Contact: Julia Boike **January 19, 2018**

Julia.Boike@awi.de

These dataset and metadata have been published in the following publication. If you use this dataset, please cite the final publication of:

Boike, J., Juszak, I., Lange, S., Chadburn, S., Burke, E., Overduin, P. P., Roth, K., Ippisch, O., Bornemann, N., Stern, L., Gouttevin, I., Hauber, E., and Westermann, S.: A 20-year record (1998–2017) of permafrost, active layer, and meteorological conditions at a High Arctic permafrost research site (Bayelva, Spitsbergen): an opportunity to validate remote sensing data and land surface, snow, and permafrost models, Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-100, in review, 2017.

Dataset Bayelva

The following tables give details on the Bayelva data.

Table 1. Variables with column name and unit

variable	column name	unit
air/snow temperature	Tair_(height in cm)	°C
relative humidity	RH_(height in cm)	%
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 direction	Dirwind_(height in cm)	°
wind direction standard deviation	Dirwind_sd_(height in cm)	°
soil/permafrost temperature	Ts_(depth in cm)	°C
soil bulk electrical conductivity	Cond_(depth in cm)	S m^{-1}
dielectric number	E2_(depth in cm)	–
soil volumetric liquid water content	Vwc_(depth in cm)	–
ground heat flux	G	W m^{-2}
precipitation (liquid)	Prec	mm
snow depth	Dsn	m

Overview of all variables provided as time series. The variables electric conductivity, dielectric number, soil temperature, and volumetric soil volumetric liquid water content include a second number or letter in the column name. In these cases, the first number is the distance in the 2D profile and the second is the depth. The single letters [a, b, c, d] refer to the different 1D profiles. Vertically oriented probes are marked with letter [v]. Additional Level 2 data is provided for the variables snow depth, soil temperature, and volumetric soil liquid water content, which is indicated by "_lv2" in the column names. If an air temperature sensor is covered by snow and thus measures snow temperature, this is indicated by flag 8 in the data.

Table 2. Quality flag scale

flag	meaning	description
0	Good data	All quality tests passed
1	No data	Missing value
2	System error	System failure led to corrupted data, e.g. when the power supply broke down, sensors were removed from their proper location, sensors broke or the data logger saved 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
4	Physical limits	Values outside the physically possible or likely limits, e.g. relative humidity should be in a range of 0-100%
5	Gradient	Values unlikely because of prolonged constant periods or high/low spikes; test within each single series
6	Plausibility	Values unlikely in comparison with other series or for a given time of the year; flagged manually by engineers
7	Decreased accuracy	Values with decreased sensor accuracy, e.g. identified when thawing soil does not have a temperature of 0°C
8	Snow covered	Good data, but the sensor is snow covered

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

Table 3. Soil temperatures and volumetric water content level2 product

1998-2012	2009-current	LV2
Ts_a_145_6	Ts_c_1	Ts_lv2_4
Ts_a_169_9	Ts_c_11	Ts_lv2_10
Ts_a_124_25	Ts_c_21	Ts_lv2_23
Ts_a_145_40	Ts_c_37	Ts_lv2_39
Ts_a_145_62	Ts_c_55	Ts_lv2_59
Ts_a_151_76	Ts_c_71	Ts_lv2_74
Ts_a_145_99	Ts_c_89	Ts_lv2_94
Ts_a_135_125_v	Ts_c_141	Ts_lv2_133
Vwc_140_6	Vwc_c_1	Vwc_lv2_4
Vwc_174_8	Vwc_c_11	Vwc_lv2_10
Vwc_115_22	Vwc_c_21	Vwc_lv2_22
Vwc_140_39	Vwc_c_37	Vwc_lv2_38
Vwc_140_62	Vwc_c_55	Vwc_lv2_59
Vwc_146_76	Vwc_c_71	Vwc_lv2_74
Vwc_95_99	Vwc_c_89	Vwc_lv2_94

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 (mud boil), in particular for sensors close to the surface; this is given for the soil sensors with x (first value followed by depth), b) R^2~1 and regression slope~1 and offset~0°C during the overlapping period 2009-2012 between old (1998-2009) and new station (2009-current), c) no obvious sensor drift, zero curtain in autumn does not deviate more than 0.3 °C from 0°C d) similar temperature level and variation in winter and summer during the overlapping period 2009-2012 between old and new station.