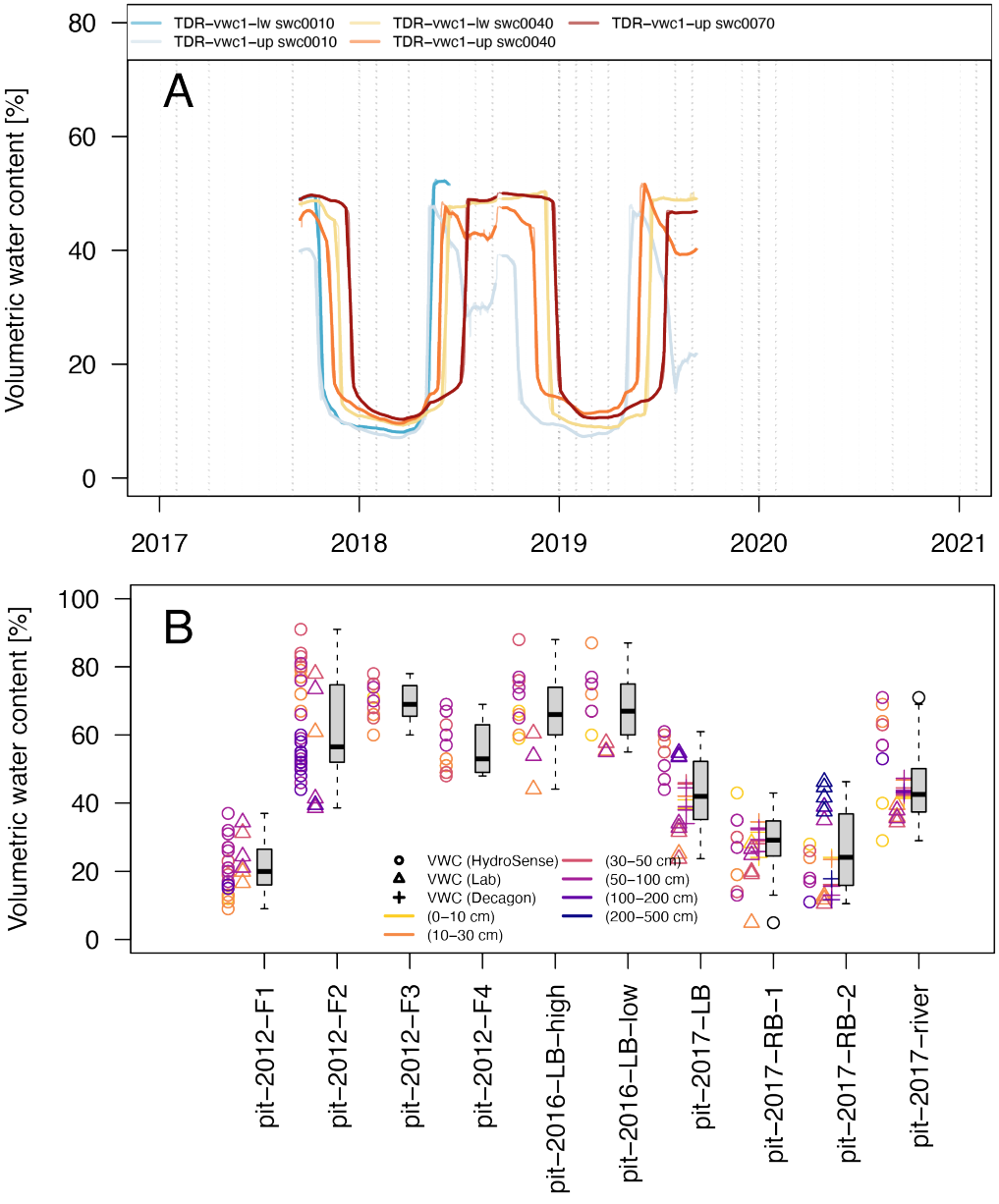

Figure 6. Soil water content measurement time-series using EC-5 probes from Decagon (A), and individual tests made in soil pits or from soil samples analysed in the field and the laboratory using multiple instruments (B).

obtained using EC-5 TDR probes (Decagon Devices) in depths of 10 cm to 70 cm, connected to a Decagon Em50 data logger. The probes were installed on two opposing sites of a soil pit made in 2017 on the left river bank. One side of the pit ("TDR-vwc1-lw") is lower and closer to the river, whereas "TDR-vwc1-up" is located about 2 m further south, further away from

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

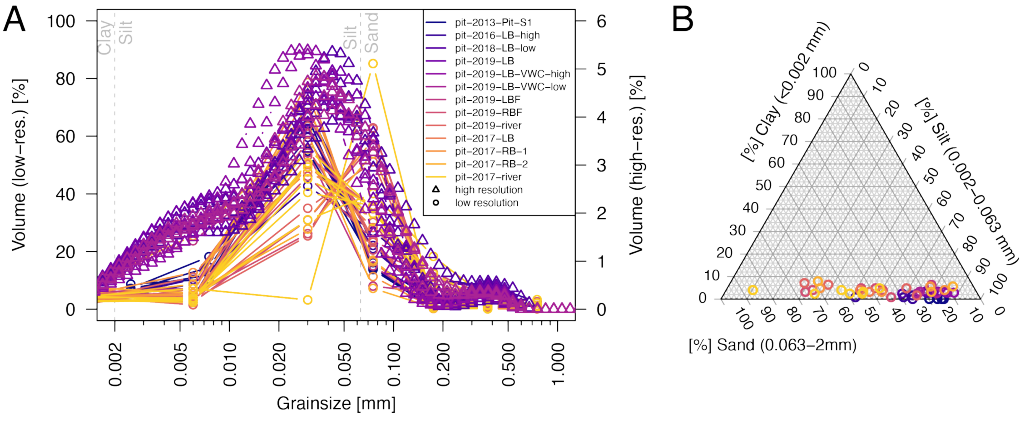

Figure 7. Grain size distributions (A) of soil samples and corresponding ternary plot (B).

the river, and slightly higher in elevation. These probes were tested in 2017 against the HydroSense II TDR probe that was used in all other years (Fig. 6). Analyses in the lab were conducted following a standard workflow of 100 mL sample weight measurements before and after oven drying at 105 °C for 24 h, following the standard by the American Society for Testing and Materials (Standard, 2019).

Soil water content is provided as volumetric water content (VWC). The HydroSense instruments are directly providing VWC estimates without any calibration performed. Within the "thermal" files, laboratory estimates (see comparison in Fig. A2), as well as EC-5 probe estimates are provided for an idea of accuracy. Single time measurements (tests) in the soil pits are included in the "thermal" files under the columns "VWC HydroSense" for the HydroSense estimates, "VWC EC-5" for EC-5 probes, and "VWC Laboratory" for the laboratory measurements at MPI or LSCE.

Uncertainty and errors: The TDR-derived VWC estimates are based on changes of the probed material's dielectric property and the instrument's internal calibration (VWC HydroSense and VWC EC-5). Therefore, differing soil types result in different results. Figure 6 shows that near the position of the continuous TDR installation "TDR-vwc1" at "pit-2016-LB", the laboratory derived estimates are consistently lower (about 40 Vol.% to 60 Vol.% vs. 60 Vol.% to 80 Vol.% for "VWC HydroSense"). The continuous TDR-based estimates of "VWC EC-5" range between 35 Vol.% to 55 Vol.% (at position "TDR-vwc1"). In different tests, where the "VWC EC-5" sensor was used, no systematic over- or underestimation can be identified, neither between the HydroSense and the laboratory estimates (Fig. 6). A maximum disagreement between individual measurements is about 30 Vol.%.

4.1.5 Soil composition - "grnl"

Grain size distributions were measured until 2017 at MPI and classified mostly using the Kachinsky classification (Kachinsky, 1968 [in Russian]) with around 10 relevant grain size classes. Grain sizes were determined using sieves with mesh sizes

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

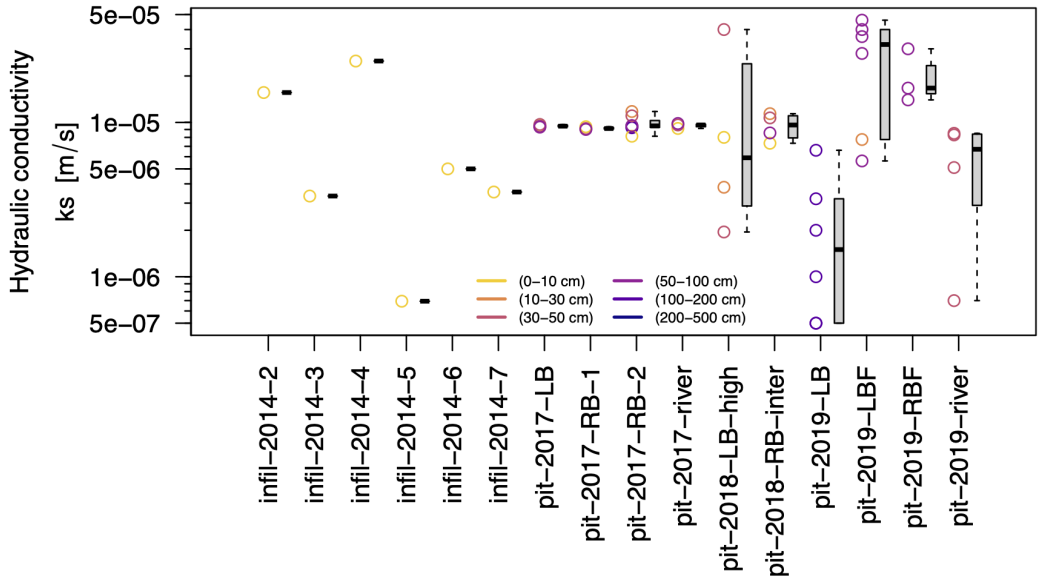

Figure 8. Hydraulic conductivity estimates determined by infiltration tests and laboratory analysis of grain size distributions using the Rosetta3 algorithm (Zhang and Schaap, 2017), at the surface and various depths within soil pits.

according to this classification (Kachinsky, 1968 [in Russian]). Deviations from the reporting classes exist. Samples after 2017 were measured using a laser diffraction grain size analyzer Mastersizer 2000 HYDRO-G (Malvern) that provides a discrimination into 100 classes. Without relevant information available to homogenize these classes, we provide the classes as they were reported. For the file structure this results in relevant grain size classes in one column and their respective weight % in a second column. Each measurement depth within a soil pit is reported as a single file. From 2012 through 2019, each year one or multiple soil pits were dug but granulometric analyses were not performed each year.

Uncertainty and errors: A comparison between the sieve- and diffraction-based results shows differences, particularly for the silt fractions (Fig. 7). As none of the samples had been measured with both methods, it remains speculative whether the differences represent a real difference in grain size distribution or whether this is the result of the method. The determined soil classes correspond to sandy silt, or silty sand soils (Fig. 7), which has also been reported by other studies for cryospheric soil types in the region (Zhirkov et al., 2021; Desyatkin et al., 2021).

4.1.6 Hydraulic conductivity - "ksat"

Multiple water infiltration tests were performed in soil pits to determine saturated hydraulic conductivity. For the years 2012 through 2016 only infiltration tests with two metal rings (13 cm inner and 20 cm outer diameter) were performed. Metal rings were pushed several millimeters into the upper soil and infiltration was measured for falling head conditions. Measurements were taken with respect to the upper edge of the inner ring and written down every few seconds to every few minutes the longer

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

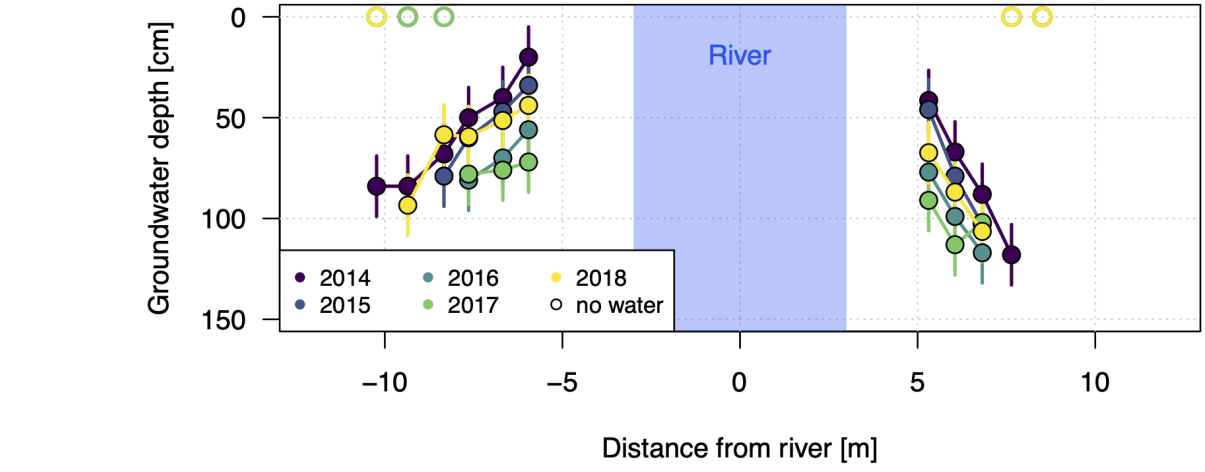

Figure 9. Groundwater depths measured in piezometer tubes at CS-9 using an electric measuring tape. Error bars represent a potential error of 15 cm (see uncertainty estimation), as was the case for the aclt measurements (see Fig. 5).

the experiment would run. Regression over the quasi-linear part for saturated conditions are reported. In addition to the *in situ* infiltration tests, for the samples from 2017 through 2019, grain sizes were determined from soil samples in the laboratory (Section 4.1.5) to derive the saturated hydraulic conductivity using the Rosetta3 algorithm (Zhang and Schaap, 2017). The applied method for the determination of ksat is reported in a separate column ("method") in the relevant files.

Uncertainty and errors: The highly compacted soils on the right bank required mechanical work to insert the infiltration rings into the soil. While no direct leakage was visible as a result of this, such leakage would result in overestimated infiltration rates and ksat. The comparison of ksat of the right bank shows in fact higher ksat compared to some of the river and left bank experiments (Fig. 8). However, the latter have experienced highly saturated conditions that might have caused the low values. For such saturated conditions, single sites (e.g. "pit-2019-LB") show a significantly wider range in ksat values, which can span one order of magnitude.

4.1.7 Ground water depth - "gwd"

During the end of summer field campaigns, attempts were made to measure the water table height inside the piezometer tubes near CS9 and CS1. The water level was determined by using a electrical measuring tape, identifying the distance between wet tape and top of the piezometer tubes. The distance was adjusted for the height of the piezometer tubes above the land surface.

Uncertainty and errors: The piezometer tubes were installed in drill holes that reached the frozen layer. Consequently, the hydraulic head measured in the tubes might be disconnected from the one in the thawed layer, and instead might represent the hydraulic head of a previous time period where the connection existed. An indication for an existing direct connection, however, is the year 2017, when the river mainly dried out and gwd showed the lowest values (Fig. 9). At the same time,

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

the actl was relatively shallow in comparison to other years (Fig. 5), suggesting no preferential hydraulic head connection compared to other years.

Additionally, some gwd measurements show a deviation from a gradient of continuously lower gwd with increasing distance from the river (Fig. 9). Based on these deviations, an error of around 15 cm is estimated.

4.1.8 Radiocarbon dating of soil organic carbon - "age"

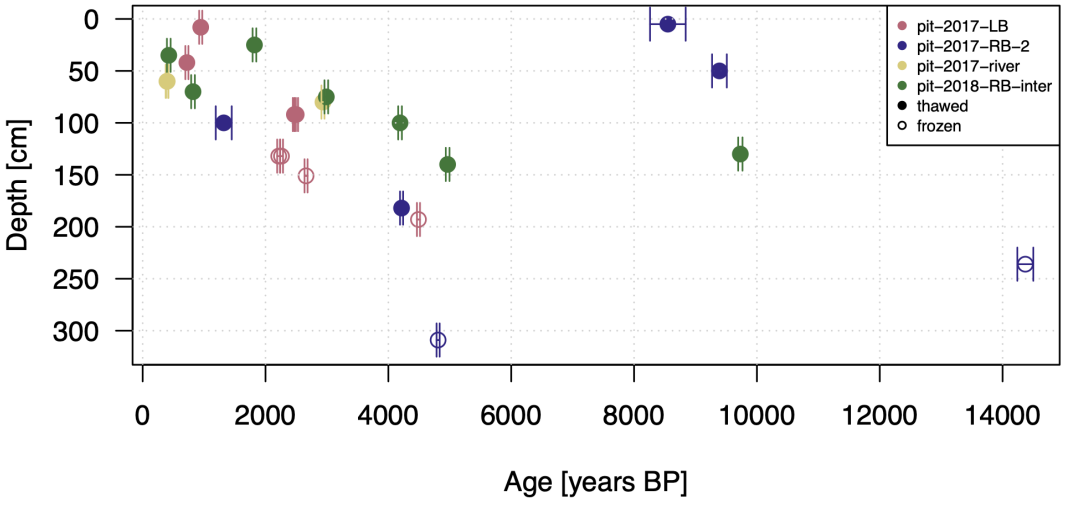

Figure 10. Radiocarbon dating (^{14}C) at Site2. Error bars represent ±1 standard deviation.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Radiocarbon dating was performed on soil samples of four soil pits at Site 2 (Fig. 10). The dating data were obtained at LSCE following standard procedures. Soil (permafrost) samples were chemically prepared according to the protocols published in Hatté et al. (2024). The preparation involved decarbonation with 1N HCl. Carbon-rich samples were subsequently oxidized into CO_2 and reduced to graphite carbon (Cgraphite) using an automated device (AGE; Wacker et al., 2010b). They were then introduced in the solid source of the MICADAS atomic mass spectrometer (AMS) instrument (Synal et al., 2007) operated at LSCE (Tisnérat-Laborde et al., 2015). Carbon-poor and small samples were introduced in an elemental analyser connected the gas source of ECHoMICADAS, through a gas interface system (EA-GIS; Ruff et al., 2010). Raw data processing and reduction were carried out using BATS software (Wacker et al., 2010c) and a custom calculation procedure (Thil et al., 2024). Results are expressed in $F^{14}C$ and normalized to the international standard, Oxalic Acid 1.

Uncertainty and errors: Measurement errors or uncertainties arise from the statistical counting of ^{14}C ions in the AMS (MICADAS) and from the propagation of errors during blank subtraction and normalization using Oxalic Acid I, the standard. Statistical counting depends on the sample mass and the measurement duration. Thus each individual age estimation is provided with an associated unceertainty in the respective files. Some old ages are apparent in the topsoil samples ("pit-2017-RB-2"). We do not investigate further the possible reasons for that but soil disturbances from road works and soil redistribution along the hillslope were visible in the field (Fig. D4). Deeper layers display continuously older ages and the frozen zone contained abundant ice lenses, both suggesting unlikely mobilization and disturbance of the soil substrate.

4.2 Atmospheric measurements - "atmo"

4.2.1 Temperature - "temp"

Air temperature is measured in the transition zone between the meadow and forest at Site2 either with a TB (Section 4.1.1) in various years, with a T&D-502 temperature sensor (TAIR) between 2017 and 2018, or with a vanEssen Micro Divers DI601 (D) (Section 4.3.1). All instruments were installed at 2 m elevation from the ground at an approximate elevation of 149 m a.s.l. The closest long-term meteorological station is located in Yakutsk (Yakutsk - WMO ID=24959, 103 m a.s.l.). The data archive for Yakutsk was obtained via the website rp5.ru (Raspisaniye Pogodi Ltd., 2004) and is provided under *auxiliary*. A correlation analysis for the period 2017-2018 (Fig. A3) shows a coefficient of determination of 0.991.

Uncertainty and errors: The high correlation with the WMO meteorological station in Yakutsk suggests a reasonable representation of air temperatures. However, stronger deviations are present in winter (Fig. A3). No analysis on possible reasons for the seasonal deviation has been performed.

4.3 Water measurements - "water"

While the focus of this database is revolving around the thermal evolution and properties of the ground, some chemical and physical analyses on the river, thermokarst-, and alas-lakes were conducted. However, a much more detailed analysis and database on river and lake physio-chemical properties exist (Hughes-Allen et al., 2020; Hughes-Allen et al., 2021).

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

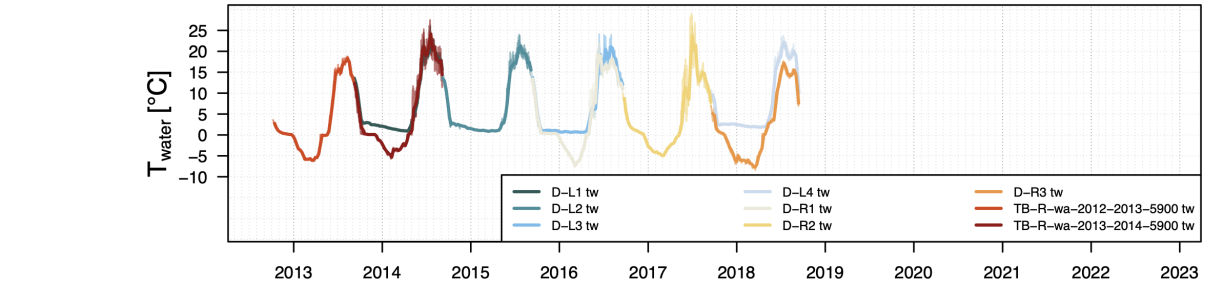

Figure 11. Water temperatures in the upstream lake and the river.

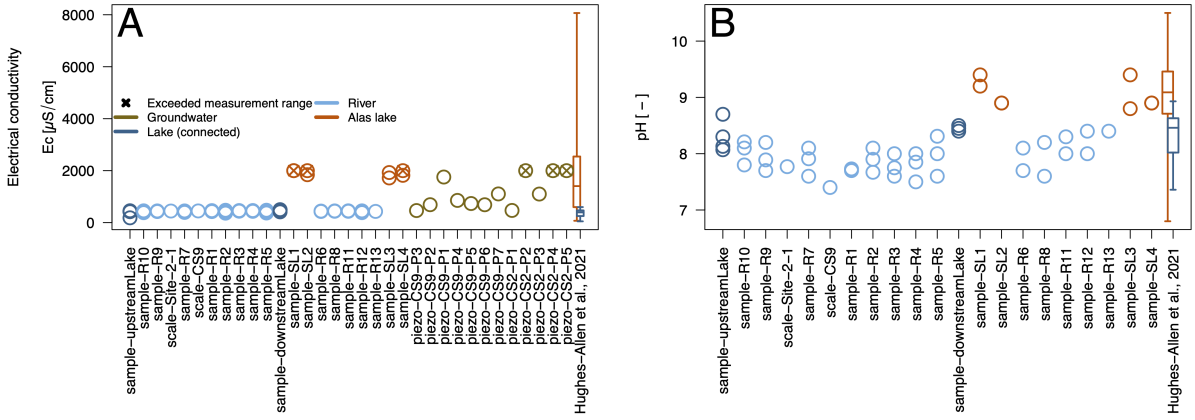

Figure 12. Water electrical conductivity (A) and pH value (B) of water samples. Additional analysis of lake samples from Hughes-Allen et al. (2020) and Hughes-Allen et al. (2021) are available in the respective datasets and publication.

4.3.1 Water temperature - "temp"

Continuous water temperatures and water level were measured using vanEssen Micro Divers DI601 (D) sensors. The sensors were installed in the upstream lake, at around two meters general depth, deep enough to be below the winter ice cover. The sensors were held in place on the lake floor with a weighted bag, and secured with a rope approximately 2 m to 5 m away from the shoreline. Additional sensors were installed inside piezometer tubes located near CS-9 in the river bed as well as in a piezometer tube on land.

Uncertainty and errors: The water temperature data were analyzed regarding the minimum and maximum temperatures during the freeze-through period for the sensors located in the river and in the piezometer tubes (groundwater). They show a total value range between -0.01°C and 0.31°C for the maximum temperatures with a standard deviation of 0.16°C, and a range between -0.03°C and 0.1°C for the minimum temperatures with a standard deviation of 0.16°C (n=6).

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

4.3.2 Water electrical conductivity - "cond"

Electrical conductivity was measured on-site using a WTW multi-parameter meter (accuracies of ±0.1 °C for temperature, ±0.2 for pH value and ±0.5% for specific conductivity) in 2017 or a YSI Pro DSS multi-parameter probe (±0.2 °C for temperature and ±1.0% for specific conductivity) (Fig. 12). Again the interested reader is referred to the much more comprehensive dataset and study in the region by Hughes-Allen et al. (2021). Some measurements in thermokarst lakes obtained with the WTW multi-parameter meter exceeded the instrument's maximum electrical conductivity range of 2000 μS cm^{-1}.

Uncertainty and errors: The YSI multi-parameter probe was calibrated at the start of each field mission, but could not be calibrated daily due to logistical constraints. However, multiple calibration sessions over the years have shown that the sensors, including the conductivity sensor, do not derive significantly.

4.3.3 pH values - "pH"

A YSI Pro DSS multi-parameter meter, as well as the WTW multi-parameter meter were used to determine the pH values.

Uncertainty and errors: As for conductivity (see above), the YSI multi-parameter probe was calibrated at the start of each field mission, but could not be calibrated daily due to logistical constraints. However, multiple calibration sessions over the years have shown that the sensors, including the pH sensor, do not derive significantly.

4.3.4 Water levels and depths - "levl"

Two types of measurement are in the following referred to as water level "wl" measurements. These include continuous monitoring of water column pressure in lakes, the river, and in piezometer tubes, as well as manual measurements of water depths using meter bands and measuring rods at selected cross-sections between the upstream lake outflow at Site1 and the highly instrumented Site2 downstream (Fig. 13, 14). Discharge measurements were conducted not in all cases when water levels were taken. Additionally, no flow conditions occurred in 2017, 2019, and 2021. Atmospheric pressure probes were only installed between September 2014 and September 2016 at location D-ATMO, at the same location as TAIR, and at an elevation of 2 m above the ground. The data file is provided in "auxiliary/external". In order to correct the water level loggers for the influence of atmospheric pressure in all years, we utilized hourly data from the Yakutsk meteorological station (World Meteorological Organisation ID: 24959, Raspisaniye Pogodi Ltd. (2004)). For this, we first derived a relationship with the atmospheric pressure time-series taken in Syrdakh between 2014 and 2016 using a vanEssen Micro Diver DI601 sensor (D-ATMO). The correlation between the Syrdakh and Yakutsk meteorological station atmospheric pressure showed an coefficient of determination of 0.96 (Fig. A1). The obtained linear relationship was then used to adjust the atmospheric pressure of the Yakutsk meteorological station data. The data is provided in "auxiliary/external". All water level data has been corrected with so-adjusted atmospheric pressure data.

Uncertainty and errors: Data points were removed for when the temperatures showed freezing temperatures. This concerns the river and groundwater measurements only. Low variation in the water level signal, together with a low signal-to-noise ratio suggest a related error of around ±10 cm (Fig. 13B, C). Additionally, the water level data show jumps that are not correlated

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

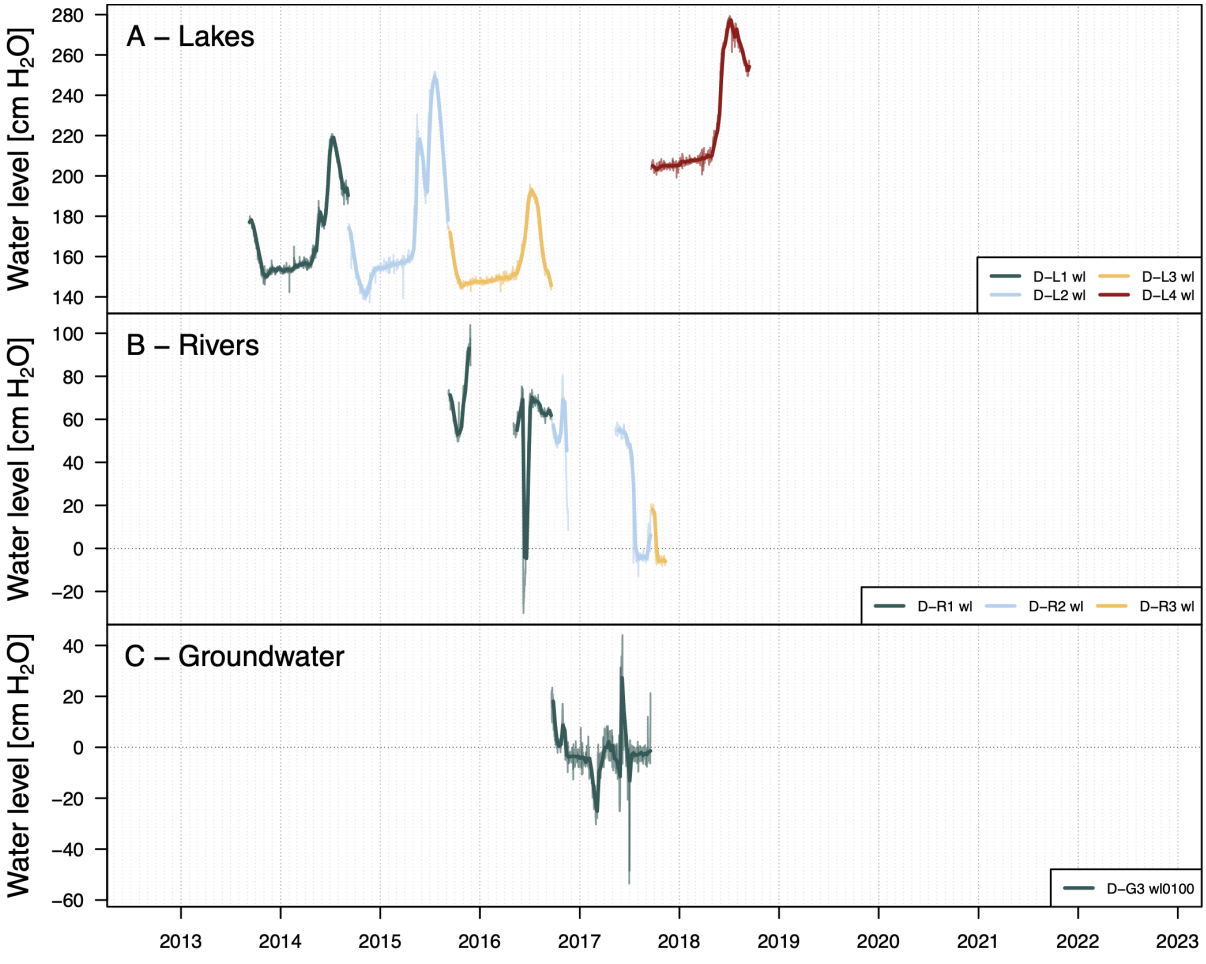

Figure 13. Water level measurements in the upstreamLake (A), in piezometer tubes along the river for river water level (B) and groundwater level (C). Extreme value differences for river and groundwater might indicate issues with pressure sensors due to freezing. Also visible are the continuous but slightly negative values around 0 cm for groundwater. It is unclear if this is a result of the extrapolated atmospheric pressure values based on the Yakutsk meteorological station (see text).

with the signal from the upper lake (Fig. 13). However, a correlation of water level peaks peaks between the river and groundwater level can be seen for the short observation period in Figure 13. We cannot resolve whether these signals represent actual water level changes in the river and groundwater or not. The user is advised to treat the data with care as they might be wrong.

4.3.5 Water velocity (discharge) - "flux"

Water velocities were measured using a handheld electromagnetic current profiler BFM 801 by HYDREKA. Discharge was calculated subsequently using Python scripts interpolating the individual measurement points to a regular grid before integrat-

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

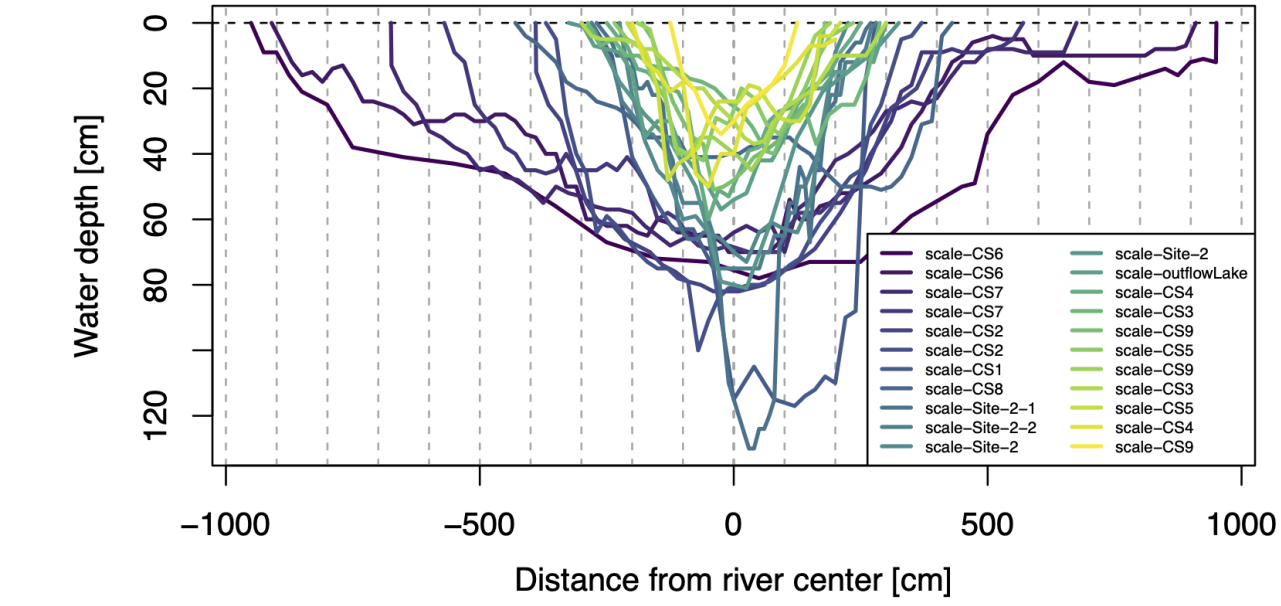

Figure 14. Water depths at cross sections along the river. Most of the profiles measured in 2018 at low flow conditions. Cross sections are ordered by cross section area.

ing (Section 4.4.3). Measurements were made along a measuring tape fixed across the stream on specific CS, at various depths measured using a measuring rod (Fig. 15). The instrument's minimum required water depth is 5 cm, and its accuracy is 0.5% of the measured discharge.

Uncertainty and errors: In the years 2016 and 2018, multiple measurements were conducted at different positions along the river within a same week. The derived discharge estimates differ by a factor of two (Fig. 15), while there was no indication in the field to expect such a difference. The measurement points had no systematic influence as the reversal of higher and lower estimates for the points near the inflow ("inflowLake", "CS-1", "CS-2"), and the points in the further downstream segment ("CS-9") show. From the presented differences between two points on the same river, the error can be assumed to be at least of the magnitude, i.e. around $301\,s-1$. Uncertainties from the interpolation to a regular grid, especially when using large cell sizes can be as much as 20%.

4.3.6 Water stable isotopes - "istp"

Water samples were taken in between 2013 and 2016 and analyzed for stable isotopes ^{18}O, and ^{2}H in at GEOPS laboratory in Paris-Saclay. The measurements were determined by cavity ring down laser absorption spectroscopy (CRDS, DLT-100 LWIA Los Gatos research; GEOPS-LSCE Panoply platform). The water isotope contents are reported in the conventional δ notation per mil (‰) as a deviation from the V-SMOW (Vienna Standard Mean Ocean Water)(Fig. 16). The accuracies are 1.5‰ for 2H and 0.2‰ for 18O (H2O).

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

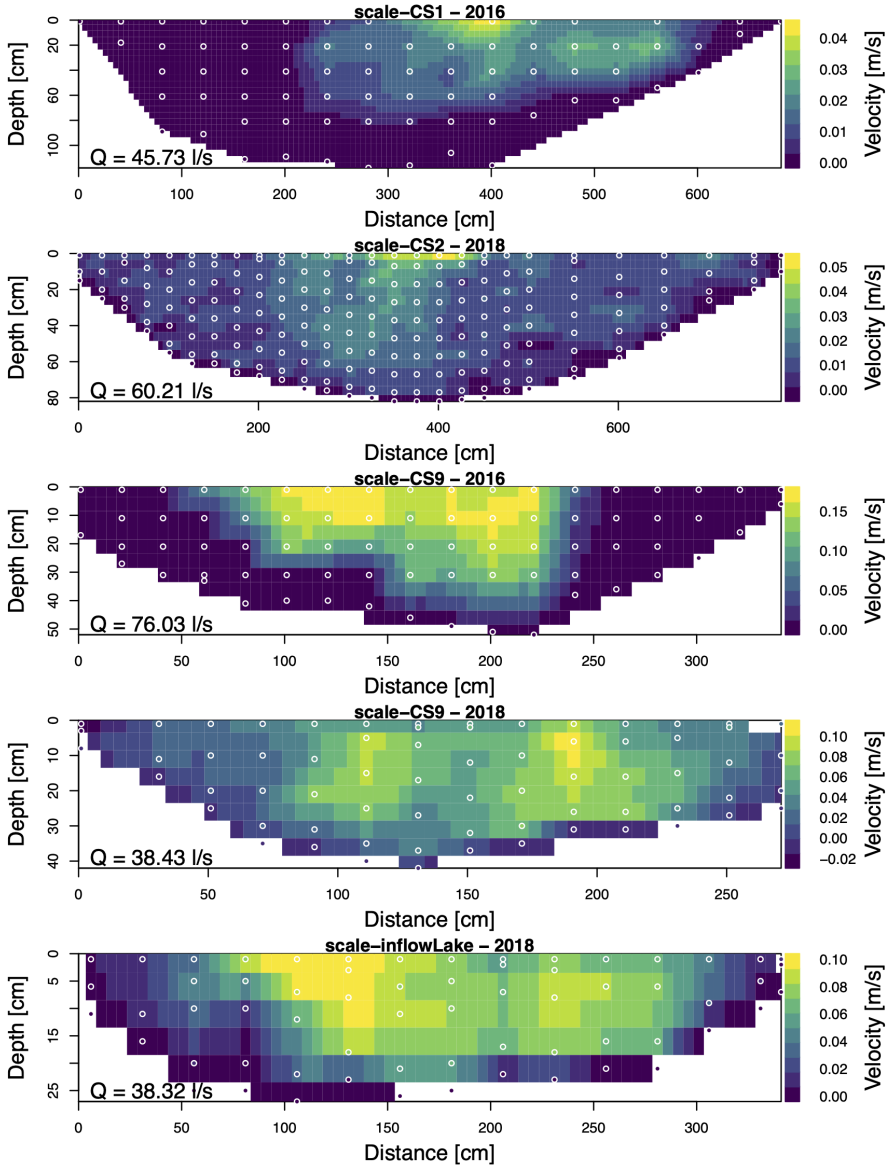

Figure 15. Water velocities at different cross-sections and years. White-circled points represent point-wise measurements made with an electromagnetic current profiler. The regular grid is produced by linear interpolation to a regular grid size resolution of 5 cm. The script to calculate the discharge ("/auxiliary/scripts") allows to set different interpolation resolutions.

Uncertainty and errors: We assume non-significant errors from the sampling, storage and and transport. The results showed consistent estimates with no deviations regarding the meteoric water lines (Fig. 16).

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

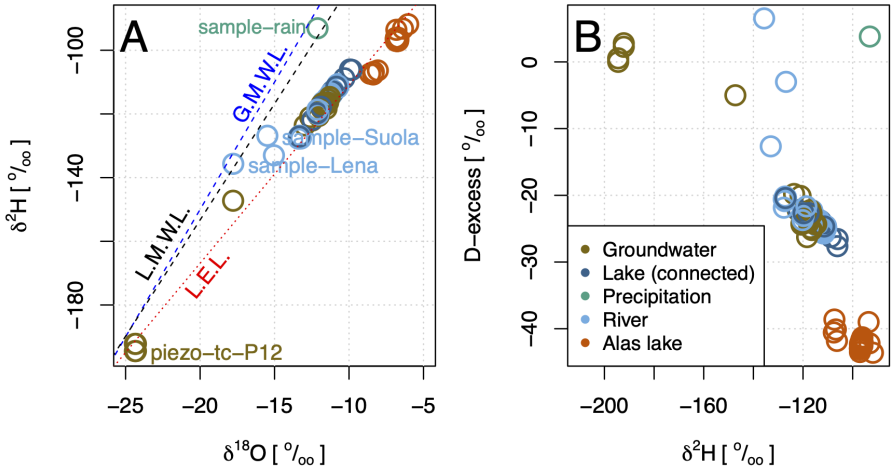

Figure 16. (A) Stable water isotope concentrations with respect to Vienna standard mean ocean water of water sources at the field site and two river samples from the Suola and Lena rivers, and (B) Deuterium excess values. Global meteoric water line (G.M.W.L.), Local meteoric water line (L.M.W.L), and Local evaporation line (L.E.L.) from Ichiyanagi et al. (2003).

4.3.7 Dissolved organic carbon - "DOC"

Data for dissolved organic carbon (DOC), dissolved oxygen saturation (DO), and other water chemistry parameters were measured in the region by Hughes-Allen et al. (2020); Hughes-Allen et al. (2021). These data are available in a PANGAEA repository (Hughes-Allen et al., 2020). The present database includes these data in a homogenized way, where file and folder structure and descriptions have a matching format. The measurement points are included in the QGIS project with a label prefix "sample-HA". Processing protocols and sampling strategy are described in detail in Hughes-Allen et al. (2021). DOC was determined using a Shimadzu TOC-L Total Organic Carbon Analyser series (SDN:L22::TOOL1760). The samples were filtered using baked glass fiber filters (Whatman GF/F, 0.7 μm), acidified to pH 2 with ultra-pure HCl and stored in baked glass vials (Hughes-Allen et al., 2020). DOC concentration was then measured in the TOC-5000A analyzer with a quantification limit of 1 mg·L^{-1}, and an associated analytical uncertainty of $\pm$0.1 mg·L^{-1} (Hughes-Allen et al., 2020). Reference material included ION-915 ([DOC]= 1.37 $\pm$ 0.41 mgC·L^{-1}) and ION 96.4 ([DOC]= 4.64 $\pm$ 0.70 mgC·L^{-1}) (Hughes-Allen et al., 2020). The exact sampling dates are not available in the original data repository (Hughes-Allen et al., 2020). Instead, only the seasons (Spring, Summer, Fall, Winter) were reported. This information is provided in a separate column in the homogenized CSV files. The different water source types are indicated in an additional column as lake (l), Alas or thermokarst lake (tl), river (r), or groundwater (g).

Uncertainty and errors: As for the isotope analysis, we do not identify any significant errors or uncertainties outside the analytical uncertainty. Lake and river samples were taken in distance of the immediate shorelines and no deviation from standard sampling protocols occurred, nor could we identify any suspicious data points within the analyses.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

4.3.8 Dissolved oxygen saturation - "DO"

DO was measured alongside DOC by Hughes-Allen et al. (2021) using a YSI Pro DSS multi-parameter meter sensor. The relevant data in their original formatting can be obtained via: Hughes-Allen et al. (2020). As for DOC, the sample locations have the prefix "sample-HA". The same additional columns as for DOC are provided in the data files (see Section 4.3.7).

Uncertainty and errors: As for conductivity and pH (see above), the YSI multi-parameter probe was calibrated at the start of each field mission, and could also be calibrated daily using a calibrating cup. Multiple calibration sessions over the years have shown that the sensors, including the DO sensor, do not derive significantly.

4.4 Auxiliary measurements and data - "auxiliary"

4.4.1 UAV derived digital surface model (DSM) - "dem"

Imagery of an UAV survey in 2021 (DJI MavicPro 2) was used to create a digital surface model (DSM) and an orthomosaic image using Pix4D software. The images have a spatial resolution of 2.2 cm pixel^{-1} and are provided in UTM zone 52N projection.

Uncertainty and errors: The DSM has a significant problem in the elevation gradient. While there is an overall decrease downstream in elevation, the middle section is slightly higher than the stream section close to the upstream lake. This would prevent any water flow downstream. Despite having used multiple ground control points in 2021, we were not able to resolve this problem. The reasons may be due to the uncertainty in GPS positions of the drone between two flights or because it is a DSM that includes the vegetation such as tall grass. The error in vertical elevation by investigating the river channel is estimated to be around 2 m. The user is advised to modify the DSM for hydrological applications that require a continuous gradient from the upstream to the downstream lake. The lateral gradients are much stronger and allow direct usage of the DSM without modifications.

4.4.2 Geographic positions - "gps"

Positions of measurement locations were recorded using one or multiple of the three devices: 1) Leica Viva Uno 10 (GPS+GNSS) differential GPS (dGPS) system with around 0.5 m absolute accuracy, 2) a Ural Optical and Mechanical Plant 3T5KP Theodolite with 2 (5) arcseconds horizontal (vertical) resolution, 3) GARMIN handheld GPS with around 10 m accuracy. The dGPS was used in individual campaigns to measure points of transects but the base stations were re-initiated for different cross-sections and different measurement locations. The resulting total differences in location coordinates required to adjust points manually. For this we used the orthomosaic photo and aligned the points to known fix points, like the river shore line or forest-meadow boundary. In case a user requires more accurate or differently-aligned points, it is possible to adjust the positions of features in the GIS project and extract the resulting adjusted positions using a provided Python script (Section 4.4.3).

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Uncertainty and errors: The uncertainties from using the dGPS with re-initializations for different measurement locations are probably within the provided accuracy estimates. By adjusting points manually to align with identified points in the ortho-mosaic, we additionally introduce a maximal horizontal error of around 3 m.

4.4.3 Processing scripts - "scripts"

The database provides four categories of scripts for (i) "plotting", (ii) extraction and management of "coordinates", (iii) data comparison and gap filling via "regression" analyses, and (iv) "discharge" calculation. The scripts to produce the figures within this manuscript, conduct the regression analysis, and to calculate discharge from water velocity measurements are provided in the R language, and scripts for the extraction of coordinates from the vector files are written in Python.

Plotting

The plotting functions used to produce the figures in this manuscript showcase the file-specific loading and processing routines for any of the CSV files in the database. The sub-folder structure of the plotting functions follows the organization of the main database by each individual variable. Additionally, one script ("/overview") extracts all meta-information of all CSV files in the database for the creation of the overview figure (Fig. 2).

Coordinates

The coordinates extraction script is designed to achieve an independence of the GIS files and analyses of measurements in a geo-spatial context. As described in Sections 3.2 and 4.4.2, point measurements inherit inaccuracies resulting from point registrations with handheld GPS, or re-initiated dGPS receivers (Section 4.4.2). The users can check within the ortho-mosaic and derived DSM the positions of points and make adjustments if needed within a GIS. The script then allows for fast extraction of so-adjusted coordinates for further processing (e.g. plotting with spatial referencing). The script automatically creates a CSV file named coordinates_<YYYY-MM-DD>.csv, where <YYYY-MM-DD> is the date of creation as year, month, and day. The file is created in the main database folder and includes the coordinates as latitude and longitude, Easting and Northing in UTM zone 52N coordinates, and the point measurement labels. The point labels are referenced in each CSV measurement file, either directly within the file name for single point measurements, or as indices in case of multiple point measurements.

Discharge

The discharge is calculated by first interpolating the existing arbitrarily distributed point-wise measurements of stream velocities over a river cross-section onto a regular 2D grid using the R library "akima" (Akima and Gebhardt, 2022). The grid resolution can be defined by the user. The interpolated values are finally summed up and its units converted to liter second^{-1}.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Regression

As highlighted in Sections 4.3.4 and 4.2.1, atmospheric pressure and temperature were not continuously available. In order to use data from the nearest meteorological station Yakutsk (WMO ID=24959, 103 m a.s.l.), the scripts in the sub-folders "Diver" and "TAIR" perform a regression analysis for the overlapping periods. From the obtained linear models, air pressure, and air temperature are reconstructed for the data periods with no local measurements. The resulting data files are contained within the respective folders under "regression".

4.4.4 Vector files for GIS - "geopackage"

Each measurement location is stored in a point or polygon vector file in GeoPackage format under "auxiliary/geopackage". The naming of GeoPackage files and contained points follows the main instrument or measurement technique applied to obtain the measurement. This means that, e.g., an infiltration test in a soil pit is not separately listed as infiltration point, but instead the location of the soil pit is provided. The relevant location names are explicitly stated in the CSV files. Polygon vector files are available for most of the soil pits. Each GeoPackage file has, in addition to the location, attributes about start and end date of the measurements conducted at an individual position, and a boolean flag whether the exact date of the measurements was known. This flag was required as some data (DO, DOC) were only associated with a season, and some measurements were conducted within a time window over multiple days.

5 Conclusions

This database shall serve specifically the development of thermo-(hydrological) modeling code, in a region where changes to permafrost under climate change are expected but data are sparse. With a focus on ground temperatures at different depths and by providing various topsoil temperatures, the difficult-to-account-for heat transfer through snow layers is avoided. Shallow soil temperatures from various landscape units with different expositions and vegetation allow for the analysis of air-ground temperature relationships and explore the inherent small-scale variability this has on heat transfer. Ultimately, this shall help improving the code within large-scale land surface models that are required to obtain more reliable future climate estimates. The rich set of supporting data on ground physical properties and water chemistry will allow setting much-needed boundary conditions and validation of modeling exercises.

6 Code and data availability

The code and data are available in combination as part of a zenodo database under CC 4.0 license (https://doi.org/10.5281/zenodo.14619854; Pohl et al. (2025)).

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

Appendix A: Data correction and validation with third-party datasets

The following shows correlation analyses to reconstruct incomplete atmospheric pressure, and air temperature data using continuous time-series from the meteorological station in Yakutsk (WMO ID=24959, 103 m a.s.l.). The data were homogenized using simple linear regression. Obtained relationships are directly shown within the figures. Furthermore, a comparison of soil water content obtained from different TDR probes against laboratory estimates are shown.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

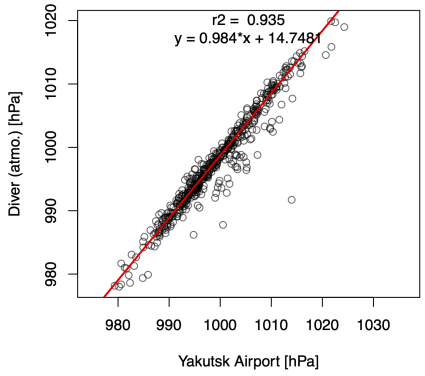

Figure A1. Comparison of atmospheric pressure measured since 2017 in the Syrdakh village and Yakutsk meteorological station (WMO ID=24959). The atmospheric pressure from Yakutsk was used then to correct the water pressure sensors for the influence of atmospheric pressure variability.

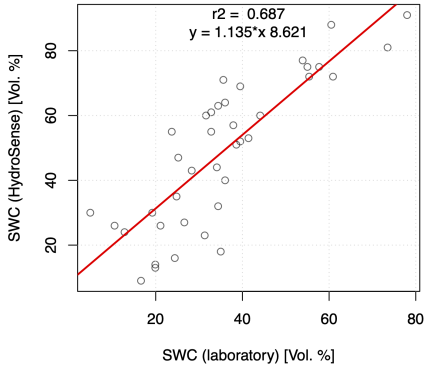

Figure A2. Comparison between soil water content (SWC) determined in the laboratory (LSCE or MPI) and in situ using the HydroSense instruments.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

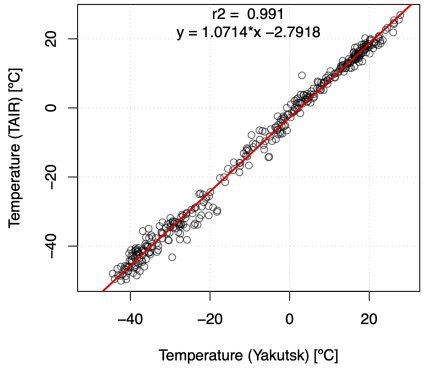

Figure A3. Comparison between meteorological station Yakutsk (WMO ID=24959) and T&D RTR-502 (TAIR) temperatures.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Appendix B: File header examples

The following section shows an example of header information within the two main data types: time-series and tests (Table B1).

Table B1. Example of CSV file metadata with 22 lines of generalized categories (left column). Two examples, one for a soil temperature time-series of a TB (middle), and one for groundwater depth measurements at various points (right column).

1	variable_id	tsl0005	gwd
2	variable_name	soil_temperature_at_5_cm	ground_water_depth
3	unit	deg.C	cm
4	experiment_id	syrdakh	syrdakh
5	source_id	TB-S-sh-2017-2019-2200	piezo-tube
6	longitude	NA	NA
7	latitude	NA	NA
8	time_start	2017-09-17T00:00:00	2015
9	time_end	2019-09-10T09:00:00	2015
10	time_res	3 (hours)	NA
11	time_format	YYYY-MM-DDTHH:MM:SS	YYYY
12	time_continuous	2.6 day(s) max gap	NA
13	source_file	ThermoBoutons_2012-2019_fixing2024-07-03.xlsx	GWD_2012-2018_EP_2023-11-30.xlsx sheet = GWD
14	comments	thermobutton time series	aclt measurements
15	type	time_series	test along transect
16	max_value	85	200
17	min_value	-40	0
18	na_value	NA	NA
19	quality_flag[a]	0 - good data;1 - missing data;2 - outside measurement range;3 - spurious data	
20	# more space for meta data ideas		
21	# more space for meta data ideas		
22	# last row of header/meta_data		

[a] Quality flag entry only for time-series data.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

Appendix C: ERT and GPR data

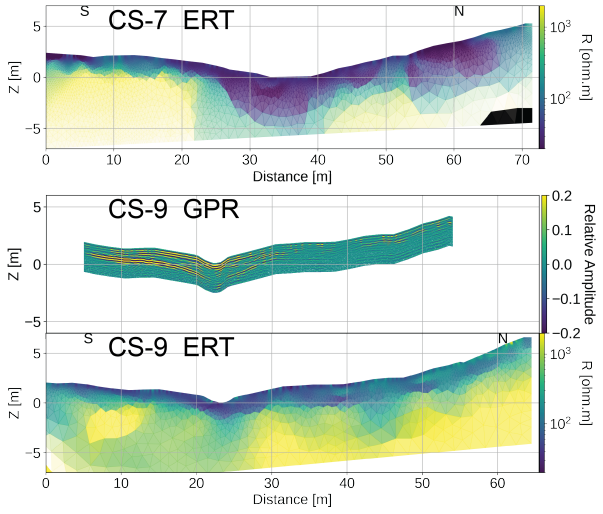

Figure C1. Geophysical measurements at CS-7 (ERT) and CS-9 (ERT+GPR) from Léger et al. (2023). For details on data processing see their publication.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

Appendix D: Photos of instruments and locations

Figure D1. Proges Plus Thermo Buttons 22L (TB) attached on wooden peg. Installation by pushing or hammering into the soil to a depth of approximately 5 cm.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

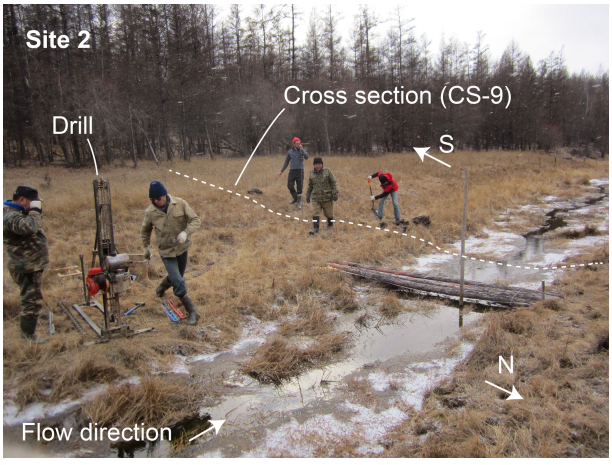

Figure D2. Installation of HOBO thermistor chains, and piezometer tubes at Site 2 (CS-9) with a petrol powered standing drill.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

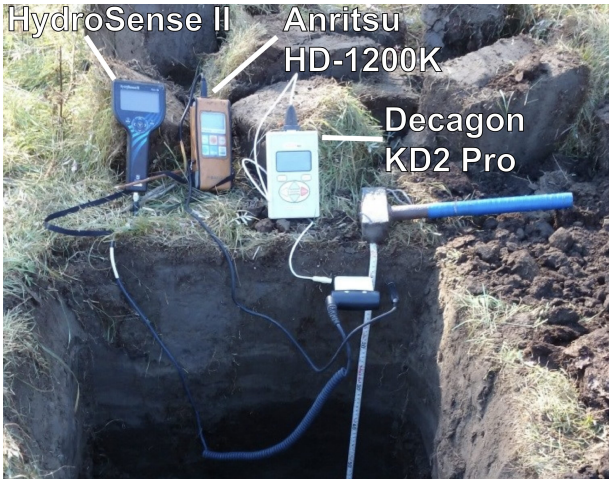

Figure D3. Thermal instruments of Anritsu (thermometer) and a Decagon Devices KD2 Pro installed in a soil pit. Volumetric water content was determined alongside using a Campbell HydroSense I or II (here II).

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Figure D4. Soil pit "pit-2017-RB-2" with metal cylinders for soil soil sampling. Old radiocarbon ages of topsoil (upper two sampling cylinders) were determined while subsequent samples show older ages with depth.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Open Access Earth System Science Data Discussions

Author contributions. EP wrote the first draft, curated the data, and developed the scripts, EP and EL produced the figures, EP, CG, ASe, FB, ASa, PK, KD, KB, and IK took measurements and produced data in the field, EP, CG, PV, AC, and CO analyzed the ground temperature data, CG, ASe, ASa, PK, KB, and IK analyzed thaw depths, CG, and PK analyzed soil water content measurements, CG, ASe, FB, and DF analyzed the soil samples regarding hydraulic and thermal properties, CH calculated the radiocarbon ages, AN, CO, ASe, and FB analyzed the water chemistry data, ASe obtained UAV imagery and produced the DSM, EM, and AF organized and provided logistics and funding, all authors contributed to the editing of the manuscript.

Competing interests. The authors declare no competing interests.

Acknowledgements. This paper is dedicated to Dr. Christophe Grenier who initialized the Syrdakh field site as a multi-disciplinary field site, and who is deeply missed. We thank the IPSL-EUR postdoc initiative that supported the work on the field site. ASe and FB acknowledge support through public funds received through the French National Research Agency (ANR) in the framework of the Make Our Planet Great Again (MOPGA) ANR project (ANR-17-MPGA-0014), ASe further received support through the PRISMARCTYC ANR project (ANR-21-SOIL-0003) of the program "Investissements d'Avenir"and AC acknowledges support from the GREENFEEDBACK (Greenhouse gas fluxes and Earth system feedbacks) project funded by the European Union's HORIZON Research and Innovation programme under grant agreement number 101056921.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

References

Aas, K. S., Martin, L., Nitzbon, J., Langer, M., Boike, J., Lee, H., Berntsen, T. K., and Westermann, S.: Thaw processes in ice-rich permafrost landscapes represented with laterally coupled tiles in a land surface model, Cryosphere, 13, 591–609, https://doi.org/10.5194/tc-13-591-2019, 2019.

Akima, H. and Gebhardt, A.: akima: Interpolation of Irregularly and Regularly Spaced Data, https://CRAN.R-project.org/package=akima, r package version 0.6-3.4, 2022.

Amante, C. and Eakins, B. W.: ETOPO1 arc-minute global relief model: procedures, data sources and analysis, Tech. rep., NOAA, National Geophysical Data Center, 2009.

Biskaborn, B. K., Smith, S. L., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D. A., Schoeneich, P., Romanovsky, V. E., Lewkowicz, A. G., Abramov, A., et al.: Permafrost is warming at a global scale, Nature communications, 10, 264, 2019.

Boike, J., Nitzbon, J., Anders, K., Grigoriev, M., Bolshiyanov, D., Langer, M., Lange, S., Bornemann, N., Morgenstern, A., Schreiber, P., Wille, C., Chadburn, S., Gouttevin, I., Burke, E., and Kutzbach, L.: A 16-year record (2002-2017) of permafrost, active-layer, and meteorological conditions at the Samoylov Island Arctic permafrost research site, Lena River delta, northern Siberia: An opportunity to validate remote-sensing data and land surface, snow, and permafrost models, Earth System Science Data, 11, 261–299, https://doi.org/10.5194/essd-11-261-2019, 2019.

Desyatkin, R., Fedorov, A., Desyatkin, A., and Konstantinov, P.: Air temperature changes and their impact on permafrost ecosystems in Eastern Siberia, Thermal Science, 19, 351–360, https://doi.org/10.2298/TSCI150320102D, 2015.

Desyatkin, R. V., Lessovaia, S. N., Okoneshnikova, M. V., and Ivanova, A. Z.: Cryosols from Tundra and Taiga Zones of Yakutia: Properties, Clay Mineralogy, and Problems of Classification, Eurasian Soil Science, 54, 1783–1794, https://doi.org/10.1134/S1064229321120048, 2021.

Fedorov, A., Ivanova, R., Park, H., Hiyama, T., and Iijima, Y.: Recent air temperature changes in the permafrost landscapes of northeastern Eurasia, Polar Science, 8, 114–128, https://doi.org/https://doi.org/10.1016/j.polar.2014.02.001, special Issue: The Third International Symposium on the Arctic Research (ISAR - 3), 2014.

Fedorov, A. N. and Konstantinov, P. Y.: Recent changes in ground temperature and the effect on permafrost landscapes in Central Yakutia, in: Proceedings of the Ninth International Conference on Permafrost, edited by Kane, D. L. and Hinkel, K. M., pp. 433–438, Institute of Northern Engineering, University of Alaska Fairbanks, Fairbanks, Alaska, 2008.

French, H. M.: Thermokarst, chap. 8, pp. 169–191, Wiley-Blackwell, https://www.wiley.com/en-us/The+Periglacial+Environment%2C+4th+Edition-p-9781119132813, 2017.

Gautier, E., Dépret, T., Costard, F., Virmoux, C., Fedorov, A., Grancher, D., Konstantinov, P., Brunstein, D., Cavero, J., Costard, F., Virmoux, C., Fedorov, A., Konstantinov, P., Jammet, M., and Brunstein, D.: Going with the flow: Hydrologic response of middle Lena River (Siberia) to the climate variability and change, Journal of Hydrology, 557, 475–488, https://doi.org/10.1016/j.jhydrol.2017.12.034, 2018.

Gorokhov, A. N. and Fedorov, A. N.: Current Trends in Climate Change in Yakutia, Geography and Natural Resources, 39, 153–161, https://doi.org/10.1134/s1875372818020087, 2018.

Grenier, C., Anbergen, H., Bense, V., Chanzy, Q., Coon, E., Collier, N., Costard, F., Ferry, M., Frampton, A., Frederick, J., Gonçalvès, J., Holmén, J., Jost, A., Kokh, S., Kurylyk, B., McKenzie, J., Molson, J., Mouche, E., Orgogozo, L., Pannetier, R., Rivière, A., Roux, N., Rühaak, W., Scheidegger, J., Selroos, J. O., Therrien, R., Vidstrand, P., and Voss, C.: Groundwater flow and heat transport for sys-

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

tems undergoing freeze-thaw: Intercomparison of numerical simulators for 2D test cases, Advances in Water Resources, 114, 196–218, https://doi.org/10.1016/j.advwatres.2018.02.001, 2018.

Günther, T. and Rücker, C.: Boundless electrical resistivity tomography (BERT) v. 2.0 open access software for advanced and flexible imaging, C. Herve, A. Adler and B. Lionhart (Eds.), 100, 177–180, 2012.

Hatté, C., Arnold, M., Dapoigny, A., Daux, V., Delibrias, G., Du Boisgueheneuc, D., Fontugne, M., Gauthier, C., Guillier, M.-T., Jacob, J., et al.: Radiocarbon dating on ECHOMICADAS, LSCE, Gif-Sur-Yvette, France: new and updated chemical procedures, Radiocarbon, 66, 1166–1181, 2024.

Hope, C. and Schaefer, K.: Economic impacts of carbon dioxide and methane released from thawing permafrost, Nature Climate Change, 6, 56–59, https://doi.org/10.1038/nclimate2807, 2016.

Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J. W., Schuur, E. A. G., Ping, C.-L., Schirrmeister, L., Grosse, G., Michaelson, G. J., Koven, C. D., O'Donnell, J. A., Elberling, B., Mishra, U., Camill, P., Yu, Z., Palmtag, J., and Kuhry, P.: Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps, Biogeosciences, 11, 6573–6593, https://doi.org/10.5194/bg-11-6573-2014, 2014.

Hughes-Allen, L., Bouchard, F., Séjourné, A., and Gandois, L.: Limnological properties of thermokarst lakes in Central Yakutia sampled between 2018-2019, https://doi.org/10.1594/PANGAEA.919907, 2020.

Hughes-Allen, L., Bouchard, F., Laurion, I., Séjourné, A., Marlin, C., Hatté, C., Costard, F., Fedorov, A., and Desyatkin, A.: Seasonal patterns in greenhouse gas emissions from thermokarst lakes in Central Yakutia (Eastern Siberia), Limnology and Oceanography, 66, S98–S116, https://doi.org/https://doi.org/10.1002/lno.11665, 2021.

Ichiyanagi, K., Sugimoto, A., Numaguti, A., Kurita, N., Ishii, Y., and Ohata, T.: Seasonal variation in stable isotopic composition of alas lake water near Yakutsk, Eastern Siberia, Geochemical Journal, 37, 519–530, 2003.

Iijima, Y., Fedorov, A. N., Park, H., Suzuki, K., Yabuki, H., Maximov, T. C., and Ohata, T.: Abrupt increases in soil temperatures following increased precipitation in a permafrost region, central Lena River basin, Russia, Permafrost and Periglacial Processes, 21, 30–41, https://doi.org/https://doi.org/10.1002/ppp.662, 2010.

Ivanov, M.: Cryogenic composition of quaternary depositions of Lena-Aldan depression, 1984.

Jan, A., Coon, E. T., and Painter, S. L.: Evaluating integrated surface/subsurface permafrost thermal hydrology models in ATS (v0.88) against observations from a polygonal tundra site, Geoscientific Model Development, 13, 2259–2276, https://doi.org/10.5194/gmd-13-2259-2020, 2020.

Kachinsky, A. N.: Physics of soil, Moscow State University, Moscow, 1968 [in Russian].

Koven, C. D., Riley, W. J., and Stern, A.: Analysis of permafrost thermal dynamics and response to climate change in the CMIP5 earth system models, Journal of Climate, 26, 1877–1900, https://doi.org/10.1175/JCLI-D-12-00228.1, 2013.

Kurylyk, B. L. and Walvoord, M. A.: Permafrost Hydrogeology, in: Arctic Hydrology, Permafrost and Ecosystems, edited by Yang, D. and Kane, D., pp. 493–523, Springer Nature Switzerland AG, Cham, 1 edn., 2021.

Léger, E., Saintenoy, A., Grenier, C., Séjourné, A., Pohl, E., Bouchard, F., Pessel, M., Bazhin, K., Danilov, K., Costard, F., Mugler, C., Fedorov, A., Khristoforov, I., and Konstantinov, P.: Comparing Thermal Regime Stages along a Small Yakutian Fluvial Valley with Point Scale Measurements, Thermal Modeling, and Near Surface Geophysics, Remote Sensing, 15, 2524, https://doi.org/10.3390/rs15102524, 2023.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Liu, W., Fortier, R., Molson, J., and Lemieux, J. M.: Three-Dimensional Numerical Modeling of Cryo-Hydrogeological Processes in a River-Talik System in a Continuous Permafrost Environment, Water Resources Research, 58, https://doi.org/10.1029/2021WR031630, 2022.

Lütjen, M., Overduin, P. P., Juhls, B., Boike, J., Morgenstern, A., and Meyer, H.: Drivers of winter ice formation on Arctic water bodies in the Lena Delta, Siberia, Arctic, Antarctic, and Alpine Research, 56, 2350 546, https://doi.org/10.1080/15230430.2024.2350546, 2024.

Meredith, M., Sommerkorn, M., Cassotta, S., Derksen, C., Ekaykin, A., Hollowed, A., Kofinas, G., Mackintosh, A., Muelbert, M., Melbourne-Thomas, J., Ottersen, G., Pritchard, H., and Schuur, E.: Chapter 3: Polar Regions, pp. 3–1–3–173, Intergovernmental Panel on Climate Change, 2019.

Miller, B. A. and Schaetzl, R. J.: Precision of soil particle size analysis using laser diffractometry, Soil Science Society of America Journal, 76, 1719–1727, https://doi.org/10.2136/sssaj2011.0303, 2012.

Miner, K. R., Turetsky, M. R., Malina, E., Bartsch, A., Tamminen, J., McGuire, A. D., Fix, A., Sweeney, C., Elder, C. D., and Miller, C. E.: Permafrost carbon emissions in a changing Arctic, Nature Reviews Earth and Environment, 3, 55–67, https://doi.org/10.1038/s43017-021-00230-3, 2022.

Mishra, U., Hugelius, G., Shelef, E., Yang, Y., Strauss, J., Lupachev, A., Harden, J. W., Jastrow, J. D., Ping, C. L., Riley, W. J., Schuur, E. A., Matamala, R., Siewert, M., Nave, L. E., Koven, C. D., Fuchs, M., Palmtag, J., Kuhry, P., Treat, C. C., Zubrzycki, S., Hoffman, F. M., Elberling, B., Camill, P., Veremeeva, A., and Orr, A.: Spatial heterogeneity and environmental predictors of permafrost region soil organic carbon stocks, Science Advances, 7, 5236–5260, https://doi.org/10.1126/sciadv.aaz5236, 2021.

Morgenstern, A., Grosse, G., Günther, F., Fedorova, I., and Schirrmeister, L.: Spatial analyses of thermokarst lakes and basins in Yedoma landscapes of the Lena Delta, Cryosphere, 5, 849–867, https://doi.org/10.5194/tc-5-849-2011, 2011.

NOAA National Geophysical Data Center: ETOPO1 1 Arc-Minute Global Relief Model, https://doi.org/http://dx.doi.org/10.7289/V5C8276M, 2009.

Pohl, E., Grenier, C., Séjourné, A., Bouchard, F., Leger, E., Saintenoy, A., Konstantinov, P., Cuynet, A., Ottlé, C., Hatté, C., Noret, A., Danilov, K., Bazhin, K., Khristoforov, I., Fortier, D., Fedorov, A., and Mouche, E.: Thermo-hydrological observatory in a permafrost river valley landscape in Syrdakh, Central Yakutia, https://doi.org/10.5281/zenodo.14619854, 2025.

QGIS Development Team: QGIS Geographic Information System, QGIS Association, https://www.qgis.org, 2024.

Raspisaniye Pogodi Ltd., St. Petersburg, R.: Reliable Prognosis, http://rp5.ru/metar.php?metar=UEEE&lang=en, 2004.

Ruff, M., Fahrni, S., Gäggeler, H. W., Hajdas, I., Suter, M., Synal, H.-A., Szidat, S., and Wacker, L.: On-line radiocarbon measurements of small samples using elemental analyzer and MICADAS gas ion source, Radiocarbon, 52, 1645–1656, 2010.

Schuur, E. A., McGuire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., Hugelius, G., Koven, C. D., Kuhry, P., Lawrence, D. M., Natali, S. M., Olefeldt, D., Romanovsky, V. E., Schaefer, K., Turetsky, M. R., Treat, C. C., and Vonk, J. E.: Climate change and the permafrost carbon feedback, Nature, 520, 171–179, https://doi.org/10.1038/nature14338, 2015.

Séjourné, A., Costard, F., Fedorov, A., Gargani, J., Skorve, J., Massé, M., and Mège, D.: Evolution of the banks of thermokarst lakes in Central Yakutia (Central Siberia) due to retrogressive thaw slump activity controlled by insolation, Geomorphology, 241, 31–40, 2015.

Soloviev, P. A.: Thermokarst phenomena and landforms due to frost heaving in Central Yakutia, Peryglacialny biuletyn, 23, 135–155, 1973.

Song, C., Wang, G., Mao, T., Dai, J., and Yang, D.: Linkage between permafrost distribution and river runoff changes across the Arctic and the Tibetan Plateau, Science China Earth Sciences, 63, 292–302, https://doi.org/10.1007/s11430-018-9383-6, 2020.

Standard, A.: D2216-19; Standard Test Methods for Laboratory Determination of Water (Moisture) Content of Soil and Rock by Mass, American Society for Testing and Materials: West Conshohocken, PA, USA, 2019.

https://doi.org/10.5194/essd-2025-134
Preprint. Discussion started: 14 October 2025

Steinert, N. J., González-Rouco, J. F., de Vrese, P., García-Bustamante, E., Hagemann, S., Melo-Aguilar, C., Jungclaus, J. H., and Lorenz, S. J.: Increasing the Depth of a Land Surface Model. Part II: Temperature Sensitivity to Improved Subsurface Thermodynamics and Associated Permafrost Response, Journal of Hydrometeorology, 22, 3231–3254, https://doi.org/10.1175/JHM-D-21-0023.1, 2021.

Strauss, J., Laboor, S., Schirrmeister, L., Fedorov, A. N., Fortier, D., Froese, D., Fuchs, M., Günther, F., Grigoriev, M., Harden, J., Hugelius, G., Jongejans, L. L., Kanevskiy, M., Kholodov, A., Kunitsky, V., Kraev, G., Lozhkin, A., Rivkina, E., Shur, Y., Siegert, C., Spektor, V., Streletskaya, I., Ulrich, M., Vartanyan, S., Veremeeva, A., Anthony, K. W., Wetterich, S., Zimov, N., and Grosse, G.: Circum-Arctic Map of the Yedoma Permafrost Domain, Frontiers in Earth Science, 9, https://doi.org/10.3389/feart.2021.758360, 2021.

Synal, H.-A., Stocker, M., and Suter, M.: MICADAS: A new compact radiocarbon AMS system, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 259, 7–13, 2007.

Tananaev, N. and Lotsari, E.: Defrosting northern catchments: Fluvial effects of permafrost degradation, Earth-Science Reviews, 228, 103 996, https://doi.org/10.1016/j.earscirev.2022.103996, 2022.

Thil, F., Tisnérat-Laborde, N., Hatté, C., Kader, E., Noury, C., Paterne, M., Phouybanhdyt, B., and Wacker, L.: 14C microsample analysis with ECHoMICADAS facilities: current state of play, Radiocarbon, 66, 1379–1394, 2024.

Tisnérat-Laborde, N., Thil, F., Synal, H., Cersoy, S., Hatté, C., Gauthier, C., Massault, M., Michelot, J., Noret, A., Siani, G., et al.: ECHoMICADAS: A new compact AMS system to measuring 14C for Environment, Climate and Human Sciences, in: 22nd International Radiocarbon Conference, Dakar, Senegal, 15-20 November, 2015.

Vonder Mühll, D., Nötzli, J., Makowski, K., and Delaloye, R.: Permafrost in Switzerland 2000/2001 and 2001/2002, Glaciological Report (Permafrost), https://doi.org/http://doi.org/10.13093/permos-rep-2004-2-3, 2004.

Wacker, L., Bonani, G., Friedrich, M., Hajdas, I., Kromer, B., Němec, M., Ruff, M., Suter, M., Synal, H.-A., and Vockenhuber, C.: MICADAS: Routine and High-Precision Radiocarbon Dating, Radiocarbon, 52, 252–262, https://doi.org/10.1017/S0033822200045288, 2010a.

Wacker, L., Christl, M., and Synal, H.-A.: Bats: A new tool for AMS data reduction, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 268, 976–979, 2010b.

Wacker, L., Němec, M., and Bourquin, J.: A revolutionary graphitisation system: fully automated, compact and simple, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 268, 931–934, 2010c.

Walvoord, M. A. and Kurylyk, B. L.: Hydrologic Impacts of Thawing Permafrost—A Review, Vadose Zone Journal, 15, 0, https://doi.org/10.2136/vzj2016.01.0010, 2016.

Westermann, S., Langer, M., Boike, J., Heikenfeld, M., Peter, M., Etzelmüller, B., and Krinner, G.: Simulating the thermal regime and thaw processes of ice-rich permafrost ground with the land-surface model CryoGrid 3, Geoscientific Model Development, 9, 523–546, https://doi.org/10.5194/gmd-9-523-2016, 2016.

Westermann, S., Peter, M., Langer, M., Schwamborn, G., Schirrmeister, L., and Boike, J.: Transient modeling of the ground thermal conditions using satellite data in the Lena River delta, Siberia, The Cryosphere, 11, 1441–1463, https://doi.org/10.5194/tc-11-1441-2017, 2017.

Zhang, Y. and Schaap, M. G.: Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3), Journal of Hydrology, 547, 39–53, https://doi.org/10.1016/j.jhydrol.2017.01.004, 2017.

Zhirkov, A., Permyakov, P., Wen, Z., and Kirillin, A.: Influence of rainfall changes on the temperature regime of permafrost in Central Yakutia, Land, 10, 1230, https://doi.org/https://doi.org/10.3390/land10111230, 2021.