Contact: Julia Boike **27 August 2019**

Julia.Boike@awi.de

This dataset has been published in the following publication. If you use this dataset, please cite this publication:

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 Syst. Sci. Data, 11, 261-299, https://doi.org/10.5194/essd-11-261-2019, 2019.

New dataset release in august 2019 with following updates:

- additional level1 data set for 2018
- precipitation (liquid) correction for 2010-2017 (see corrigendum)

Table 1. Variables with column name and unit (level 1)

Variable	Column name	Units
air / snow-covered air temperature	Tair_(height in cm)	°C
relative humidity	RH_(height in cm)	%
atmospheric pressure	PA	kPa
incoming shortwave radiation	SwIn	W m^{-2}
outgoing shortwave radiation	SwOut	W m^{-2}
incoming longwave radiation	LwIn	W m^{-2}
outgoing longwave radiation	LwOut	W m^{-2}
net radiation	RadNet	W m^{-2}
wind speed	Vwind_(height in cm)	m s^{-1}
wind speed maximum	Vwind_max_(height in cm)	m s^{-1}
wind speed minimum	Vwind_min_(height in cm)	m s^{-1}
wind direction	Dirwind_(height in cm)	°
wind direction standard deviation	Dirwind_sd_(height in cm)	°
active layer thaw depth	Dal_(ID)	cm
soil/permafrost temperature	Ts_(depth in cm)	°C
soil bulk electrical conductivity	Cond_(depth in cm)	S m^{-1}
soil dielectric number	E2_(depth in cm)	–

soil volumetric liquid water content	Vwc_(depth in cm)	–
ground heat flux	G_(depth in cm)	W m^{-2}
precipitation (liquid)	Prec	mm

Overview of all variables published as time series. Some variables have _center, _rim and _slope as location index (see chapter 3.2. of the ESSD publication Boike et al., 2018). Additional Level 2 data is published for the variables air temperature, relative humidity, precipitation, wind speed and direction, net radiation, soil temperatures, and soil volumetric liquid water content, which is indicated by "_lv2" in the column name. If an air temperature sensor is covered by snow and thus measures not ambient air temperature, a flag 8 in the data indicates this.

The grid points of the CALM site are classified according to micro-topographic characteristics: (1) "center" refers to the depressed central region of the polygonal structures; (2) "slope" refers to the transition zone between the centers and the rims; (3) "rim_high" refers to polygon rims which are significantly elevated above the remaining terrain; (4) "rim_flat" refers to shallower polygon rims between the centers; (5) "crack" are measurement points within a frost crack which are typically located within the rims.

Table 2. Quality flag scale

Flag	Meaning	Description
ONL	Online data	Data from online stations, daily download, used for online status check
RAW	Raw data	Base data from offline stations, 3-monthly backup of online data, used for maintenance check in the field
LV0	Level 0	Standardized format with data in equal time steps (UTC), filled with NA for data gaps
LV1	Level 1	Quality-controlled data including flags; quality control includes maintenance periods, physical plausibility, spike/constant value detection, sensor drifts, and snow on sensor detection
LV2	Level 2	Modified data compiled for special purposes such as combined data series from multiple sensors and gap-filled data
0	Good data	All quality tests passed
1	No data	Missing value
2	System error	System failure led to corrupted data, e.g. due to power failure, sensors being removed from their proper location, broken or damaged sensors, or the data logger saving error codes
3	Maintenance	Values influenced by the installation, calibration, and cleaning of sensors or programming of the data logger; information from field protocols of engineers

Description quality control for data flags. Most flags are run automatically, few are done manually (3: maintenance, 6: plausibility).

Table 3. Meteorological variables and soil temperatures (Level 2)

1998–2002	2002–current	Level 2 (1998–current)
Tair_200	Tair_a_200	Tair_a_lv2_200
Tair_50	Tair_a_50	Tair_a_lv2_50
RH_200	RH_a_200	RH_a_lv2_200
RH_50	RH_a_50	RH_a_lv2_50
Prec	Prec	Prec_lv2
Vwind_300	Vwind_300	Vwind_lv2_300
Dirwind_300	Dirwind_300	Dirwind_lv2_300
RadNet	RadNet	RadNet_lv2
Ts_center_7	Ts_center_5	Ts_center_lv2_5
Ts_center_13	Ts_center_10	Ts_center_lv2_10
Ts_center_32	Ts_center_30	Ts_center_lv2_30
Ts_center_42	Ts_center_40	Ts_center_lv2_40
Ts_rim_9	Ts_rim_6	Ts_rim_lv2_6
Ts_rim_15	Ts_rim_16	Ts_rim_lv2_16
Ts_rim_28	Ts_rim_27	Ts_rim_lv2_27
Ts_rim_37	Ts_rim_38	Ts_rim_lv2_38
Ts_rim_47	Ts_rim_51	Ts_rim_lv2_51

The data has not been changed in any way (e.g. no interpolation or regression). The choice of data series was based on the following criteria: a) same surface type, in particular for sensors close to the surface, b) no obvious sensor drift, zero curtain in autumn does not deviate more than ±0.3 °C from 0 °C, c) similar temperature level and variation in winter and summer.