1. The lack of focus criticized by reviewer nr. 2 has still not been addressed thoroughly. Please sharpen and distill your own results further in light of the previous and similar work done by Park et al. (2016) and André et al. (2015). What does your paper provide in terms of substantial progress beyond the current scientific understanding? What are the necessary implications? Please rewrite especially abstract and conclusions accordingly.

   Reply: Both, Park et al. and Andre et al. do not address MAGT with microwave records. The latter only derived LST for snow free months (day 180 - 250). A filling of all snow covered months with re-analyses data would be required to eventually obtain a full annual record which can be then used by a model. Park et al. does estimate permafrost extent by a threshold, but not MAGT. Permafrost extent is only a partial aspect in our study, a possible subsequent analyses of MAGT retrieval.
   We have now adjusted the abstract, conclusions as well as state of the art and title to better highlight the focus of our study on MAGT. In addition, the methods and results sections have been restructured (change of order): first the MAGT estimation as primary analyses and then the permafrost extent as additional analyses. See track changes version. (Note that due to restructuring, further changes within the moved sections cannot be tracked separately. They are however limited to changes listed below and change of tenses.)

   E.g. the text related to Andre et al. 2015 and Park et al has been adjusted:
   Old: Improved temperature records can be obtained for the snow free period from combination with passive microwave data (Andre et al. 2015) but they need to be complemented with re-analysis data for the remaining year in order to derive the required indices.
   A different and simpler approach using only satellite records has been recently described by Park et al. (2016).

   New: Improved temperature records can be obtained for the snow free period from combination with passive microwave data (Andre et al. 2015). Such records need to be, however, complemented with re-analysis data for the remaining year in case of further analyses which targets MAGT.
   An approach for permafrost extent determination without estimation of MAGT and use of satellite records only has been recently described by (Park et al. 2016).

2. The English language needs clarification, the format of the figures and tables needs improvement. I have listed a few examples below. Please go through the entire manuscript and check thoroughly.

   All below changes have been made, tenses adjusted and the entire manuscript revised for language.

Specific comments

Page 2: laps rate should be lapse rate
         Reply: changed

Page 3: "The mean annual ground temperature (MAGT) is currently only in some cases provided by the data owners within the Global Terrestrial Network on Permafrost (GTN-P) database.
This sentence is unclear- do you mean that "Only the mean annual ground temperature…."?

Are data series available, just not the means?
Please clarify the sentence.

 Reply: Yes, data are available but not the means.
 New: The mean annual ground temperature (MAGT) derived from these records is currently only in some cases provided by the data owners within the Global Terrestrial Network on Permafrost (GTN-P) database.

Page 4. ECMWF ReAnalyis should be reanalysis
 Reply: Changed

Page 4: "The most frequent sensor depth is found at 5 m. Permanently installed multi-thermistor cables showing an accuracy between 0.002 - 0.1°C are most commonly used for measuring continuous ground temperatures at specific depths (Romanovsky et al., 2010)."
Please correct: Romanovsky et al 2010 report "Measurement systems currently in use generally provide an accuracy and precision of 0.1°C or better."

 Reply: changed as suggested

Page 4: which suggest es
Please correct typo.
 Reply: changed as suggested

Page 4: "Most records of North America are accompanied with meta records which suggest es sensor depth for approximation of the MAGT".
Please be specific: Do you mean approximation of the MAGT at zero annual amplitude?
Please also specify the MAGT – which depth you are referring to?
 Reply: This paragraph only refers to MAGT at zero annual amplitude. We have now moved the sentence with MAGT range to the end of the paragraph (after mentioning ZAA) and changed the above sentence to:
 New: Most records of North America are accompanied with meta records which suggest a sensor depth (closest to zero annual temperature amplitude) for approximation of the MAGT.

Page 6: Only sensors with more than 1m depth have been used as the MAGT near the surface can be much colder than at larger depth.
Awkward sentence: Do you mean sensors instrumented below 1 m depth…?

 Reply: Changed as suggested

Page 6: "To explore the dependency of the results on snow melting events, the permafrost extent estimation from ASCAT data was carried out excluding the melt days in the count of frozen days (FT) and a second analysis counting the melt days as frozen grid cells (FM).
Sentence unclear: do you mean that in the second analysis the melt days were counted as frozen days? Please clarify sentence.

 Reply: 'frozen grid cells' has been now replaced with 'frozen days'

Page 6: "In the study of Park et al. (2016) a threshold of half a year was chosen for the delineation of permafrost extent."
Threshold of half a year of "frozen days"? Please clarify.

 Reply: 'frozen days' added as suggested

Page 7: "The relationship between MAGT and frozen day of year was further examined.."
MAGT at which depth?
Overall, please specify in text and all figure captions at what depth you refer to when reporting
MAGT, including all figure captions (example Figure 6).

Reply: We have now introduced $MAGT_c$ for clarification the following sentence was inserted in
the methods section:
The MAGT at coldest sensor depth is in the following referred to as $MAGT_c$

Page 7: "R2 between MAGT from meta records and MAGT at coldest sensor depth is 0.994 (Fig. 1)."
MAGT from meta records- what source and which depth? Be specific and also add this information in
caption of figure 1.

Reply: We added now 'GTN-P' and '(at or close to zero annual amplitude)' in all cases. The
assumption is that GTN-P meta records (provided by the local experts) refers to ZAA. See also
data section.

Page 10: The results based on SSM/I are mostly warmer in the continuous zone.
Sentence unclear- how can results be warmer?

Reply: 'Results' have been replaced with modelled $MAGT_c$

Page 10: "The model result for ASCAT FM is about two degree colder than what is obtained from the
in situ records in the first case".
Please clarify sentence.

New: The model result for ASCAT FM is about two degree colder than in the case of exclusion of
melting snow. In contrast, $MAGT_c$ increases for Alaska.

Page 11: "…order of the actual extent but the error of commission is relatively large".
Awkward wording.

Reply: We have now split this sentence in two parts:
The extent of permafrost estimated with the initial threshold is in the order of the actual extent.
The error of commission is, however, relatively large.

Page 11: …"Westermann et al. (2015) reported a model accuracy of 2.5°C."
Model accuracy of what?

Added: .. for MAGT

Page 12: "…200 days corresponds to approximately 0.5°C in case of ASCAT FT."
0.5°C of model derived MAGT? Which depth?
Added: modelled $MAGT_c$

Added: for permafrost extent retrieval

Page 14: Please complete reference to (e.g. Smith and Riseborough).

Reply: Year added

Overall format issues
-Check correct spelling: degree C and degree Celsius is used throughout the paper, please unify.
Reply: All changed to C

-Check the tenses throughout the paper- it switches between present and past throughout

Reply: We checked the tenses throughout the paper to ensure a consistent usage of past/present tense and adapted the text where it was necessary. As this journal does not provide specific instruction concerning the usage of tenses in the guidelines and house standards (https://www.the-cryosphere.net/for_authors/manuscript_preparation.html) we followed the following generally agreed upon structure of tense usage in scientific manuscripts:

- Abstract: Past tense when reporting the work and results, present tense for general statements and implications of the results
- Introduction: Past tense and present perfect for describing related works and previous studies, present tense for general statements, information
- Datasets: Past tense when reporting on previous studies and related works, present tense when talking about the characteristics of the used data sets
- Methods: Past tense to describe the followed methodology and workflow, present tense to describe general facts about used data sets etc.
- Results: Past tense when reporting results, present tense when directly commenting on figures and tables
- Discussion: Past tense to briefly summarize findings, present tense to interpret and discuss findings
- Conclusion: Past tense to refer to the research done in this study, present tense to talk about implications of the results

See also:

- https://www.nature.com/scitable/topicpage/effective-writing-13815989
- https://www.publishingcampus.elsevier.com/websites/elsevier_publishingcampus/files/Skills%20training/Elements_of_Style.pdf
- http://services.unimelb.edu.au/__data/assets/pdf_file/0009/471294/Using_tenses_in_scientific_writing_Update_051112.pdf
- https://wordvice.com/which-verb-tense-to-use-in-a-research-paper/

-Change all figures to digital decimals (not commas, used in Figure 14 for temperature and standard deviation)
   Reply: Commas in figures 6 and 14 have been replaced

-Provide all information required for understanding the data presented in the figures.
Examples: Figure 1. MAGT at coldest sensor depth- please add the source of data
   Reply: Added: based on GTN-P borehole records.

-Improve quality of Tables and Figures (example Table 3, 4, Figure 7)
   Reply: Tables have been converted to Latex-Format and Figure 7 improved.

~~Potential permafrost distribution and~~ Circumpolar patterns of potential mean annual ground ~~temperatures~~ temperature based on surface state obtained from microwave satellite data

Christine Kroisleitner[1,2], Annett Bartsch[1,2,3], and Helena Bergstedt[4,2]

[1]Zentralanstalt für Meteorologie und Geodynamik, Vienna, Austria
[2]Austrian Polar Research Institute, Vienna, Austria
[3]b.geos, Korneuburg, Austria
[4]Department of Geoinformatics - Z_GIS, University Salzburg, 5020 Salzburg, Austria

Correspondence to: Annett Bartsch (annett.bartsch@zamg.ac.at)

Abstract.

Gap filling is required for temporally and spatially consistent records of land surface temperature from satellite data due to clouds or snow cover. Land surface state, frozen versus unfrozen conditions, can be, however, captured globally with satellite data obtained by microwave sensors. ~~It has been previously suggested that the~~ The number of frozen days per year ~~can be used to determine permafrost extent~~ has been previously proposed to be used for permafrost extent determination based on a ~~threshold method. We have tested the validity of this simplification~~ single threshold of freeze duration. This suggests an underlying relationship between number of frozen days and mean annual ground temperature (MAGT). We tested this hypothesis for the northern hemisphere north of 50°~~and identified regions where more advanced modelling is essential. Coarse N using~~ coarse spatial resolution microwave satellite data (Metop ~~ASCAT and~~ Advanced Scatterometer - ASCAT - and Special Sensor Microwave Imager - SSM/I; 12.5km and 25km nominal resolution; 2007 - 2012) which provide the necessary temporal sampling~~have been used. Borehole measurements across the entire domain as well as a conventional permafrost zone map have been chosen for assessment with four different methods: 1) quantification of map area overlap, 2) Kendall's τ test with in situ data, 3) in situ data confusion matrix and 4) determination of an empirical relationship between mean annual ground temperature (MAGT) and number of frozen days. The MAGT~~. The MAGT from GTN-P (Global Terrestrial Network - Permafrost) borehole records at coldest sensor depth ~~has been~~ was tested for validity in order to build a comprehensive in situ dataset for calibration and validation ~~. Mean annual ground temperature can be obtained from ASCAT~~ and was eventually applied. Results are discussed with respect to snow water equivalent, soil properties, land cover and permafrost type. The obtained temperature maps were classified for permafrost extent and compared to alternative approaches.

An R^2 of 0.99 was found for correlation between and MAGT at zero annual amplitude provided in GTN-P meta data and MAGT at coldest sensor depth. The latter could be obtained with an RMSE of 2.2°C ~~compared to~~ from ASCAT and 2.5°C ~~in case of~~ from SSM/I ~~. This is weaker than what can be obtained using more complex models at medium resolution (appr. 1km), but comparably good taking into account the very coarse spatial resolution of 25km.~~ surface state records using a linear model. The average deviation within the validation period is less than ~~one degree Celsius~~ 1°C at locations without glaciers

and coastlines within the resolution cell ~~. The ASCATdataset also allows to distinguish between melting and dry snow~~in case of ASCAT. The exclusion of snow melt days ~~leads~~ (available for ASCAT) led to better results~~using the selected evaluation method (Pearson correlation) and zonal assessment (by region as well as snow water equivalent classes)~~. This suggests that soil warming under wet snow cover needs to be accounted for in this context. Specifically Scandinavia and Western Russia

5 are affected. ~~Further regions which show high differences between the tested methods of potential permafrost extent retrieval are e.g. Central Russia and Central Siberia.~~ In addition, MAGT ~~is~~ at coldest sensor depth was overestimated in areas with a certain amount of organic material and in areas of cold permafrost. ~~In conclusion, microwave satellite records can~~ A region which showed high differences in derived potential permafrost extent is e.g. central Siberia. We show that microwave satellite derived surface state records can not only provide an estimation of permafrost extent ~~and even~~ but also MAGT without the need

10 for gap filling~~, but results need to be treated with care~~. This applies specifically to C-band. The deviations between the tested datasets, their spatial patterns as well as in relation to environmental conditions, revealed areas which need special attention for modelling of MAGT.

1 Introduction

Permafrost covers large parts of the Earth's surface and is defined as ground that remains at or below 0°C for at least two consecutive years. The impact of climate change on the Arctic, as well as other permafrost dominated environments, is thought to be more severe compared to the rest of the world (Schuur et al., 2015; National research Council, 2013). Warming and with that thawing of permafrost impacts multiple environmental processes ranging from surface and sub-surface hydrology (O'Donnell et al., 2012; Woo et al., 2008), ecological changes (Schuur et al., 2008) to processes like carbon exchange (Hayes et al., 2014). Knowledge of the ~~temperature of the ground, the~~ extent of permafrost and possible changes in its distribution are therefore crucial for climate modelling and prediction (Cheng and Wu, 2007; O'Connor et al., 2010). So far, the exact extend of permafrost is unknown and has only been approximated (Zhang et al., 2008; Nguyen et al., 2009).

A range of approaches exist for ~~modelling of mean annual ground temperature (MAGT) and subsequent~~ permafrost extent determination. They vary in complexity, accuracy and their parameterization differs depending on the size of the area of interest and data availability. Different temperature data sources are employed for the approximation. E.g. Gruber (2012) ~~uses~~ _have used_ re-analyses records, which are based on spatially interpolated in situ air temperature data (mean annual air temperature) together with elevation data (~~laps~~ _lapse_ rate), to generate a global map of permafrost probabilities. Temperature records from re-analyses and land surface temperature from satellites together with basic vegetation information and satellite derived snow properties have been investigated for modelling larger areas by Westermann et al. (2015)~~. Satellite~~ _as satellite_ data alone do not provide the required spatially and temporally consistent temperature values. Cloud cover is problematic when using thermal infrared. Clear sky bias is an additional problem (Soliman et al., 2012). Improved temperature records can be obtained for the snow free period from combination with passive microwave data (André et al., 2015)~~but they~~. _Such records_ need to be, _however,_ complemented with re-analysis data for the remaining year in ~~order to derive the required indices~~_case of further analyses which targets MAGT._

~~A different and simpler approach using only satellite records~~ _An approach for permafrost extent determination without estimation of MAGT and use of satellite records only_ has been recently described by Park et al. (2016). They ~~hypothesize~~ _have hypothesized_ that the number of frozen days per year from passive microwave satellite data (SSM/I - Special Sensor Microwave Imager) can be used as indicator for permafrost extent. A 30 year record was analyzed for trends and compared to the map of Brown et al. (1997) - often referred to as the IPA (International Permafrost Association) map - and results of the CHANGE progress model (Park et al., 2011a). A threshold of half a year of frozen days during at least two consecutive years was chosen to delineate possible permafrost areas. This value was justified with reference to Dobinski (2011), Nelson and Outcalt (1987), Saito et al. (2013) and Zhang et al. (2005). These studies ~~use~~ _have used_ measures derived from actual temperature records, specifically the use of mean annual air temperature and the concept of thawing/freezing degree days. Results showed differences between the model and microwave product, especially at the start of the record which could be due to sparse satellite data during the beginning of the chosen time period (Trofaier et al., 2017). The comparison with the IPA permafrost map revealed an overestimation of permafrost extent around 65°N. Overall, agreement regionally differed, especially over non-continuous permafrost. Luoto et al. (2004) ~~suggest~~ _suggested_ that the minimum number of frozen days is

200 for new permafrost to develop (in form of palsas) in the transition zone of Scandinavia. Additionally, local factors like a low annual mean air temperature (<1°C), water-saturated peat, moss patches and low vegetation have to be present (Seppälä, 1986; Harris, 1981). This number is considerably higher, compared to the selection of Park et al. (2016). The utility of a single threshold may, however, be the result of an underlying relationship between MAGT and number of frozen days.

5   Surface state information ~~can be also derived from active microwave sensors operating at various frequencies (Park et al., 2011b; Naeimi~~ and with that the calculation of frozen and thawed days per year, can also be derived from active microwave sensors operating at various frequencies (Park et al., 2011b; Naeimi et al., 2012; Bartsch et al., 2007; Wang et al., 2008; Zwieback et al., 2015). Frequencies are usually lower than SSM/I and especially C-band (5.3 GHz) scatterometer like the ASCAT (Advanced Scatterometer) sensor on-board the Metop satellites have already been shown to be applicable for freeze-thaw information retrieval

10 in permafrost regions by validation with near surface soil temperature from borehole records (Naeimi et al., 2012). Multiannual statistics of thaw and freeze-up timing based on these records have been e.g. applied for the retrieval of circumpolar landscape units (Bartsch et al., 2016).

The number of frozen days can be observed consistently from space. Freezing degree days, as conventionally used for MAGT retrieval, require spatial and temporal interpolation of the input data. The use of frozen days would therefore allow a

15 purely observation based assessment. This does, however, require the assumption that changes at the surface (as represented by the satellite at a certain frequency) are uniformly and linearly related to sub-ground temperatures. This neglects effects such as insulation by snow as well as varying soil thermal conductivity. The validity of the approach may therefore be limited. In cases where it is applicable, it may, however, allow the estimation of actual ground temperatures and not only extent.

A further issue is the available data for calibration and validation of datasets spanning the entire Arctic. The map of Brown

20 et al. (1997) does provide zones of permafrost occurrence which correspond to area fraction of permanently frozen ground. The actual patterns within the non-continuous zone are unknown, different sources have been used and it represents the stage of the second half of the 20th century. Results need to be therefore treated with care. It is however used in many studies for evaluation of modelling results (e.g. in Park et al., 2016; Matthes et al., 2017).

An alternative are actual ground temperature measurements. Borehole data represent, however, only point information, with

25 uneven distribution (Biskaborn et al., 2015) and provide measurements at selected depths only. The mean annual ground temperature (MAGT) derived from these records is currently only in some cases provided by the data owners within the Global Terrestrial Network on Permafrost (GTN-P) database. A practical method which allows use of all freely available data is required for circumpolar applications.

The objective of this study is to investigate the applicability of the frozen days approach based on satellite data for potential

30 permafrost extend determination as well as MAGT retrieval. Special emphasis is given on suitability of calibration and validation data, differences between microwave sensors and uncertainties with respect to environmental conditions including snow, land cover and ground ice content. Regional patterns of agreement with the map of Brown et al. (1997) as well as with in situ data are discussed.

2 Datasets

2.1 Satellite records

We used two microwave remote sensing data sets derived from globally available records and with similar classification accuracy obtained by comparison to air temperature data. They ~~are~~ were derived from sensors with different frequencies, acquisition
5 methods (active versus passive) and timing. The first was derived from the ASCAT sensor on-board the Metop satellites. The ASCAT sensor is a C-band (5.255 GHz) scatterometer (Figa-Saldaña, et al., 2002), providing almost daily coverage of the Earth's surface. The equator overpass time is 9:30 am. The surface state information (freeze/thaw) ~~has been~~ was derived from the ASCAT sensor as a surface status flag for a soil moisture product specifically post-processed for high latitudes (Paulik et al., 2014). The circumpolar data set covers the years 2007 - 2013. It was developed for permafrost monitoring and climate
10 modelling purposes (Bartsch and Seifert, 2012; Naeimi et al., 2012; Reschke et al., 2012), and covers the area above 50°N with a grid spacing of 12.5 km and an up to daily temporal resolution (Paulik et al., 2014). This includes the parameters frozen and thawed ground, temporary water (including snow melt) and frozen water/permanent ice. The surface status information was derived using a step-wise threshold algorithm based on ASCAT backscatter values and ECMWF ~~ReAnalyis~~ reanalyis data (Naeimi et al., 2012). The accuracy was assessed with in situ surface air temperature measurements from the global weather
15 station network and found to be about 82% overall (Naeimi et al., 2012). Up to 92% agreement was found for near surface temperature measurements from boreholes of GTN-P located in Siberia.

The second remotely sensed data set used in this study was derived from global daily (ascending and descending orbit) 37 GHz, vertically polarized brightness temperature observations from calibrated SMMR (scanning multi-channel microwave radiometer) and SSM/I (special sensor microwave imager) satellite sensor records by Kim et al. (2014). There have been a
20 series of those sensors carried on-board satellites from the Defense Meteorological Satellite Program. They are passive radiometric systems that measure atmospheric, ocean and terrain microwave temperature (Hollinger et al., 1990). SSM/I equatorial crossings are 6 am and 6 pm. The freeze-thaw status was analyzed globally from 1979 - 2012 by Kim et al. (2014). The data set has a nominal resolution of 25 km and covers the Arctic terrestrial drainage basin. The actual footprint size at 37GHz is 37km x 28km. The threshold approach produces two classes: frozen and non-frozen. The estimated classification accuracy
25 is approximately 85% (am passes) to 92% (pm passes) compared to in situ surface air temperature measurements from the global weather station network (Kim et al., 2012). This dataset ~~has been~~ was used by Park et al. (2016) for the assessment of permafrost extent changes.

2.1.1 ~~In situ records from boreholes~~

2.2 In situ records from boreholes

30 All borehole records above a latitude of 50° North with available time series of ground temperature were retrieved from the Global Terrestrial Network for Permafrost database (GTN-P, 2016). The network collects measurements of the Thermal State of Permafrost (TSP) in polar and mountain regions. 277 borehole sites have temperature data, single sites often

comprise more then one measurement unit or period, which leads to a total sum of 1062 ground temperature data sets (Biskaborn et al., 2015). The depth of most boreholes is less than 25 m, although the average ~~lies at~~ is 53 m. The most frequent sensor depth ~~is~~ can be found at 5 m. ~~Permanently installed multi-thermistor cables showing an accuracy between 0.002 -~~ Romanovsky et al. (2010) reported that measurement systems currently in use generally provide an accuracy and precision of

5 0.1°C ~~are most commonly used for measuring continuous ground temperatures at specific depths (Romanovsky et al., 2010)~~or better. The time series are available in hourly, daily or annual resolution and cover different time periods. Deepest sensor depths of the used data set vary between 1m and 99m. In total, 216 boreholes ~~have been considered. They represent a MAGT range from -15°C to 6°C. There are~~ were considered. There were however inconsistencies in sensor spacing and the MAGT at zero annual temperature amplitude ~~is~~ was not directly measured. Most records of North America are accompanied with meta

10 records which suggest ~~es sensor depth~~ a sensor depth (closest to zero annual temperature amplitude) for approximation of the MAGT. But this information is unavailable for the majority of records from Asia. MAGT together with the year which they represent ~~has been~~ is available for 64 sites only (24% of all sites relevant for the analyses period). The boreholes represent a MAGT range from -15°C to 6°C.

2.2.1 ~~Spatial information on environmental conditions~~

15 ## 2.3 Spatial information on environmental conditions

The circumpolar permafrost map by Brown et al. (1997) depicts the permafrost extent divided in different classes as well as the ground ice content for the Northern Hemisphere (20°N to 90°N). The data set defines permafrost as frozen ground that remains at or below 0°C for at least two years. Areas are classified as continuous, discontinuous, sporadic or isolated permafrost with differing ground ice content. The classes correspond to percent area categories: 90-100%, 50-90%, 10-50%, <10%, and no

20 permafrost. These classes ~~have been~~ were compared separately and as one aggregated class (excluding areas of no permafrost) to the results of the frozen day classifications. Zones with specified ground ice content are also supplied with the permafrost map and used in this study. Classes are: high >20%, medium 10-20% and low <10%.

The Global Snow Monitoring for Climate Research (GlobSnow) dataset provides information about SWE and snow extent for the Northern Hemisphere (25°N - 84°N) (Metsämäki et al., 2015; GlobSnow). The products are based on Scanning Multi-

25 channel Microwave Radiometer (SMMR), Special Sensor Microwave Imager (SSM/I), and Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) sensor data in combination with ground based measurements (GlobSnow; Takala et al., 2011). The SWE data product used in this study has a spatial resolution of 25km. The SWE values are provided as daily SWE, weekly aggregated SWE and monthly aggregated SWE. In this study, we used the monthly aggregated SWE which provides a maximum SWE value for each month to determine the maximum SWE for each winter.

30 The Global Land Cover 2000 Project (GLC 2000) provides global landcover information at 1km resolution (GLC2000, 2003). The data is mainly based on Satellite Pour l'Observation de la Terre-4 (SPOT-4) observations, partially supported by other Earth observing sensors (Bartholomé and Belward, 2005).

We extracted the soil texture for all points from the Harmonized world soil data base Fischer et al. (2008) and soil organic carbon (SOC) from the Northern Circumpolar Soil Carbon Database (NCSCD) by Hugelius et al. (2013, 2014) to include the soil properties in our analysis. The NCSCD is a polygon-based digital database compiled from harmonized regional soil classification maps in which data on soils have been linked to pedon data from the northern permafrost regions to calculate

5   SOC content and mass. It includes SOC values for 0-30 cm, 0-100 cm, 0 - 200 cm and 0-300 cm.

3   Methods

3.1   Preparation of borehole temperature data

The MAGT is usually calculated at the depth of zero annual amplitude (ZAA) for permafrost studies. As the availability of data at specific depths is limited, representing or reaching the depth of ZAA is not possible for all cases. The MAGT, defined as

10   the temperature at a specific site in this study, was therefore calculated for each borehole location at the depth of the minimum MAGT. The minimum MAGT, in a stable climate, would be the same as the MAGT at the depth of ZAA (Bodri and Cermak, 2007). Where available, we have collected also meta data for all GTN-P boreholes regarding MAGT, the year of the calculation and depth of sensor representing MAGT in order to test the validity of this approach. ~~Only sensors with more than~~ The MAGT at coldest sensor depth is in the following referred to as $MAGT_c$. Only sensors instrumented below 1m depth ~~have been~~ were

15   used as the MAGT near the surface can be much colder than at larger depth. Temperature time series of the different sensors ~~have been~~ were tested for gaps and inconsistencies. Only years with complete records ~~have been~~ were considered for calibration and validation. The used sites are located within an area with 150 to 330 number of frozen days as observed in the satellite records in order to account for artifacts which can occur due to e.g. large water bodies within the footprint.

3.2   Pre-processing of satellite records

20   Both surface status data sets underwent post-processing before being used in our permafrost extent and temperature estimation.
The ASCAT surface status flag (SSF) data contain cells with no data where the algorithm failed to produce a result. Gaps were filled by surrounding values (class with majority), ensuring a complete data set. The sum of frozen days per year for every pixel was determined for both satellite records, according to the method of Park et al. (2016). Grid cells where the number of frozen days exceeds the number of thawed days during two consecutive years ~~are~~ were classified as permafrost. We ~~have~~

25   defined the averaging period with respect to the water year from September 1 to August 31 as suggested by Park et al. (2016).
To explore the dependency of the results on snow melting events, the permafrost extent estimation from ASCAT data was carried out excluding the melt days in the count of frozen days (FT) and a second analysis counting the melt days as frozen ~~grid cells~~ days (FM).

3.3   ~~Frozen days threshold determination for potential permafrost extent~~

In the study of Park et al. (2016) a threshold of half a year was chosen for the delineation of permafrost extent. Half year corresponds to 180 or 182.5 days in climate models (Saito et al., 2013). In a first step, we extended the analysis to 210 frozen days to test the validity of the suggested threshold. The cross-comparison with the permafrost extent classes considers the information of four thresholds (180, 190, 200, 210) for the entire study area which allowed us to analyze the difference in estimated permafrost extent and the sensitivity of this approach to the chosen threshold. The results from both active and passive microwave freeze/thaw data sets were compared with the permafrost map by Brown et al. (1997).

To evaluate the initial threshold of half a year with in situ data, ASCAT and SSM/I data were extracted for each of the borehole locations. We classified borehole (coldest sensor) derived MAGT > 0 as 0 and MAGT <= 0 as 1 and the DOY > "threshold" as 1 and DOY < "threshold" as 0. Kendall's Tau (τ) was used to examine the correlation coefficient between the in situ and satellite derived data set. It was chosen as it provides a method to measure the ordinal association to measured or calculated quantities.

In order to determine the most suitable limit to map permafrost extent, thresholds where varied from 180 to 210 days in one day steps. For the evaluation of the initial threshold of 180 days we divided the dataset (all years and stations) into four sectors; positive and negative MAGT versus DOY >= 180 and DOY < 180. The number of considered data points differ between SSM/I and ASCAT due to different masking schemes.

The classified maps (initial half year threshold) have been summed up for each data type for the four periods to obtain information on interannual and spatial variability. ASCAT FT and SSM/I results have been compared by deriving the difference between the individual sums.

3.3 Model parameterization for potential mean annual ground temperature retrieval

The relationship between $MAGT_c$ and frozen day of year was further examined for the retrieval of ground temperature and consecutive determination of permafrost extent by using the 0°C threshold. Only data inside the range of 150 to 330 frozen DOY have been were considered. Additionally, sites which are located on islands in the high Arctic have been were excluded with respect to microwave sensor footprint size. The remaining 168 sites haven been were used to fit the model. The records from ASCAT as well as SSM/I have been were split into two parts by defining a calibration (2009-2011) and a validation period (2007-2008). We tested linear, logarithmic and polynomial functions on their ability to describe the relationship between $MAGT_c$ and day of year frozen. We found no significant fit for polynomial functions and a slightly weaker fit for logarithmic functions compared to a simple linear regression. Therefore a linear model has been was applied to the frozen days for the years 2009 to 2011 for the determination of an empirical relationship. The resulting formula was used to estimate the $MAGT_c$ from the day of year data set for the years 2007 and 2008. The differences between the modelled and in situ MAGT$_c$ have been were calculated separately for the two years in order to assess the capability of the approach to capture inter-annual variations, to investigate various environmental impacts (snow water equivalent, landcover type), differences between previously defined permafrost zones and specific regions. The Arctic has been was split up into 14 regions in the latter case (Fig. A1).

The average $MAGT_c$ for the entire time period (2007-2011) has been was calculated for ASCAT FT and SSM/I results and compared. The standard deviation has been was derived in addition.

~~The MAGT values have been eventually~~

3.4 Frozen days threshold determination for potential permafrost extent

The modelled $MAGT_c$ values were also classified for each year in order obtain permafrost extent maps (binary maps of values below and above ~~zero degree Celsius). They have been~~ 0°C). Results were compared to two further approaches for threshold

5   determination. In the study of Park et al. (2016) a threshold of half a year frozen days was chosen for the delineation of permafrost extent. Half year corresponds to 180 or 182.5 days in climate models (Saito et al., 2013). In a first step, we extended the analysis to 210 frozen days to test the validity of the suggested threshold. The cross-comparison with the permafrost extent classes considered the information of four thresholds (180, 190, 200, 210) for the entire study area which allowed us to analyze the difference in estimated permafrost extent and the sensitivity of this approach to the chosen threshold. The results from both

10   active and passive microwave freeze/thaw data sets were compared with the permafrost map by Brown et al. (1997). To further evaluate the initial threshold of half a year, in situ data as well as ASCAT and SSM/I number of drozen days were extracted for each of the borehole locations. We classified borehole (coldest sensor) derived $MAGT_c > 0$ as 0 and $MAGT_c <= 0$ as 1 and the DOY > "threshold" as 1 and DOY < "threshold" as 0.

Kendall's Tau (τ) analyses was used as an alternative approach to determine a suitable threshold. The correlation coefficient

15   between the in situ records and satellite derived number of frozen days has been examined. It was chosen as it provides a method to measure the ordinal association to measured or calculated quantities. In order to determine the most suitable limit to map permafrost extent, thresholds where varied from 180 to 210 days in one day steps.

Eventually, the classified maps (classified modelled $MAGT_c$ and half year threshold approach) were summed up for each data type for the four periods to obtain information on interannual and spatial variability. ASCAT FT and SSM/I results ~~have~~

20   ~~been compared similarly as for the test of the initial threshold (see section 3.4)~~were compared by deriving the difference between the individual sums.

4   Results

4.1   Evaluation of in situ MAGT at coldest sensor depth

64 borehole locations had meta data suitable for the assessment. 20 of the boreholes belong to the Vorkuta GTN-P site. R^2

25   between MAGT from meta records ~~and MAGT at coldest sensor depth~~ (GTN-P) and $MAGT_c$ is 0.994 (Fig. 1). The slope of the linear fit ~~is~~ was close to one and RMSE ~~is~~ was 0.38 °C. Derived values ~~tend~~ tended to be colder by on average 0.2 °C. The coldest sensor depth and sensor to be used agreed in 58 (90%) of the samples. Deviations of sensor depth occurred only for the Alaskan sites Bonanza Creek and Atmautluak as well as Arctic Bay (Nunavut Canada), Kursflaket 2 (Sweden), Ishim (Tyumen, Russia) and Kotkino (Western Russia), however, with only little deviations in MAGT. They ~~have been~~ were below

30   0.1°C except for Bonanza Creek (-0.39°C). MAGT~~at cold sensor depth~~$_c$ seems to be also valid as a substitute for MAGT at

ZAA in non-permafrost regions based on the available records.

4.2 ~~Threshold sensitivity analyses and permafrost extent~~ Potential mean annual ground temperature at coldest sensor depth

5   The Pearson correlation for the linear fit decreased slightly from 0.66 to 0.64 if snow melting days were included (Fig. 2 and Tab. 1). The same applied to the residual standard error (2.15 versus 2.21) for the validation years. The slope of the linear fit also differed slightly between the two ASCAT data sets (Fig. 2). It was steeper in case of exclusion of melting days as well as for SSM/I. The spread of in situ temperature values was higher for conditions below 0°C than above. This was similar for all data sets. The majority of $MAGT_c$ values (in a wide range of -10°C to 5°C) were found in the sector between 170 and 250

10   frozen DOY for ASCAT. The Pearson correlation in case of SSM/I was lower with 0.39 and RMSE was higher with 2.53°C (Fig. 2). A difference of 11 frozen days of near surface soil corresponded to 1°C in case of ASCAT FT.

    Deviations in derived potential $MAGT_c$ between ASCAT and SSM/I occurred within and outside the continuous permafrost region (Fig. 3). Larger differences occurred over mountain ranges and lake rich areas. Modelled $MAGT_c$ based on SSM/I were mostly warmer in the continuous zone. The temperature difference between ASCAT (FT) and SSM/I exceeded 2°C in the

15   Central Siberian transition zone, where SSM/I retrievals resulted in colder $MAGT_c$ than for ASCAT.

4.2.1 Evaluation with in situ data with respect to region

The regionally averaged difference was mostly below 1°C for all three products (Fig. 4). The deviations of ASCAT results from in situ records differed for most regions also between the two validation years. Large differences occurred specifically in Western Russia and across Siberia for ASCAT FM as well as SSM/I. The exclusion of snow melting days reduced these

20   differences. The deviations of the SSM/I product were largest in central Russia and the Yakutsk region. This was consistent with the patterns observed in comparison to the permafrost map (Fig. 5) and the deviations between the permafrost extent maps (Fig. 6 and Fig. 7).

    The inclusion of days with melting snow reduced the deviations in Yamal-Nenets district, Western Russia, and the Canadian High Arctic and Greenland (Fig. 4). The model result for ASCAT FM was about 2°C colder than in the case of exclusion

25   of melting snow. In contrast, $MAGT_c$ increased for Alaska. Median values were in general similar between inclusion and exclusion of melting snow (Fig. 8). All model results, ASCAT FT, FM and SSM/I, showed the highest deviations for the Northern Russian Far East. Modelled temperatures were up to 7°C too warm (median difference 3 - 5°C). Most boreholes in the region are located in comparably cold permafrost.

4.2.2 Evaluation with in situ data with respect to environmental conditions

30   Deviations of $MAGT_c$ were larger at sites with SWE larger than 150 mm (Fig. 9). Model results were colder than retrievals from the in situ records. This applied also for exclusion of the snow melting period. However, the exclusion reduced deviations

in all medium SWE classes. Sites with high SWE were mostly located in Western Russia (Vorkuta region, Bolvansky) and Western Siberia (Nadym region, Urengoy).

The regions identified as sporadic and isolated permafrost in Brown et al. (1997) showed comparably large offsets (Fig. 9). The isolated class did however only include five samples. The exclusion of snowmelt specifically reduced the deviations in the discontinuous zone. In case of the ASCAT products, average deviation was always below 1°C in the continuous permafrost zone. The exclusion of melting snow clearly reduced deviations in case of high ground ice content.

Figure 14 shows results for different landcover classes. This seemed to result in higher deviations for specifically the SSM/I results. Only two samples were available in the burned area class. The model results were warmer in both cases. Sites which were contained in the water class were located in proximity to larger lakes or ocean.

Deviations seemed to be larger for sites with a certain soil organic carbon content (> 10kg m^{-2} within the top meter; Fig. 10). Modelled MAGT$_c$ was in the order of 2°C higher than derived from in situ data. There was, however, strong variation between the years.

4.3 Permafrost extent estimation

4.3.1 Threshold sensitivity analyses

The area mapped as permafrost with a frozen day threshold of 180 days for the period 2007-2011 differed regionally between the ASCAT and the SSM/I data set (Fig. 6). The extent of areas where only ASCAT determined permafrost was about four times higher than for SSM/I. The latter largely occurred outside the expected permafrost region. Deviations in the transition zone were low for Western Siberia but still present. The classification using SSM/I data resulted in a smaller permafrost extent for the Canadian Arctic compared to Brown et al. (1997). Most of the transition zone (including sporadic and isolated) in this region had more than 180 days frozen in the ASCAT dataset. The Canadian High Arctic was also largely not covered by the SSM/I dataset. ASCAT overestimated the permafrost extent in Scandinavia. In general, high latitudinal lowland permafrost was overestimated with ASCAT whereas the extent in more southern uplands and mountain regions was overestimated with SSM/I using the initial threshold.

Mapped permafrost extent varied for each threshold step across the different zones. The largest deviation in area extent for continuous permafrost occurred in the SSM/I result with more than 2.5 Mio km^2 lower values than in Brown et al. (1997) (more than 12 %, Tab. 2). Less than 1 Mio km^2 (less than 5% of total continuous permafrost area) was missed by ASCAT. This applied to thresholds between 180 and 200 frozen days (excluding snow melt days). Matching extent was lowest for 200 days in case of inclusion of snow melt days. ASCAT mapped more permafrost outside the boundaries of Brown et al. (1997) than SSM/I in case of a 180 day threshold but agreed better than SSM/I for discontinuous and isolated permafrost area. This resulted in lower percentage agreement of the SSM/I product for the total permafrost extent (Tab. 3). The false permafrost detection by ASCAT was out-weighted by a significantly higher detection performance for the total extent. There was also more year to year variability in the results from the analysis using SSM/I data than from ASCAT.

4.4 ~~Determination of optimal threshold with Kendall's τ test and in situ measurements~~

4.3.1 Determination of optimal threshold with Kendall's τ test and in situ measurements

The comparison of the satellite records with in situ data ~~demonstrates~~ demonstrated differences between the datasets. For the ASCAT data sets the best τ was found to coincide with 204 (FM) and 203 (FT) frozen DOY (Tab. 1), whereas for the SSM/I,

5   190 DOY showed the highest τ with a steep gradient before and after the peak (Fig. 11).

With the DOY threshold of 180 days, the algorithm ~~tends~~ tended to overestimate the amount of negative $MAGT_c$ values, while the threshold found by the Kendall's τ ~~leads~~ led to an underestimation of negative $MAGT_c$ values below the threshold. (Tab. 4). It ~~is~~ was highest for exclusion of snow melting days. On the other hand the ASCAT results ~~show~~ showed nearly no negative $MAGT_c$ below the threshold of 180. The thresholds delineated by the best correlation coefficient (Kendall's τ)

10   ~~indicate~~ indicated for all data sets a better performance regarding positive $MAGT_c$ in areas below the threshold. However, the higher thresholds ~~lead~~ led also to more negative $MAGT_c$ in these areas. For the SSM/I 75% of MAGT~~temperatures can~~$_c$ ~~temperatures could~~ be correctly allocated with both thresholds. The ASCAT results ~~show~~ showed a higher accuracy with more than 80% correctly assigned values.

~~Comparison matrix between in situ Mean Annual Ground Temperature (MAGT, at coldest sensor depth) classes (below or~~

15   ~~above zero °C) and classified number of frozen days per year from ASCAT (FT - excluding snow melt days, FM - including snow melt days) and SSM/I for 2007-2012. Both, the initial 180 Days of Year (DOY) threshold and the optimal threshold based on best Kendall's τ are assessed. Values represent numbers of sites for multiple years. Rows indicating false classifications are in italic. Sensor ASCAT FT ASCAT FM SSM/IDOY 180203180204180190Positive MAGT & DOY <thres. 142 211 127 199 115 148 Negative MAGT & DOY <thresh. 2 70 0 35 14 41 Positive MAGT & DOY >= thresh. 138 69 153 81 157 124~~

20   ~~Negative MAGT & DOY >= thresh. 494 426 496 461 398 371 Sum of data points 776 776 776 776 684 684~~

~~Comparison between in situ Mean Annual Ground Temperature (MAGT, at coldest sensor depth) with number of frozen days per year from ASCAT (FT - excluding snow melt days, FM - including snow melt days) and SSM/I for 2007-2012.~~

~~Parameter ASCAT FT ASCAT FM SSM/I Pearson correlation for linear fit 0.66 0.64 0.39 RMSE validation years (2007/2008) 2.15 2.21 2.53 RMSE all years (2007 - 2011) 2.22 2.24 2.57 DOY for zero degree C from linear fit 207 212~~

25   ~~205 DOY for best Kendall's τ203 204 190~~

ASCAT and SSM/I maps derived with the initial threshold of 180 days mostly ~~agree~~ agreed in the continuous permafrost zones over the four analyses periods (Fig. 6). ASCAT ~~overestimates~~ overestimated permafrost extent in Northern America, Scandinavia and Western Siberia. SSM/I ~~overestimates~~ overestimated specifically in southern Central Siberia.

4.4 ~~Mean annual ground temperature determination for retrieval of permafrost extent~~

30   #### 4.3.1 Classification of modelled MAGT for coldest sensor depth

The classification of the derived MAGT~~dataset provides~~$_c$ dataset provided different results than the other methods for permafrost extent determination. The DOY for intersection of the derived linear function at ~~zero degree C corresponds~~ 0°C corresponded

to a higher number of frozen days in all cases (Tab. 1). ~~The Pearson correlation for the linear fit decreases slightly from 0.66 to 0.64 if snow melting days are included~~

The FT and FM methods differed more for the year 2007 than for 2008 (Fig. ~~2~~ 12 and Tab. ~~1~~5). The ~~same applies to the residual standard error (2.15 versus 2.21) for the validation years. The slope of the linear fit also differs slightly between the two ASCAT data sets (Fig. 2). It is steeper in case of exclusion of melting days as well as for~~ SSM/I ~~. The spread of in situ temperature values is higher for conditions below 0°C than above. This is similar for all data sets. The majority of MAGT values (in a wide range of -10°C to 5°C)are found in the sector between 170 and 250 frozen DOY for ASCAT. The Pearson correlation in case of SSM/I is lower with 0.39 and RMSE is higher with 2.53°C (Fig. 2). A difference of 11 frozen days of near surface soil corresponds to 1°C in case of ASCAT FT.~~ extent missed a large proportion of continuous permafrost and mapped additional area as permafrost (more pronounced in 2007, Tab. 5), similarly to the initial threshold based comparison to the permafrost extent map (Tab. 2).

The number of years which have MAGT~~below zero degree Celsius are~~ below 0°C were similar between ASCAT (FT) and SSM/I in the lowland continuous permafrost zones (Fig. 6). Deviations ~~are~~ were high in transition zones, especially southern Central Siberia, with a similar spatial pattern as observed already in the case of initial threshold comparison (Fig. 5). The results ~~deviate~~ deviated by four periods in this region, meaning that SSM/I always ~~remains~~ remained below a MAGT~~below zero degree Celsius~~ below 0°C, and ASCAT in none of the periods.

~~Deviations in derived potential MAGT between ASCAT and SSM/I occur within and outside the continuous permafrost region (Fig. 3). Larger differences occur over mountain ranges and lake rich areas. The results based on SSM/I are mostly warmer in the continuous zone. The temperature difference between ASCAT (FT) and SSM/I exceeds two degree Celsius in the Central Siberian transition zone, where SSM/I retrievals result in colder MAGT than for ASCAT.~~

4.3.2 Comparison between permafrost extent maps

~~The FT and FM methods differ more for the year 2007 than for 2008 (Fig. 12 and Tab. 5). The SSM/I extent misses a large proportion of continuous permafrost and maps additional area as permafrost (more pronounced in 2007, Tab. 5), similarly to the initial threshold based comparison to the permafrost extent map (Tab. 2).~~

4.3.3 ~~Evaluation with in situ data with respect to region~~

~~The regionally averaged difference is mostly below one degree Celsius for all three products (Fig. 4). The deviations of ASCAT results from in situ records differ for most regions also between the two validation years. Large differences occur specifically in Western Russia and across Siberia for ASCAT FM as well as SSM/I. The exclusion of snow melting days reduces these differences. The deviations of the SSM/I product are largest in central Russia and the Yakutsk region. This is consistent with the patterns observed in comparison to the permafrost map (Fig. 5) and the deviations between the permafrost extent maps (Fig. 6 and Fig. 7).~~

~~The inclusion of days with melting snow reduces the deviations in Yamal-Nenets district, Western Russia, and the Canadian High Arctic and and Greenland, but increases for Alaska (Fig. 4). The model result for ASCAT FM is about two degree~~

colder than what is obtained from the in situ records in the first case. Median values are in general similar between inclusion and exclusion of melting snow (Fig. 8). All model results, ASCAT FT, FM and SSM/I, show the highest deviations for the Northern Russian Far East. Modelled temperatures are are up to seven degrees too warm (median difference 3 - 5°C). Most boreholes in the region are located in comparably cold permafrost.

4.3.3 Evaluation with in situ data with respect to environmental conditions

Deviations of MAGT are larger at sites with SWE larger than 150 mm (Fig. 9). Model results are colder than retrievals from the in situ records. This applies also for exclusion of the snow melting period. However, the exclusion reduces deviations in all medium SWE classes. Sites with high SWE are mostly located in Western Russia (Vorkuta region, Bolvansky) and Western Siberia (Nadym region, Urengoy).

The regions identified as sporadic and isolated permafrost in Brown et al. (1997) show comparably large offsets (Fig. 9). The isolated class does however only includes five samples. The exclusion of snowmelt specifically reduces the deviations in the discontinuous zone. In case of the ASCAT products, average deviation is always below 1°C in the continuous permafrost zone. The exclusion of melting snow clearly reduces deviations in case of high ground ice content.

Figure 14 shows results for different landcover classes. This seems to result in higher deviations for specifically the SSM/I results. Only two samples were available in the burned area class. The model results are warmer in both cases. Sites which are contained in the water class are located in proximity to larger lakes or ocean.

Deviations seem to be larger for sites with a certain soil organic carbon content (> 10kg m^{-2} within the top meter; Fig. 10). Modelled MAGT is in the order of 2°C higher than derived from in situ data. There is, however, strong variation between the years.

4.4 Comparison between permafrost extent maps

The results from all The results from all three tested methods for potential permafrost determination (half year threshold, best Kendall's τ and from modelled MAGT$_c$) are shown in Fig. 5. The largest spatial differences between included/excluded melting days occur occurred in the same regions (Northeastern Canada and Scandinavia), which show showed the largest differences between the ASCAT and SSM/I 180 day threshold maps. An additional region of disagreement is was Yakutsk. The inclusion of snow melting days leads led here to a reduction of permafrost extent which is was contradictory to most other regions.

The modelled extent and best Kendall's τ threshold results includes included the discontinuous permafrost zone in Northern America for ASCAT (Fig. 5). The exclusion of melting snow has had a large impact over Scandinavia. It leads led to results which are more similar to the map of Brown et al. (1997). It is was also one of the regions with extensive snow melt detected by ASCAT (Fig. A2). Patterns of continuous permafrost boundaries in central Asia (e.g. region around lake Baikal) are were better represented in the ASCAT maps than in the SSM/I retrievals.

Variation of potential MAGT$_c$ from year to year are were much larger for SSM/I. Standard deviation is was comparably high for the transition zone in Western Siberia and also the Canadian Arctic (Fig. 13). This pattern agrees agreed with results obtained with the initial threshold in comparison to Brown et al. (1997) (see Section 4.3).

5 Discussion

5.1 General issues

The ~~overall performance of permafrost extent mapping using number of frozen days is limited but reveals regional patterns in uncertainties. The extent of permafrost estimated with the initial threshold is in the order of the actual extent but the error of commission is relatively large. This problem can be tackled by adjustment of the threshold and use of different type of satellite acquisitions (ASCAT versus SSM/I). The performance of the~~ performance of the empirical model for MAGT ~~(based on coldest sensor,$_c$~~ (Tab. 1) ~~is~~ was partially lower than what can be achieved with the more complex TTOP model (Westermann et al., 2015) which considers terrain, snow, land cover and land surface temperature measurements from satellite data. Westermann et al. (2015) reported a model accuracy of 2.5°C for MAGT. Permafrost temperatures are also not always representing current climate conditions (Lachembruch and Marshall, 1986). This may regionally impact the comparability of the borehole records with surface observations. Estimates of permafrost temperatures as well as extent ~~are therefore providing~~ therefore provide a potential distribution only. Results may however support the identification of regions where permafrost extent maps, including continuity classes, need to be treated with care. This includes Western and Central Siberia. The overall performance of permafrost extent mapping using number of frozen days is limited but reveals regional patterns in uncertainties. The extent of permafrost estimated with the initial threshold is in the order of the actual extent. The error of commission is, however, relatively large. This problem can be tackled by adjustment of the threshold and use of different types of satellite acquisitions (ASCAT versus SSM/I).

The mean annual ground temperature from boreholes was calculated from the coldest sensor below 1 m. The depth of these sensors ~~varies~~ varied from borehole to borehole which may impact the empirical model representativity. A high number of sites ~~has been~~ was however chosen for calibration which may weaken the impact. The evaluation results with MAGT (expected to be at or close to zero annual amplitude) in the meta records for 24% of the sites support the assumption that the sensor at coldest depth can be used for approximation. Uncertainties introduced by variable sensor spacing can, however, not be addressed with the available data.

The in all cases (ASCAT and SSM/I, for all permafrost retrieval methods) higher accounts of number of frozen days than the previously suggested half year threshold agrees with field observations by Luoto et al. (2004), who estimated a minimum number of 200 days for permafrost formation for Scandinavia. 200 days corresponds to approximately 0.5°C modelled MAGT$_c$ in case of ASCAT FT. Considering local factors such as water-saturated peat and organic layer as well as uncertainties in the retrieval (difference to actual mean temperature), this might still be sufficient for permafrost formation. Variations of topography within the footprint lead also to local deviations from days to weeks (Bergstedt and Bartsch, 2017). We therefore suggest the consideration of temperature buffers when such data are applied. Boreholes with in situ MAGT$_c$ below 0°C located in these buffer zones may represent sites where local factors are important. In addition, the role of past climate conditions for present ground temperatures need to be considered as well as location specific soil and snow properties (see Section 5.4).

The classes in Brown et al. (1997) correspond to the area fraction of permanently frozen ground. In case of the isolated permafrost class it can be assumed that at least 10% are below 0°C, but the actual mean temperature for a region in these areas

(as e.g. represented by an ASCAT or SSM/I cell) can be below or above 0°C depending on local parameters such as topography and soil properties which impact thermal conductivity. The latter especially plays a role for occurrence of permafrost in the transition zone. Data sets of higher spatial resolution would be required. Relevant measurements from microwave data are only available from active systems due to technical constraints. Synthetic Aperture Radar (SAR) instruments could be used in case

5   of sufficient sampling intervals (Park et al., 2011b). Current systems and acquisition plans do however not provide a sufficient temporal and spatial sampling (Bartsch et al., 2016).

5.2   Regional issues

A threshold higher than the previously suggested half year leads to better performance of ASCAT than for SSM/I for permafrost extent retrieval especially over Scandinavia, Western Russia and Southern Russian Far East (Fig. 5; region overview in Fig.

10   A1). ASCAT better captures the regional patterns as in Brown et al. (1997) with the exception of Scandinavia. The actual temperature amplitude (freezing and thawing degree days) may need to be considered in this region. The longer snow melt period (Fig. A2) also indicates a certain amount of snow which may lead to decoupling of air and ground temperatures.

The highest density of boreholes with available data is in the Vorkuta region in Western Russia. This region shows the largest sensitivity to inclusion/exclusion of snow melting days (Fig. 4). Here, the ASCAT $MAGT_c$ results differ by more than

15   ~~two degree Celsius~~ 2°C (lower temperature). This might have an impact on the average deviation derived from the entire borehole records. It is likely larger for the ASCAT FM result than calculated as most other regions show better agreement.

The validation results in the regions 'Central Russia' and 'Central Siberia' differ from those of the other areas. SSM/I results suggest between ~~one to two degree~~ 1 to 2°C lower regionally averaged $MAGT_c$ values (Fig. 4 and Fig. 8). The majority of boreholes located in these regions show higher than ~~zero degree MAGT at coldest sensor depth~~ 0°C $MAGT_c$ (Fig. A1). This

20   deviation therefore impacts the permafrost boundary retrieval based on freeze/thaw records from SSM/I. These regions are characterized by a longer than average period of diurnal thaw and refreeze cycling during spring melt (Bartsch et al., 2007; Bartsch, 2010). Acquisition timing may therefore play a role for the determination of the length of the frozen period.

The Greenland and Svalbard sites are expected to have the highest variations due to the mixture of glaciers, land area and ocean within the ASCAT as well as SSM/I pixels. There is actually no coverage of the Greenland and several Canadian High

25   Arctic sites in the SSM/I records.

5.3   Performance differences between ASCAT and SSM/I

Results suggest that SSM/I freeze/thaw records are less suitable to derive actual $MAGT_c$ ~~values below zero degree Celsius~~ values below 0°C compared to ASCAT (Fig. 3). The thresholds obtained for SSM/I are considerably lower than for ASCAT, what might be the result of the different wavelength, the sensing technique (passive/active), overpass timing and classification

30   methods used to create these data sets. The considerable lower number of frozen days in regions with low $MAGT_c$ might be the result of the retrieval method (treatment regarding acquisition timing) and sensitivity to soil state changes. The instruments also differ in wavelength apart from the fact that one is active and the other passive. ASCAT uses C-band with about 5.7 cm and the SSM/I channel used by Park et al. (2016) about 0.8 cm. This results in different signal interactions with objects on the

earth surface including snow and vegetation. It can be expected that the C-band signal is less sensitive to interactions, although present. The latter issue could be addressed by L-band missions with even lower frequency than ASCAT such as SMOS (Soil Moisture and Ocean Salinity Mission) or SMAP (Soil Moisture Active and Passive). In general, the role of acquisition timing and sampling rate needs to be investigated in more detail for permafrost related applications for ASCAT as well as SSM/I.

5   The lower performance of SSM/I might be also attributed to the fact that it has an even larger footprint (although gridded to 25km) than ASCAT. The validation results are therefore not fully comparable between the sensors for the entire Arctic, only on regional level. This may also affect the calibration, since the number of available samples is lower for SSM/I. It affects especially colder sites (Fig. 2). In general, less areas are masked in the ASCAT product. This applies to especially lake rich regions (Fig. 5). Findings of Bergstedt and Bartsch (2017) suggest an offset of the state change in the resolution cell due to

10   lakes. This may lead to lower accuracy in these regions.

5.4   The role of environmental conditions

The amount of snow seems to play the most important role for the applicability of the frozen day approach. One may expect warmer modelled $MAGT_c$ than in situ values in transitions zones due to the fact that boreholes in transition zones often represent isolated patches of permafrost. However, the opposite is the case (Fig. 9). This may relate to the importance of the

15   insulation effect of snow cover in these regions, e.g. as known for Scandinavia (Luoto et al., 2004).

The number of snow melting days is in general highly variable in the Arctic (Bartsch, 2010) but the melting period is comparably short (Zhang et al., 2005). Snowmelt is expected to delay the soil surface warming due to latent heat and therefore cools the soils (Zhang et al., 2005). Latent heat released due to refreeze of meltwater may have a warming effect after a few days (Dingman et al., 1980). Dingman et al. (1980) also report start of soil thaw before the end of snowmelt at Barrow. The

20   overall impact of snowmelt is expected to be dependent on local conditions (Zhang et al., 2005). Our results suggest that there is a warming effect with an impact on $MAGT_c$. Days with melting snow should be therefore treated as unfrozen. This leads to higher $MAGT_c$ (on average 1°C for considered borehole locations) and better agreement with in situ measurements. Exclusion of the snowmelt period is also consistent with calculation of thawing and freezing degree days from air temperature data. The snowmelt period does also count as unfrozen in this case.

25   Snow melting days are, however, not mapped for all grid points in case of the ASCAT data set (Fig. A2). This may depend on snow depth as well as acquisition timing. The coverage pattern is irregular across the Arctic with a mix of morning and evening (ascending and descending orbits) measurements (Bergstedt and Bartsch, 2017). Usually only evening measurements capture the melt as diurnal freeze and thaw cycles are common (Bartsch et al., 2007). The differentiation between frozen and melting days may however be valid in regions with prolonged melt and high SWE. Areas with melting snow in the ASCAT

30   data set are common in the high Arctic, in areas with low $MAGT_c$ (Fig. A2) as well as areas in the transitions zone (such as Scandinavia). This pattern differs from the length of the snow melt period detected with Seawinds QuikScat, a Ku-band scatterometer (Bartsch, 2010), which provides several measurements per day. The number of days with snow melt are much lower in C-band than in the Ku-band product. This could be attributed to the lower sensitivity of C-band to melting processes

and the limited temporal sampling. In addition, the QuikScat results reported by Bartsch (2010) are representing periods of freezing and thawing which can contain breaks (with frozen conditions) of up to ten days.

Landcover seems to be of minor importance (Fig. 14). It only plays a role when comparing performance of SSM/I versus ASCAT. Boreholes which fall into the water class according to GLC2000 are located close to coasts. The coarser resolution SSM/I datasets is here more affected than ASCAT. Deviations are therefore larger for SSM/I. Modelled results are on average warmer in all cases for the class 'shrubs'. It mostly represents the tundra biome and is also the largest sample. It overlaps with continuous permafrost which shows a similar bias (Fig. 9). The difference is larger in case of SSM/I which might relate to the fact that the variation of below zero $MAGT_c$ is not well reflected in the SSM/I derived number of frozen days, and also fewer samples have been available for colder sites due to masking (Fig. 2).

An influence of soil organic carbon on deviations between the modelled temperature and in situ measurements cannot be clearly exemplified (Fig. 10). This might be partially influenced by spatial inconsistencies and the nature of the used database (Bartsch et al., 2016). SOC is represented by areal averages only, but it is the most detailed dataset available to date. The tendency for warmer modelled temperature in case of sites with more than 10kg m^{-2} agrees, however, with the expected effect of organic soils on ground thermal regime (e.g. Smith and Riseborough, 2002). Sites with sandy loam and loamy sand seem to differ between ASCAT and SSM/I results. They represent different regions across Central Siberia and Western Russia, but the sample size is much smaller than for the other categories. It, however, agrees with the observed regional patterns for differences between the methods.

6 Conclusions

Conventional approaches for spatially continuous mapping of permafrost temperatures require gap filling or spatial interpolation. ~~The~~ This applies to use of in situ temperature measurements as well as to satellite derived land surface temperature (thermal and passive microwave). The direct comparison of microwave satellite derived surface status (frozen/unfrozen), rather than actual temperatures, to borehole temperatures revealed the potential of such information for ground temperature estimation. C-band ~~radar record provide a purely observational account. However, a number of issues need to be considered.~~

~~Differences between the tested methods for identification of permafrost extent are low within the continuous permafrost zone (as defined in Brown et al. (1997))~~ backscatter based records performed better than surface status derived from passive microwave brightness temperature. The relationship between MAGT at coldest sensor depth with SSM/I derived surface status is comparably weak, especially for lower temperatures. ASCAT can capture variations over the full MAGT range investigated. The C-band scatterometer record can therefore provide a purely observational estimate of MAGT. This refers to temperatures at coldest sensor depth and not necessarily zero annual temperature range, due to limitations of the in situ data records. It could be, however, shown that MAGT at coldest sensor depth can be used as substitute for actual MAGT for validation and calibration purposes. ~~Exceptions are mountain ranges. Deviations in transition areas are largest in Central Siberia and areas with high snow depth . This underlines the importance of snow and suggests that advanced models should be applied in the areas of the mountain ranges in central Asia, including southern and central Yakutia and Mongolia.~~

~~The direct comparison to borehole temperatures reveals the potential of C-band scatterometer based surface status detection for ground temperature estimation~~Our study also points to the role of snow status (dry or melting) for the temperature in the soil beneath and subsequent impact on $MAGT_C$. A linear empirical model ~~performs with an RMSE of less than 2.2 °C~~ performed best when days with melting snow ~~are~~ were excluded. The ~~modelled temperature deviates~~ overall RMSE was 2.2 °C with ASCAT but the modelled temperature deviated on average by less than ~~one degree Celsius~~ 1°C in footprints without glaciers and mix of land and water. Especially regions with large variations of frozen days between the years and/or between the three different ~~products~~ analyses (ASCAT with snow melting days and without, the SSM/I records) need to be further investigated with respect to the representativity of the borehole records and derived temperatures respectively (e.g. Western Siberia). They mostly correspond to the permafrost transition zones. The validity of the coarse resolution microwave satellite records for the point locations needs to be confirmed by e.g. using higher spatial resolution synthetic aperture radar (SAR) records. More detailed analyses of the ~~role~~ impact of melting snow conditions is required in addition in order to clarify underlying processes. Results also exemplify the role of organic material for thermal conductivity which is not accounted for with the application of a global empirical relationship.

In addition, the suitability of surface state information from satellite data for permafrost extent estimation could be confirmed, but differences between the tested methods and datasets were also evident. Agreement was high within the continuous permafrost zone (as defined in Brown et al. (1997)), except for mountain ranges. Deviations in transition areas were largest in Central Siberia and areas with high snow depth. This underlines the importance of snow and suggests that advanced models should be applied in the areas of the mountain ranges in central Asia, including southern and central Yakutia and Mongolia. These regional differences should be considered in case of interpretation of especially long-term trends.

7 Data availability

The average annual sums of frozen and snow melting days derived from Metop ASCAT are available via the ESA DUE GlobPermafrost project WebGIS and catalogue. www.globpermafrost.info.

Author contributions. CK has performed all data analyses and contributed to the preparation of figures. The concept of the paper was jointly developed by CK and AB. AB wrote the majority of manuscript. CK contributed to the writing of the in situ, methods and results sections, HB to the satellite data description and discussion.

Competing interests. The authors declare no conflict of interest

Acknowledgements. This work was supported by the Austrian Science Fund (Fonds zur Förderung der wissenschaftlichen Forschung, FWF) under Grant [I 1401] (Joint Russian–Austrian project COLD-Yamal), Grant [DK W1237-N23] (Doctoral College GIScience) as well as the European Space Agency project DUE GlobPermafrost (contract number 4000116196/15/I-NB).

References

André, C., Ottlé, C., Royer, A., and Maignan, F.: Land surface temperature retrieval over circumpolar Arctic using SSM/I–SSMIS and MODIS data, Remote Sensing of Environment, 162, 1 – 10, doi:10.1016/j.rse.2015.01.028, 2015.

Bartholomé, E. and Belward, A. S.: GLC2000: a new approach to global land cover mapping from Earth observation data, International Journal of Remote Sensing, 26, 1959–1977, doi:10.1080/01431160412331291297, https://doi.org/10.1080/01431160412331291297, 2005.

Bartsch, A.: Ten Years of SeaWinds on QuikSCAT for Snow Applications, Remote Sensing, 2, 1142–1156, doi:10.3390/rs2041142, 2010.

Bartsch, A. and Seifert, F. M.: The ESA DUE Permafrost project - A service for high latitude research, in: 2012 IEEE International Geoscience and Remote Sensing Symposium, pp. 5222–5225, doi:10.1109/IGARSS.2012.6352432, 2012.

Bartsch, A., Kidd, R. A., Wagner, W., and Bartalis, Z.: Temporal and Spatial Variability of the Beginning and End of Daily Spring Freeze/Thaw Cycles Derived from Scatterometer Data, Remote Sensing of Environment, 106, 360–374, doi:10.1016/j.rse.2006.09.004, 2007.

Bartsch, A., Höfler, A., Kroisleitner, C., and Trofaier, A. M.: Land Cover Mapping in Northern High Latitude Permafrost Regions with Satellite Data: Achievements and Remaining Challenges, Remote Sensing, 8, 979, doi:10.3390/rs8120979, 2016.

Bartsch, A., Kroisleitner, C., and Heim, B.: Circumpolar Landscape Units, links to GeoTIFFs, doi:10.1594/PANGAEA.864508, https://doi.pangaea.de/10.1594/PANGAEA.864508, supplement to: Bartsch, A et al. (2016): An Assessment of Permafrost Long-term monitoring sites with circumpolar satellite derived datasets. Proceedings of the ESA Living Planet Symposium 2016. ESA SP-740, Prague, Czech Republic, 6 pp, hdl:10013/epic.48548.d001, 2016.

Bartsch, A., Widhalm, B., Kuhry, P., Hugelius, G., Palmtag, J., and Siewert, M. B.: Can C-band synthetic aperture radar be used to estimate soil organic carbon storage in tundra?, Biogeosciences, 13, 5453–5470, doi:10.5194/bg-13-5453-2016, https://www.biogeosciences.net/13/5453/2016/, 2016.

Bergstedt, H. and Bartsch, A.: Surface State across Scales; Temporal and SpatialPatterns in Land Surface Freeze/Thaw Dynamics, Geosciences, 7, 65, doi:10.3390/geosciences7030065, 2017.

Biskaborn, B. K., Lanckman, J.-P., Lantuit, H., Elger, K., Streletskiy, D. A., Cable, W. L., and Romanovsky, V. E.: The new database of the Global Terrestrial Network for Permafrost (GTN-P), Earth System Science Data, 7, 245–259, doi:10.5194/essd-7-245-2015, 2015.

Bodri, L. and Cermak, V.: Borehole Climatology: a new method how to reconstruct climate, Elsevier Science, 1st edn., 2007.

Brown, J., Ferrians, Jr., O., Heginbottom, J., and Melnikov, E.: Circum-Arctic map of permafrost and ground-ice conditions, 1997.

Cheng, G. and Wu, T.: Responses of permafrost to climate change and their environmental significance, Qinghai-Tibet Plateau, Journal of Geophysical Research: Earth Surface, 112, F02S03, doi:10.1029/2006JF000631, 2007.

Dingman, S., Barry, R., Weller, G., C.Benson, LeDrew, E., and Goodwin, C.: Climate, Snow Cover, Microclimate, and Hydrology, in: An Arctic Ecosystem: the Coastal Tundra at Barrow, edited by Brown, J., Miller, P., and Bunnell, F., chap. 2, pp. 30–65, Hutchinson and Ross, Stroudsburg, PA, 1980.

Dobinski, W.: Permafrost, Earth-Science Reviews, 108, 158 – 169, doi:10.1016/j.earscirev.2011.06.007, 2011.

Figa-Saldaña,, J., Wilson, J., Attema, E., Gelsthorpe, R., Drinkwater, M., and Stoffelen, A.: The advanced scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for European wind scatterometers, Canadian Journal of Remote Sensing, 28, 404–412, doi:10.5589/m02-035, 2002.

Fischer, G., Nachtergaele, F., Prieler, S., van Velthuizen, H., Verelst, L., and Wiberg, D.: Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008), iIASA, Laxenburg, Austria and FAO, Rome, Italy [Online; accessed 06-October-2017], 2008.

GLC2000: The Global Land Cover Map for the Year 2000, European Commission Joint Research Centre, http://www-gem.jrc.it/glc2000. Access date: 12/01/2017, 2003.

GlobSnow: Global Snow Monitoring for Climate Research, http://www.globsnow.info.Accessdate:12/01/2017.

Gruber, S.: Derivation and analysis of a high-resolution estimate of global permafrost zonation, The Cryosphere, 6, 221–233, doi:10.5194/tc-6-221-2012, 2012.

GTN-P: Global Terrestrial Network for Permafrost Database: Permafrost Temperature Data (TSP Thermal State of Permafrost), 2016.

Harris, S. A.: Distribution of zonal permafrost landforms with freezing and thawing indices, Erdkunde, 35, 81–90, doi:10.3112/erdkunde.1981.02.01, 1981.

Hayes, D. J., Kicklighter, D. W., McGuire, A. D., Chen, M., Zhuang, Q., Yuan, F., Melillo, J. M., and Wullschleger, S. D.: The impacts of recent permafrost thaw on land–atmosphere greenhouse gas exchange, Environmental Research Letters, 9, 045 005, doi:10.1088/1748-9326/9/4/045005, 2014.

Hollinger, J. P., Peirce, J. L., and Poe, G. A.: SSM/I instrument evaluation, IEEE Transactions on Geoscience and Remote Sensing, 28, 781–790, doi:10.1109/36.58964, 1990.

Hugelius, G., Tarnocai, C., Broll, G., Canadell, J. G., P.Kuhry, and Swanson, D. K.: The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions, Earth System Science Data, 5, 3–13, doi:10.5194/essd-5-3-2013, 2013.

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, doi:10.5194/bg-11-6573-2014, http://www.biogeosciences.net/11/6573/2014/, 2014.

Kim, Y., Kimball, J. S., Zhang, K., and McDonald, K. C.: Satellite detection of increasing Northern Hemisphere non-frozen seasons from 1979 to 2008: Implications for regional vegetation growth, Remote Sensing of Environment, 121, 472–487, doi:10.1016/j.rse.2012.02.014, 2012.

Kim, Y., Kimball, J. S., Glassy, J., and McDonald, K. C.: MEaSUREs Global Record of Daily Landscape Freeze/Thaw Status, Version 3, doi:10.5067/MEASURES/CRYOSPHERE/nsidc-0477.003, 2014.

Lachembruch, A. H. and Marshall, B. V.: Changing climate: geothermal evidence from permafrost in the Alaskan Arctic, Science, 234, 689 – 696, 1986.

Luoto, M., Fronzek, S., and Zuidhoff, F. S.: Spatial modelling of palsa mires in relation to climate in northern Europe, Earth Surface Processes and Landforms, 29, 1373–1387, doi:10.1002/esp.1099, 2004.

Matthes, H., Rinke, A., Zhou, X., and Dethloff, K.: Uncertainties in coupled regional Arctic climate simulations associated with the used land surface model, Journal of Geophysical Research: Atmospheres, 122, 7755–7771, doi:10.1002/2016JD026213, http://dx.doi.org/10.1002/2016JD026213, 2016JD026213, 2017.

Metsämäki, S., Pulliainen, J., Salminen, M., Luojus, K., Wiesmann, A., Solberg, R., Böttcher, K., Hiltunen, M., and Ripper, E.: Introduction to GlobSnow Snow Extent products with considerations for accuracy assessment, Remote Sensing of Environment, 156, 96 – 108, doi:https://doi.org/10.1016/j.rse.2014.09.018, http://www.sciencedirect.com/science/article/pii/S0034425714003630, 2015.

Naeimi, V., Paulik, C., Bartsch, A., Wagner, W., Kidd, R., Boike, J., and Elger, K.: ASCAT Surface State Flag (SSF): Extracting Information on Surface Freeze/Thaw Conditions from Backscatter Data Using an Empirical Threshold-Analysis Algorithm, IEEE Transactions on Geoscience and Remote Sensing, 50, 2566–2582, doi:10.1109/TGRS.2011.2177667, 2012.

National research Council: Abrupt impacts of climate change: Anticipating surprises, National Academies Press, 2013.

Nelson, F. E. and Outcalt, S. I.: A Computational Method for Prediction and Regionalization of Permafrost, Arctic and Alpine Research, 19, 279–288, 1987.

Nguyen, T., Burn, C., King, D., and Smith, S.: Estimating the extent of near surface permafrost using remote sensing, Mackenzie Delta, Northwest Territories, Permafrost and Periglacial Processes, 20, 141–153, doi:10.1002/ppp.637, 2009.

O'Connor, F. M., Boucher, O., Gedney, N., Jones, C. D., Folberth, G. A., Coppell, R., Friedlingstein, P., Collins, W. J., Chappellaz, J., Ridley, J., and Johnson, C. E.: Possible role of wetlands, permafrost, and methane hydrates in the methane cycle under future climate change: A review, Reviews of Geophysics, 48, RG4005, doi:10.1029/2010RG000326, 2010.

O'Donnell, J. A., Jorgenson, M. T., Harden, J. W., McGuire, A. D., Kanevskiy, M. Z., and Wickland, K. P.: The Effects of Permafrost Thaw on Soil Hydrologic, Thermal, and Carbon Dynamics in an Alaskan Peatland, Ecosystems, 15, 213–229, doi:10.1007/s10021-011-9504-0, 2012.

Park, H., Iijima, Y., Yabuki, H., Ohta, T., Walsh, J., Kodama, Y., and Ohata, T.: The application of a coupled hydrological and biogeochemical model (CHANGE) for modeling of energy, water, and CO2 exchanges over a larch forest in eastern Siberia, Journal of Geophysical Research: Atmospheres, 116, D15 102, doi:10.1029/2010JD015386, 2011a.

Park, H., Kim, Y., and Kimball, J.: Widespread permafrost vulnerability and soil active layer increases over the high northern latitudes inferred from satellite remote sensing and process model assessments, Remote Sensing of Environment, pp. 349–358, doi:10.1016/j.rse.2015.12.046, 2016.

Park, S.-E., Bartsch, A., Sabel, D., Wagner, W., Naeimi, V., and Yamaguchi, Y.: Monitoring Freeze/Thaw Cycles Using ENVISAT ASAR Global Mode, Remote Sensing of Environment, 115, 3457–3467, doi:10.1016/j.rse.2011.08.009, 2011b.

Paulik, C., Melzer, T., Hahn, S., Bartsch, A., Heim, B., Elger, K., and Wagner, W.: Circumpolar surface soil moisture and freeze/thaw surface status remote sensing products (version 4) with links to geotiff images and NetCDF files (2007-01 to 2013-12), doi:10.1594/PANGAEA.832153, https://doi.pangaea.de/10.1594/PANGAEA.832153, 2014.

Reschke, J., Bartsch, A., Schlaffer, S., and Schepaschenko, D.: Capability of C-Band SAR for Operational Wetland Monitoring at High Latitudes, Remote Sensing, 4, 2923–2943, doi:10.3390/rs4102923, 2012.

Romanovsky, V., Smith, S., and Christiansen, H.: Permafrost Thermal State in the Polar Northern Hemisphere during the International Polar Year 2007-2009: a Synthesis, Permafrost and Periglac. Process., 21, 106–116, doi:10.1002/ppp.689, 2010.

Saito, K., Sueyoshi, T., Marchenko, S., Romanovsky, V., Otto-Bliesner, B., Walsh, J., Bigelow, N., Hendricks, A., and Yoshikawa, K.: LGM permafrost distribution: how well can the latest PMIP multi-model ensembles perform reconstruction?, Climate of the Past, 9, 1697–1714, doi:10.5194/cp-9-1697-2013, 2013.

Schuur, E., Bockheim, J., Canadell, J., and Euskirchen, E.: Vulnerability of permafrost carbon to climate change: Implications for the global carbon cycle, BioScience, doi:10.1641/B580807, 2008.

Schuur, E. A. G., 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, doi:10.1038/nature14338, 2015.

Seppälä, M.: The origin of palsas, Geografiska Annaler. Series A. Physical Geography, 1986.

Smith, M. W. and Riseborough, D. W.: Climate and the limits of permafrost: a zonal analysis, Permafrost and Periglacial Processes, 13, 1–15, doi:10.1002/ppp.410, 2002.

Soliman, A., Duguay, C., Saunders, W., and Hachem, S.: Pan-Arctic Land Surface Temperature from MODIS and AATSR: Product Development and Intercomparison, Remote Sensing, 4, 3833–3856, doi:10.3390/rs4123833, 2012.

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J.-P., Koskinen, J., and Bojkov, B.: Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements, Remote Sensing of Environment, 115, 3517 – 3529, doi:https://doi.org/10.1016/j.rse.2011.08.014, http://www.sciencedirect.com/science/article/pii/S0034425711003166, 2011.

Trofaier, A. M., Westermann, S., and Bartsch, A.: Progress in space-borne studies of permafrost for climate science: Towards a multi-ECV approach, Remote Sensing of Environment, 203, 55–70, doi:10.1016/j.rse.2017.05.021, 2017.

Wang, L., Derksen, C., and Brown, R.: Detection of Pan-Arctic Terrestrial Snowmelt from QuikSCAT, 2000-2005, Remote Sensing of Environment, 112, 3794–3805, doi:10.1016/j.rse.2008.05.017, 2008.

Westermann, S., Østby, T., Gisnås, K., Schuler, T., and Etzelmüller, B.: A ground temperature map of the North Atlantic permafrost region based on remote sensing and reanalysis data, Cryosphere, 9, 1303–1319, doi:10.5194/tc-9-1303-2015, 2015.

Woo, M., Kane, D. L., Carey, S. K., and Yang, D.: Progress in permafrost hydrology in the new millennium, Permafrost and Periglacial Processes, 19, 237–254, doi:10.1002/ppp.613, 2008.

Zhang, T., Frauenfeld, O. W., Serreze, M. C., Etringer, A., Oelke, C., McCreight, J., Barry, R. G., Gilichinsky, D., Yang, D., Ye, H., Ling, F., and Chudinova, S.: Spatial and temporal variability in active layer thickness over the Russian Arctic drainage basin, Journal of Geophysical Research: Atmospheres, 110, D16 101, doi:10.1029/2004JD005642, 2005.

Zhang, T., Barry, R., Knowles, K., Heginbottom, J., and Brown, J.: Statistics and characteristics of permafrost and ground-ice distribution in the Northern Hemisphere, Polar Geography, 31, 47–68, doi:10.1080/10889370802175895, 2008.

Zwieback, S., Paulik, C., and Wagner, W.: Frozen Soil Detection Based on Advanced Scatterometer Observations and Air Temperature Data as Part of Soil Moisture Retrieval, Remote Sens, 7, 3206–3231, doi:10.3390/rs70303206, 2015.

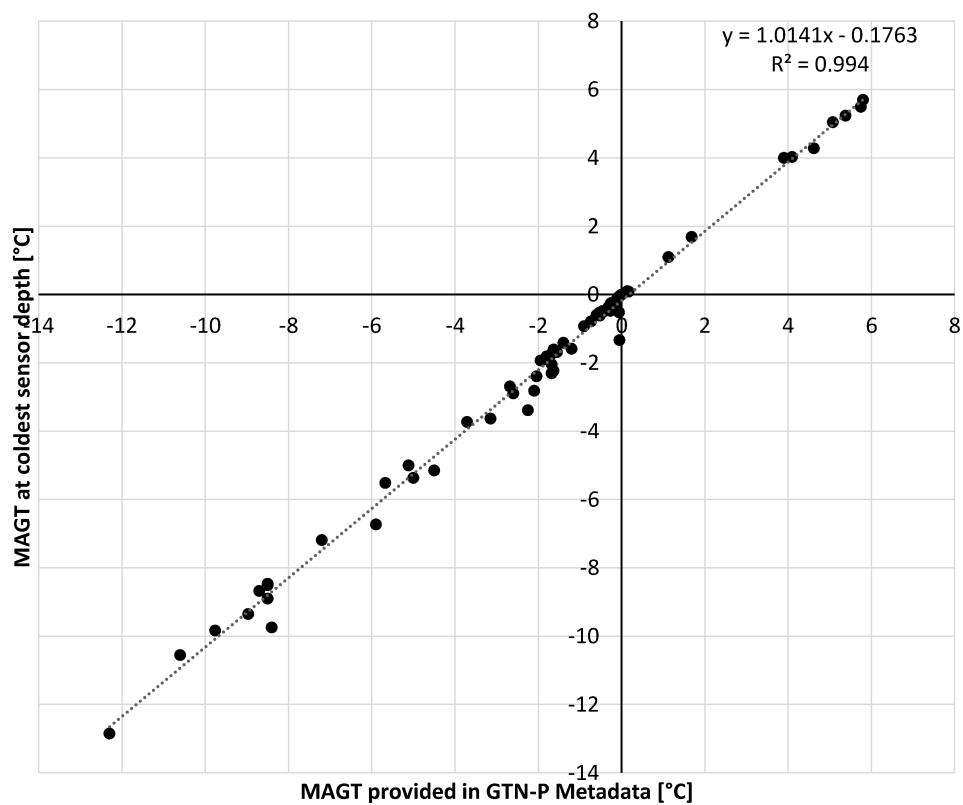



Figure 1. Mean Annual Ground Temperature (MAGT) derived from ~~GTN-P~~ metadata (at or close to zero annual amplitude) versus derived $MAGT_c$ from coldest sensor (for year of MAGT specified in the meta data) based on GTN-P borehole records. Dotted line represents linear fit.

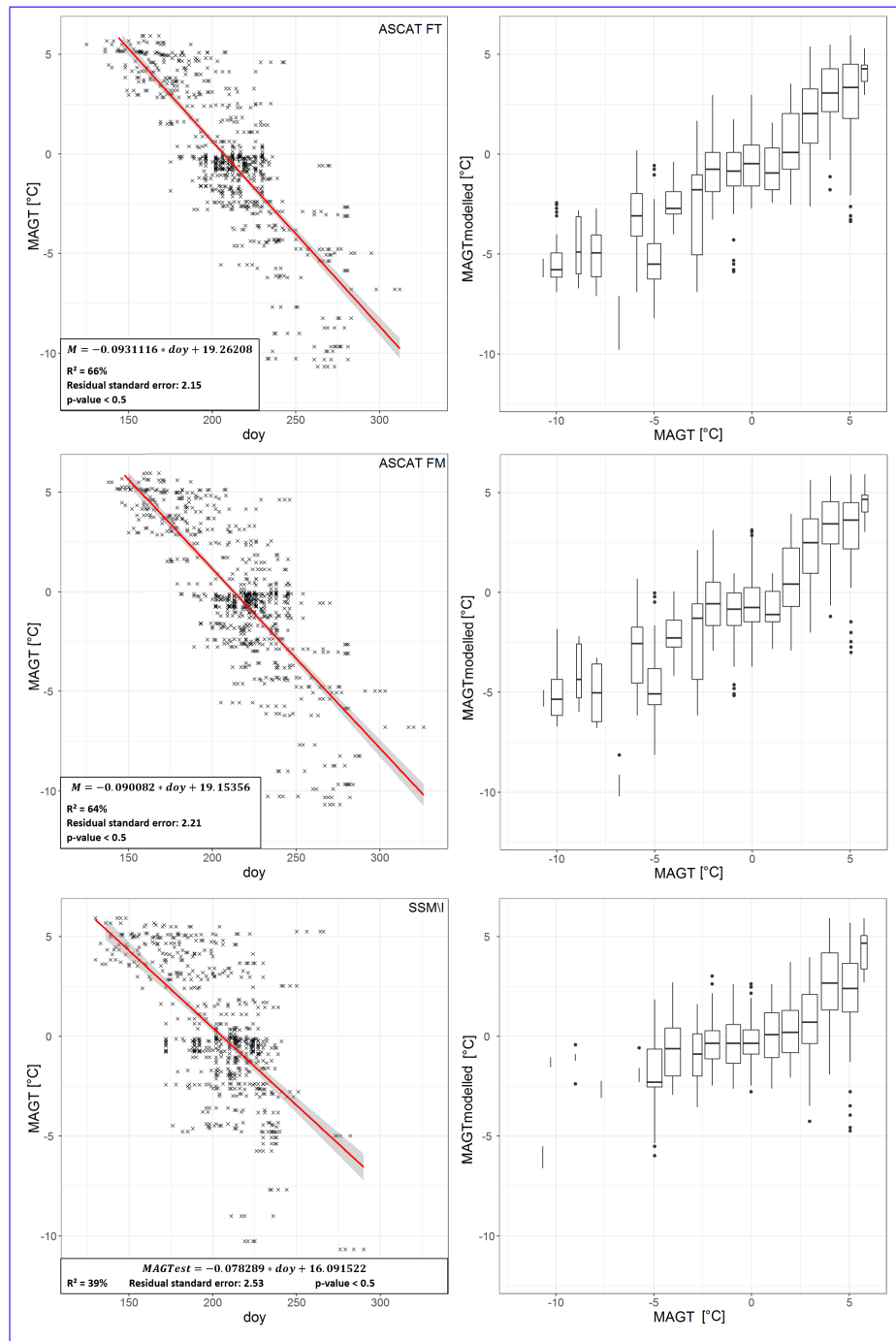



Figure 2. Left: Comparison of number of frozen days (doy - days of year) from satellite records and Mean Annual Ground Temperature for GTN-P boreholes (at depth of coldest sensor, years 2010-2012). Red line represents linear fit. Right: Box plots of modeled versus Mean Annual Ground Temperature from GTN-P boreholes (at depth of coldest sensor, years 2007/8 and 2008/9) Top: ASCAT excluding snow melt days, middle: ASCAT with snow melt days, bottom: SSM/I.

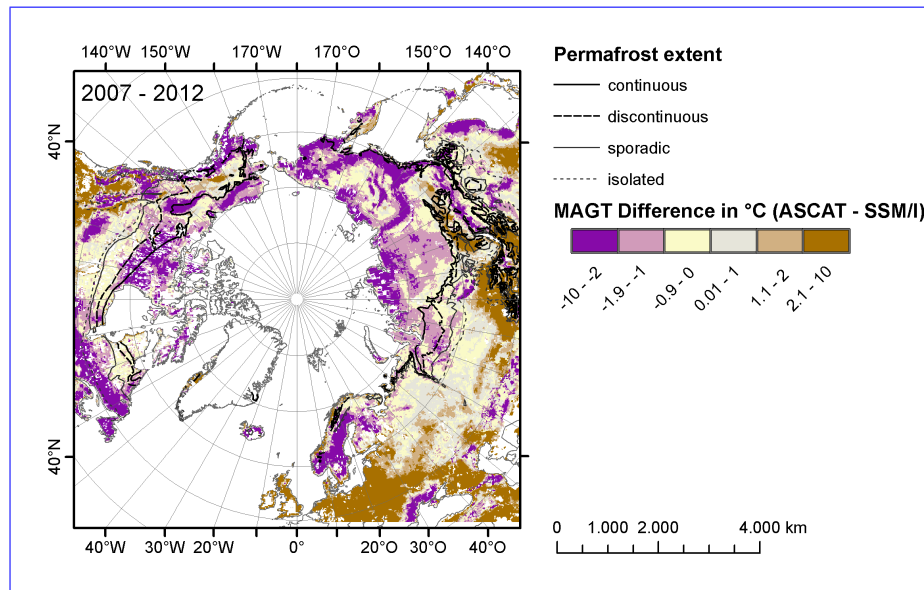



Figure 3. Comparison of MAGT$_c$ (coldest sensor) in °C of all years between Metop ASCAT (FT) and SSM/I for 2007-2012. Source of permafrost extent classes: Brown et al. (1997).

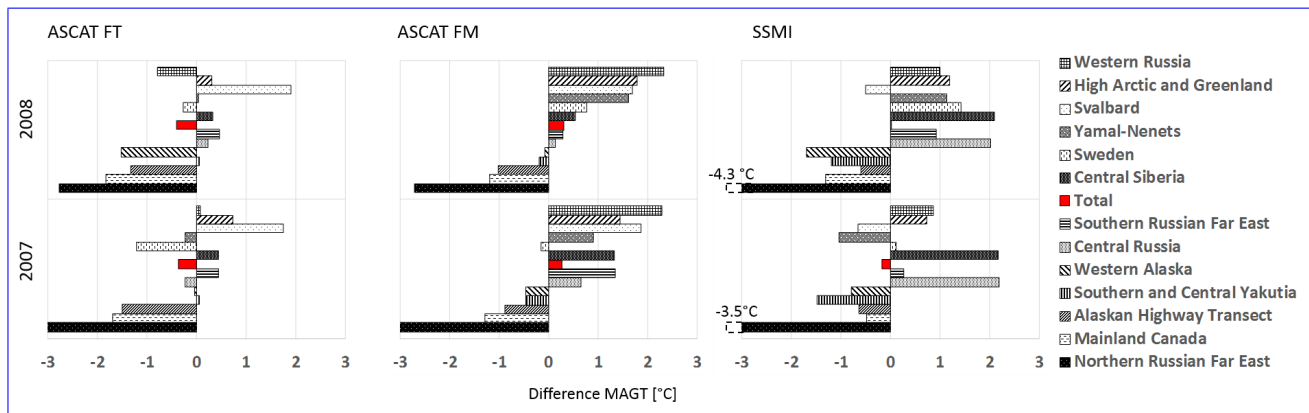



Figure 4. Difference between in situ MAGT$_c$ (coldest sensor) and modelled MAGT$_c$ by region (see Fig. A1) in 2007/8 and 2008/9. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground in ASCAT.

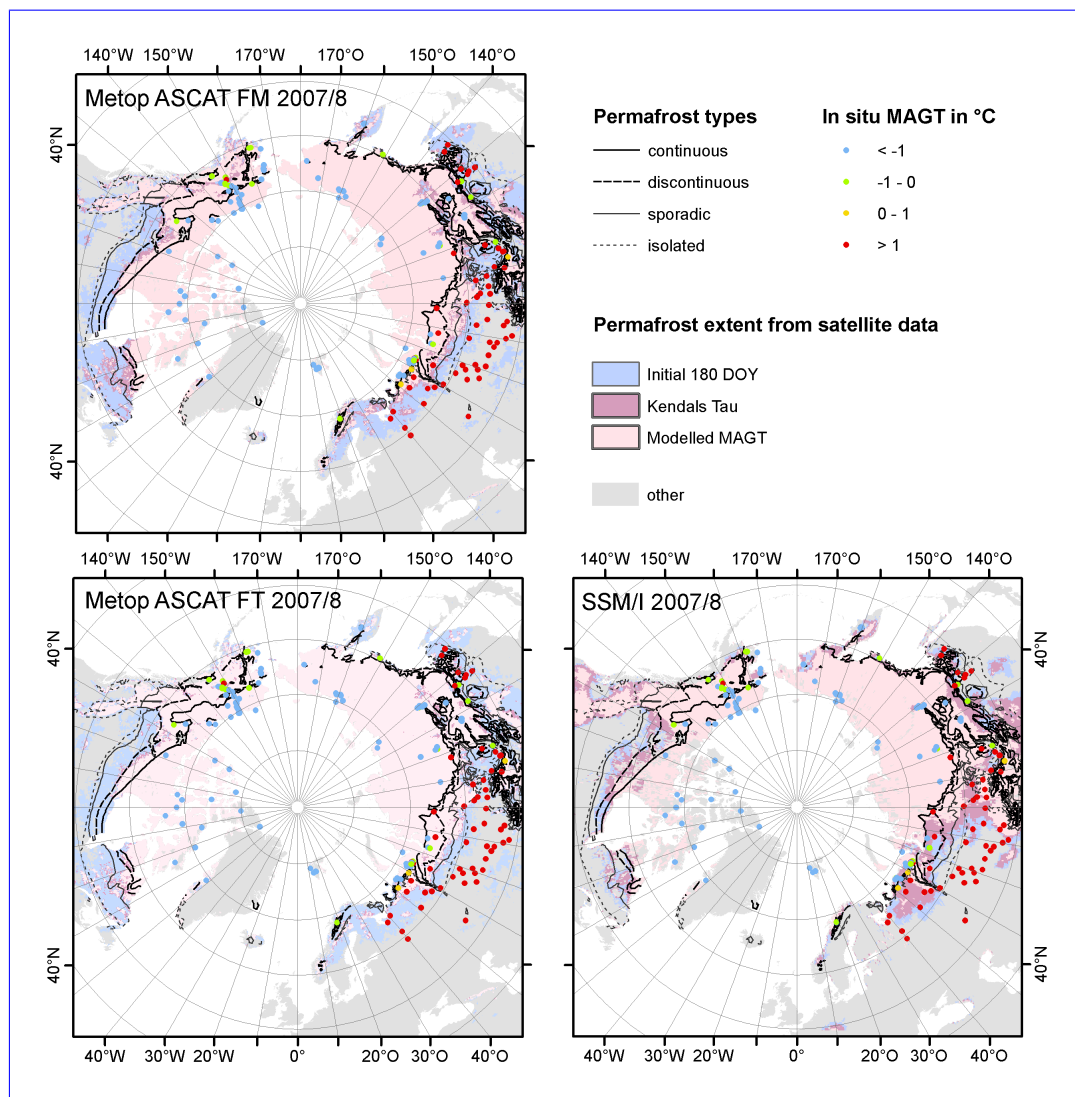



Figure 5. Permafrost extent maps based on thresholds applied to frozen days of year (DOY) to ASCAT excluding melt days (FT), thresholds applied to frozen days of year (DOY) to ASCAT including melt days (FM), and thresholds applied to frozen days of year (DOY) to SSM/I. The initial threshold is 180 days. The value for best Kendall's τ represents the best fit with borehole measurements. The highest threshold has been determined using an empirical model calibrated with borehole measurements. See also Table 1. All satellite results are based on 2007/8-2008/9 records. Source for permafrost extent classes: Brown et al. (1997).

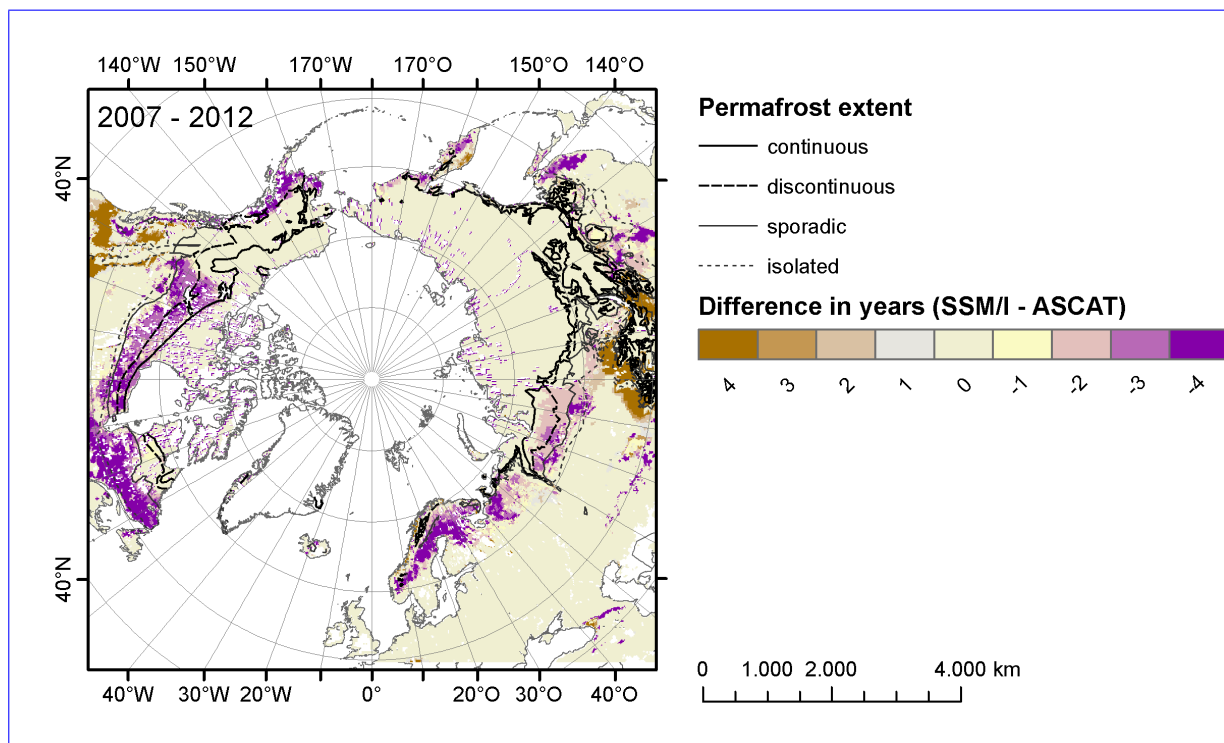
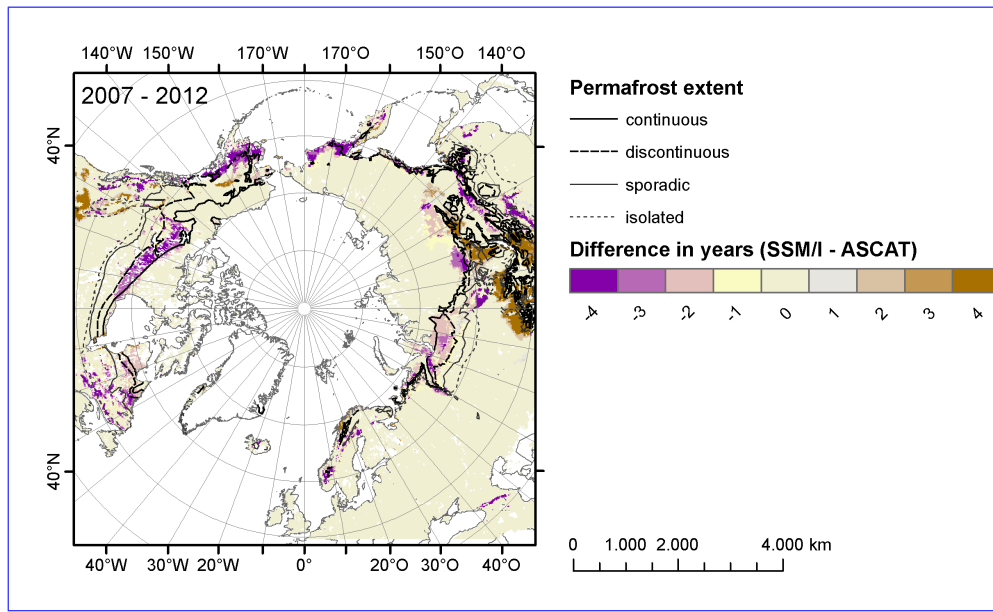



Figure 6. Comparison of the total number of years classified as permafrost between Metop ASCAT and SSM/I for 2008-2012 based on the 180 day threshold method applied to all years (minimum of two consecutive years with at least 180 days frozen).

~~Comparison of satellite data results~~ 2008-2012 based on ~~a 180 day~~ the 0°C temperature threshold method applied to all years (minimum of two consecutive years with below 0°C). Source of permafrost extent ~~from~~ classes: Brown et al. (1997).~~Covered area~~

~~Comparison of satellite data results~~ 2008-2012 based on ~~a 180 day~~ the 0°C temperature threshold method applied to all years (minimum of two consecutive years with below 0°C). Source of permafrost extent ~~from~~ classes: Brown et al. (1997).~~Covered area~~

Figure 7. ~~Permafrost extent comparison~~ Comparison of ~~satellite data results from~~ the total number of years classified as permafrost between Metop ASCAT ~~based on modified thresholds (180, 190, 200 and 210 days) with permafrost extent classes from Brown et al. (1997). Covered area is provided in km^2 within each class. FT~~ ~~days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground with Metop ASCAT.)~~ and SSM/I ~~classification results based on 180 days is included~~ for ~~comparison.~~ ~~Comparison of satellite data results~~ 2008-2012 based on ~~a 180 day~~ the 0°C temperature threshold method applied to all years (minimum of two consecutive years with below 0°C). Source of permafrost extent ~~from~~ classes: Brown et al. (1997).~~Covered area~~

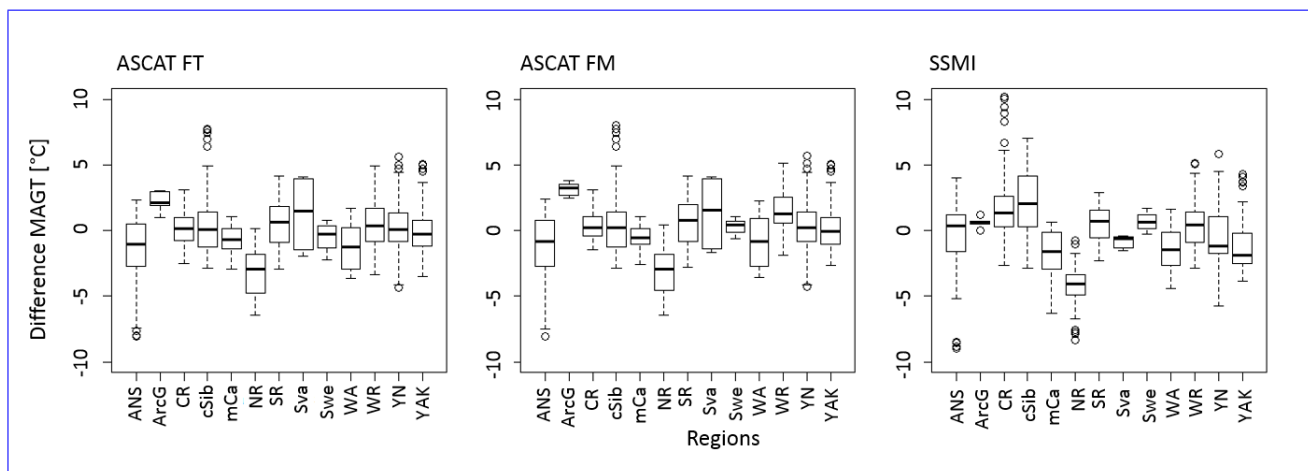



Figure 8. Difference between in situ MAGT_c (coldest sensor) and modelled MAGT_c by region (see Fig. A1 for abbreviations and map) for all years. Only stations which overlap with records from all datasets are considered. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground in ASCAT.

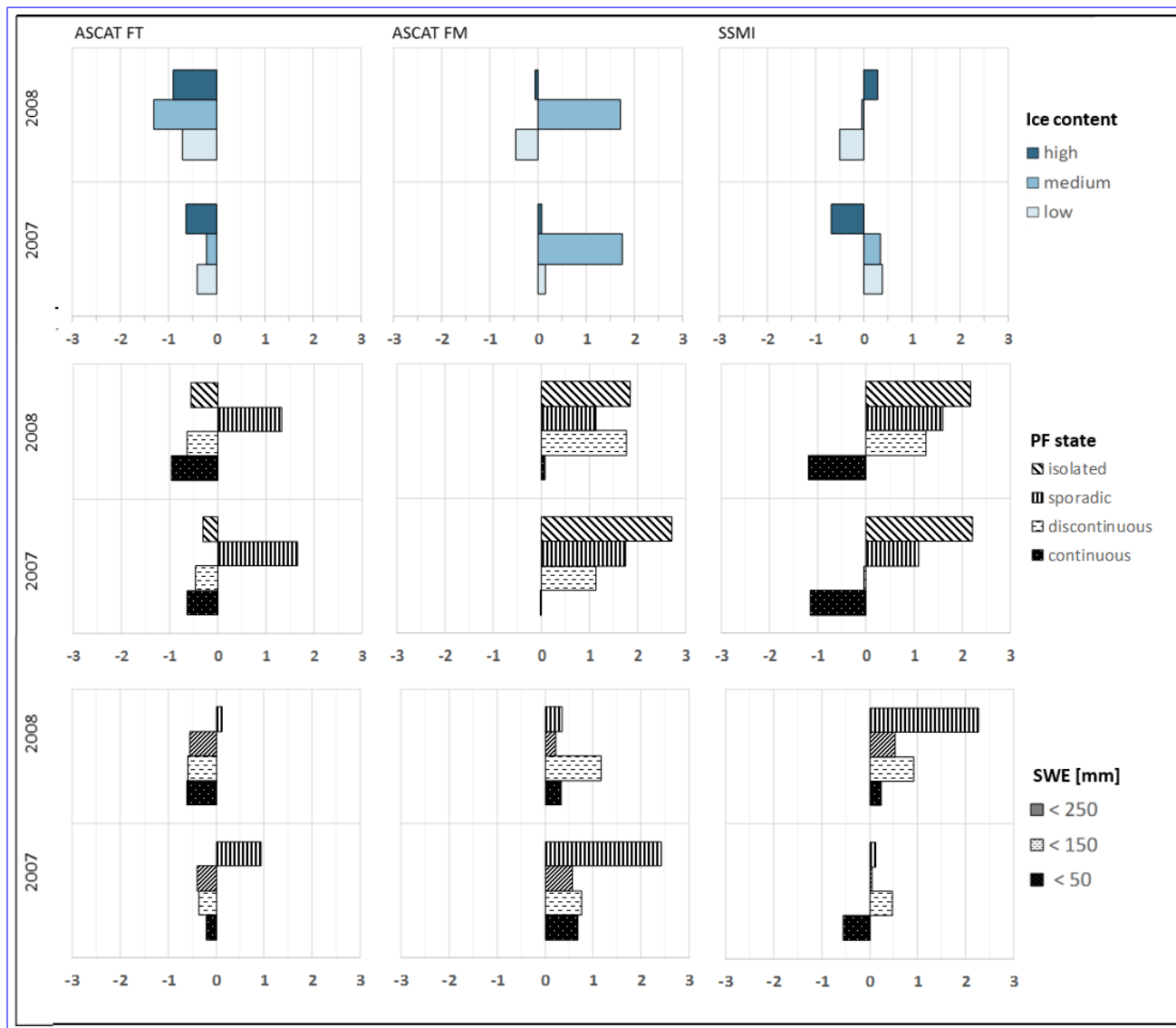



Figure 9. Difference between in ~~% inside~~ situ MAGT$_c$ (coldest sensor) and ~~outside~~ modelled MAGT$_c$ (coldest sensor) by permafrost ~~regions is provided.~~ zone and ice content (source: Brown et al. (1997)), and snow water equivalent (SWE, source: GlobSnow) categories in 2007/8 and 2008/9. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground ~~with Metop~~ in ASCAT.

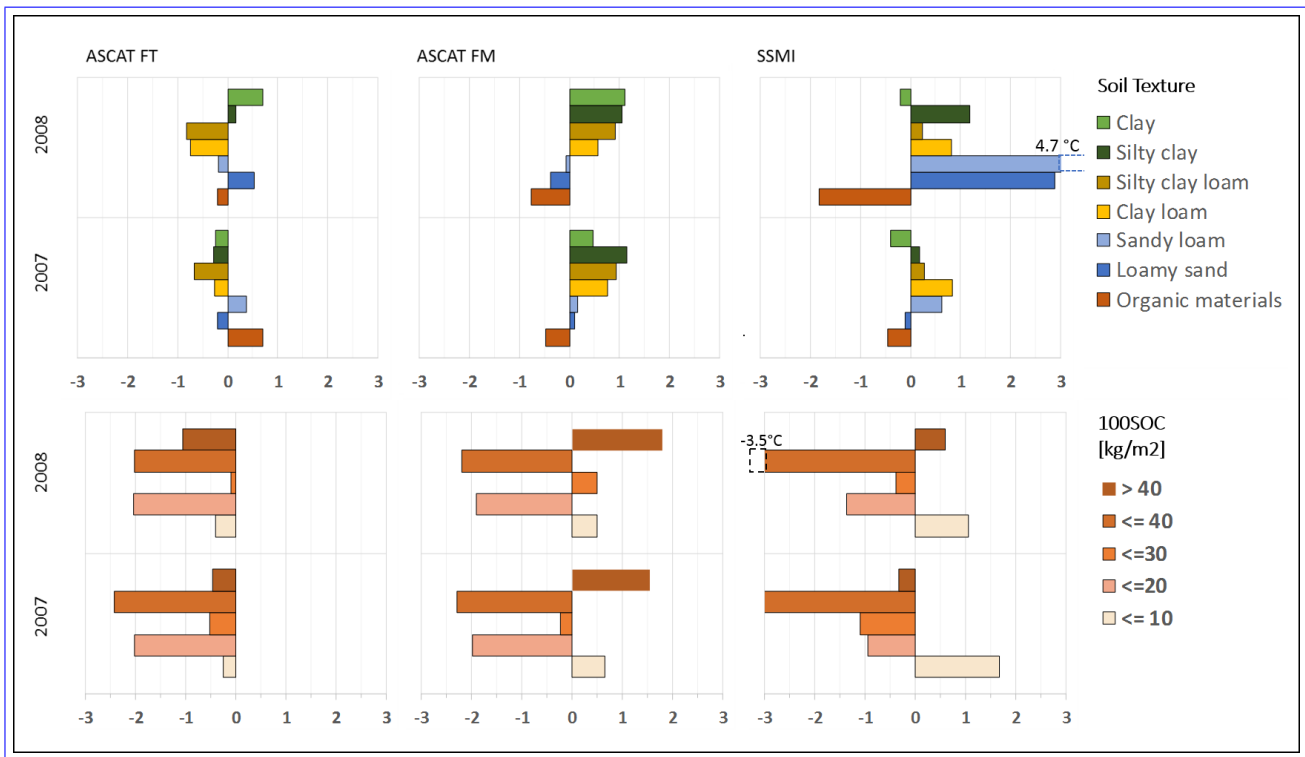



Figure 10. Difference between in situ MAGT$_c$ (coldest sensor) and modelled MAGT$_c$ by soil texture (source: Fischer et al. (2008)) and soil organic carbon content within the top 100 cm (100SOC; source: Hugelius et al. (2013, 2014)) in 2007/8 and 2008/9. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground in ASCAT.

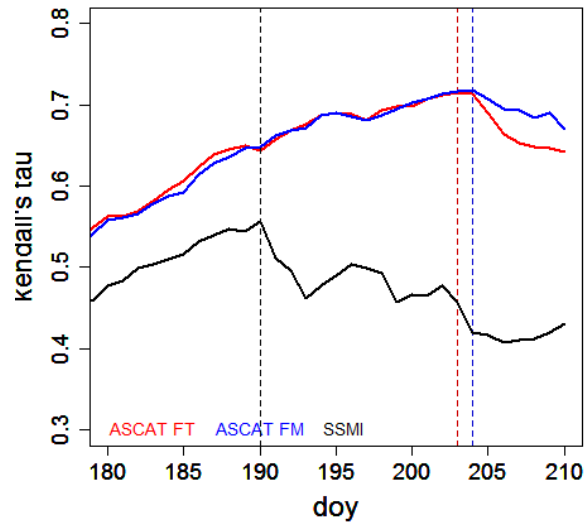



Figure 11. Curves of correlation coefficient (Kendall's τ) between 180 and 210 day of year (doy) frozen from satellite products and positive or negative MAGT at depth of coldest sensor from borehole data. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground with Metop ASCAT.

Left: Comparison of number of frozen days (doy - days of year) from satellite records and Mean Annual Ground Temperature for GTN-P boreholes (at depth of coldest sensor, years 2010-2012). Red line represents linear fit. Right: Box plots of modeled versus Mean Annual Ground Temperature from GTN-P boreholes (at depth of coldest sensor, years 2007/8 and 2008/9) Top: ASCAT excluding snow melt days, middle: ASCAT with snow melt days, bottom: SSM/I.

5   Permafrost extent maps based on thresholds applied to frozen days of year (DOY) to ASCAT excluding melt days (FT), thresholds applied to frozen days of year (DOY) to ASCAT including melt days (FM), and thresholds applied to frozen days of year (DOY) to SSM/I. The initial threshold is 180 days. The value for best Kendall's τ represents the best fit with borehole measurements. The highest threshold has been determined using an empirical model calibrated with borehole measurements. See also Table 1. All satellite results are based on 2007/8-2008/9 records. Source for permafrost extent classes: Brown et al. (1997).

10   Comparison of MAGT in °C of all years between Metop ASCAT (FT) and SSM/I for 2007-2012. Source of permafrost extent classes: Brown et al. (1997).

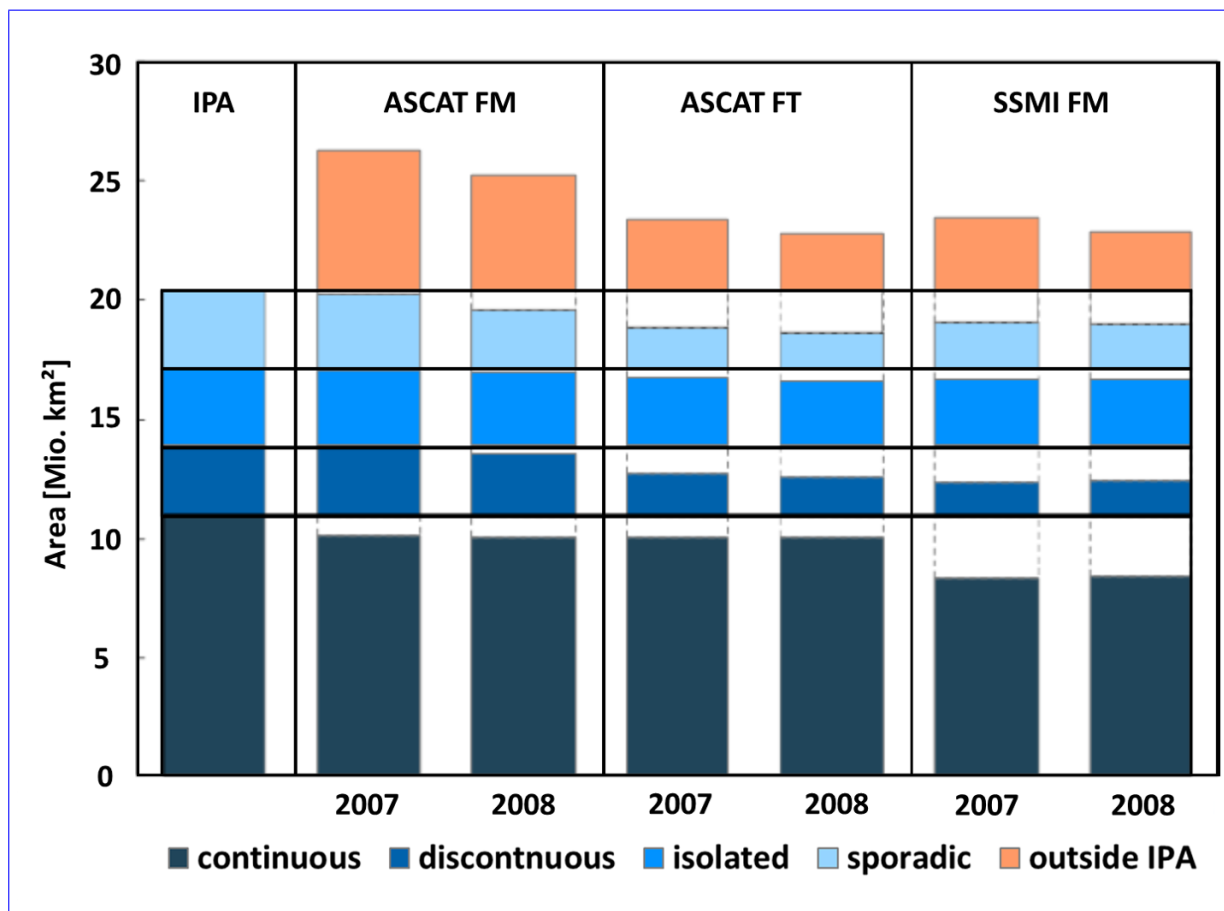



Figure 12. Permafrost extent comparison of satellite data results based on modelled ground temperatures (LM - Linear Model) with permafrost classes from Brown et al. (1997). Covered area in km^2 within each class and outside. FT - days identified as frozen with ASCAT, FM - days with melting snow are considered as frozen ground with ASCAT.

~~Difference between in situ MAGT (coldest sensor) and modelled MAGT by region (see Fig. A1) in 2007/8 and 2008/9. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground in ASCAT.~~
~~Comparison of the total number of years classified as permafrost between Metop ASCAT (FT) and SSM/I for 2008-2012 based on the zero degree temperature threshold method applied to all years (minimum of two consecutive years with below zero degree). Source of permafrost~~
5 ~~extent classes: Brown et al. (1997).~~

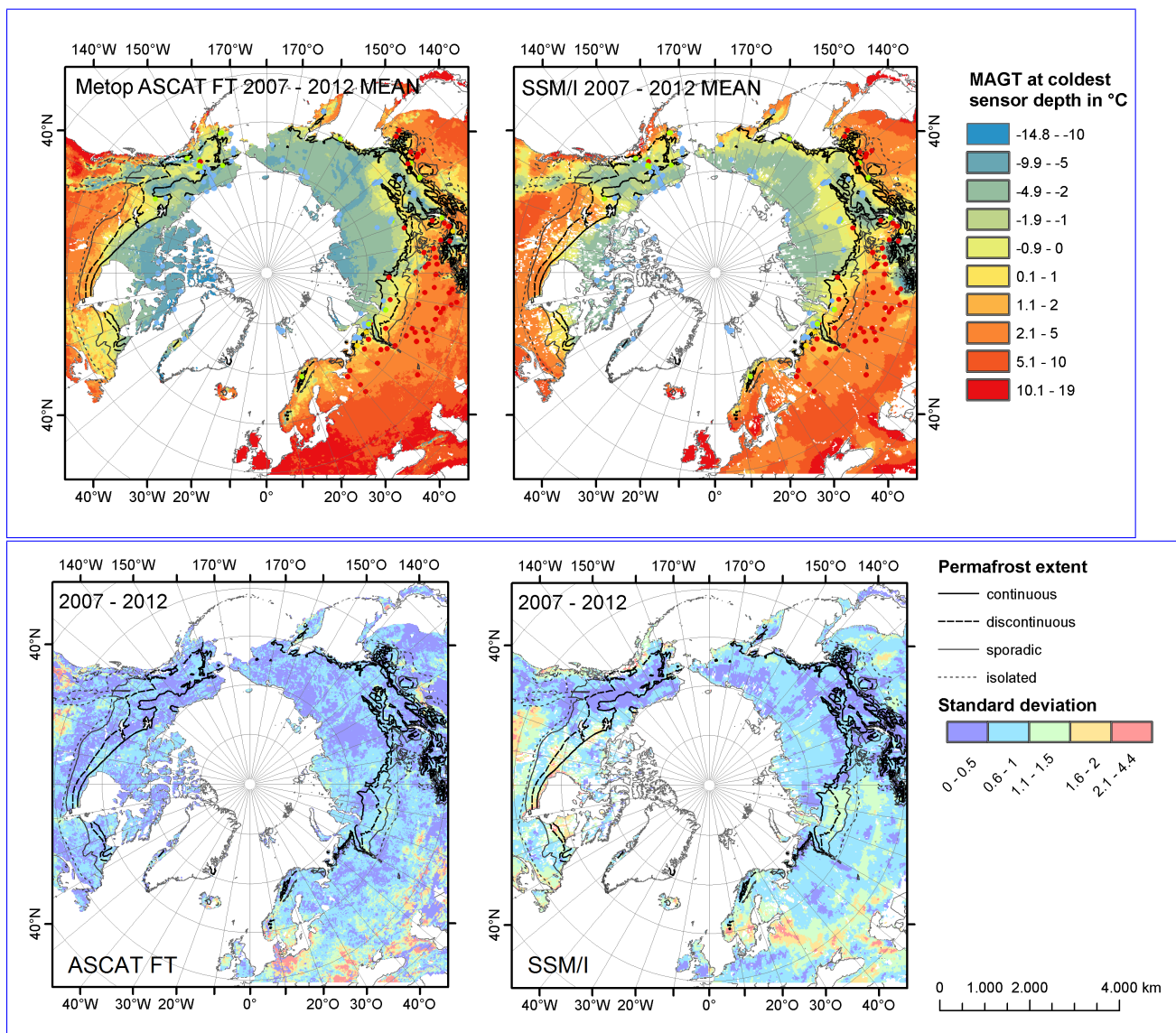



Figure 13. ~~Comparison~~ Maps of ~~satellite data results based on~~ modelled mean annual ground ~~temperatures with~~ temperature (top) and standard deviation of modelled MAGT$_c$ (bottom) for ASCAT excluding melt days (FT; left) and SSM/I (right) based on all analyzed years. Source for permafrost extent ~~from~~ classes: Brown et al. (1997).~~Covered area in % inside and outside permafrost regions. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground, LM - model results.~~

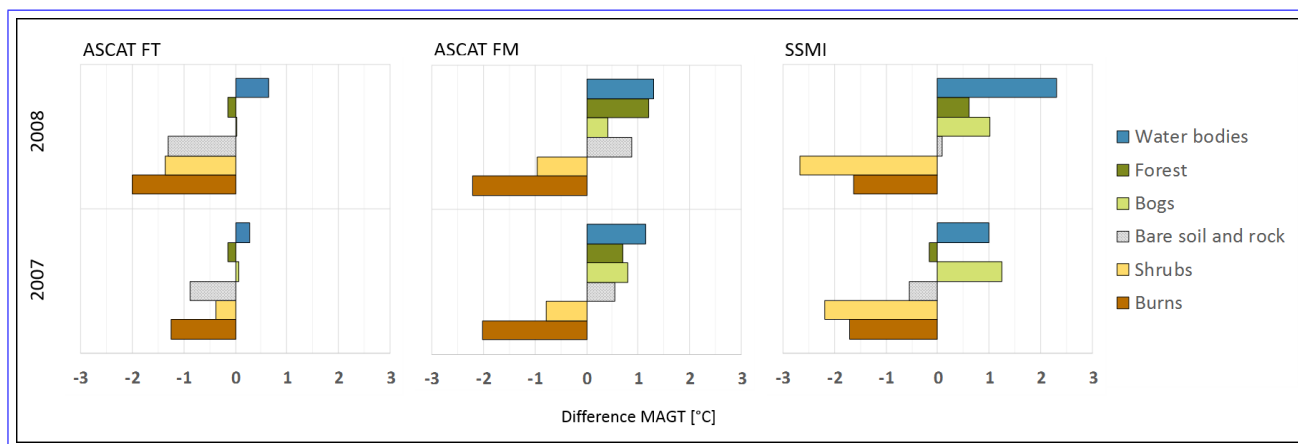



Figure 14. Difference between in situ MAGT$_c$ (coldest sensor) and modelled MAGT$_c$ by land cover type (source: GLC2000 (2003)) in 2007/8 and 2008/9. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground in ASCAT.

Table 1. Comparison between in situ Mean Annual Ground Temperature (MAGT$_c$, at coldest sensor depth) with number of frozen days per year from ASCAT (FT - excluding snow melt days, FM - including snow melt days) and SSM/I for 2007-2012.

Parameter	ASCAT FT	ASCAT FM	SSM/I
Pearson correlation for linear fit	0.66	0.64	0.39
RMSE validation years (2007/2008)	2.15	2.21	2.53
RMSE all years (2007 - 2011)	2.22	2.24	2.57
DOY for 0°C from linear fit	207	212	205
DOY for best Kendall's τ	203	204	190

Table 2. Permafrost extent comparison of satellite data results from Metop ASCAT based on modified thresholds (180, 190, 200 and 210 days) with permafrost extent classes from Brown et al. (1997). Covered area is provided in km^2 within each class. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground with Metop ASCAT. SSM/I classification results based on 180 days is included for comparison.

	IPA	FT 180	FT 190	FT 200	FT 210	FM 180	FM 190	FM 200	FM 210	SSMI 180
continuous [Mio. km^2]	10.90	10.31	10.25	10.16	9.93	10.31	10.26	9.96	9.97	8.49
(% of IPA)		(95%)	(94%)	(93%)	(91%)	(95%)	(94%)	(91%)	(92%)	(78%)
discontinuous [Mio. km^2]	2.97	2.17	2.13	1.49	0.86	2.73	2.20	1.57	1.03	2.11
(% of IPA)		(26%)	(20%)	(14%)	(8%)	(25%)	(20%)	(14%)	(9%)	(19%)
isolated [Mio. km^2]	3.24	3.10	2.94	2.70	2.21	3.11	2.96	2.74	2.32	2.74
(% of IPA)		(28%)	(27%)	(25%)	(20%)	(28%)	(27%)	(25%)	(21%)	(25%)
sporadic [Mio. km^2]	3.29	2.54	1.95	1.35	0.73	2.58	2.02	1.39	0.87	1.71
(% of IPA)		(23%)	(18%)	(12%)	(7%)	(24%)	(19%)	(13%)	(8%)	(16%)

Difference between in situ MAGT (coldest sensor) and modelled MAGT by region (see Fig. A1 for abbreviations and map) for all years. Only stations which overlap with records from all datasets are considered. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground in ASCAT.

Table 3. Comparison of satellite data results based on a 180 day threshold with permafrost extent from Brown et al. (1997). Covered area in % inside and outside permafrost regions is provided. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground with Metop ASCAT.

	ASCAT FT 180	ASCAT FM 180	SSMI FM 180
inside IPA (% of IPA)	20.40 (100%)	17.41 (85%)	12.54 (61%)
outside IPA (% of IPA)	0.00 (0%)	-1.97 (-10%)	-1.09 (-5%)

Difference between in situ MAGT (coldest sensor) and modelled MAGT by permafrost zone and ice content (source: Brown et al. (1997)), and snow water equivalent (SWE, source: GlobSnow) categories in 2007/8 and 2008/9. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground in ASCAT.

Table 4. ~~Difference~~ Comparison matrix between in situ ~~MAGT~~ Mean Annual Ground Temperature (MAGT$_c$, at coldest sensor depth) ~~and modelled MAGT by land cover type~~ classes (~~source: GLC2000 (2003)~~below or above zero °C) ~~in 2007/8~~ and ~~2008/9.~~ classified number of frozen days per year from ASCAT (FT - ~~days identified as frozen without melting~~ excluding snow ~~are used~~melt days, FM - ~~days with melting~~ including snow melt days) and SSM/I for 2007-2012. Both, the initial 180 Days of Year (DOY) threshold and the optimal threshold based on best Kendall's τ are ~~considered as frozen ground~~ assessed. Values represent numbers of sites for multiple years. Rows indicating false classifications are in ~~ASCAT~~*italic*.

Sensor	ASCAT FT		ASCAT FM		SSM/I	
DOY	180	203	180	204	180	190
Positive MAGT$_c$ & DOY <thres.	142	211	127	199	115	148
Negative MAGT$_c$ & DOY <thresh.	*2*	*70*	*0*	*35*	*14*	*41*
Positive MAGT$_c$ & DOY >= thresh.	*138*	*69*	*153*	*81*	*157*	*124*
Negative MAG$_c$ & DOY >= thresh.	494	426	496	461	398	371
Sum of data points	776	776	776	776	684	684

Table 5. ~~Difference between in situ MAGT (coldest sensor) and~~ Comparison of satellite data results based on modelled ~~MAGT by soil texture (source: Fischer et al. (2008)) and soil organic carbon content within the top 100 cm (100SOC; source: Hugelius et al. (2013, 2014))~~ ground temperatures with permafrost extent from Brown et al. (1997). Covered area in ~~2007/8~~ % inside and ~~2008/9.~~ outside permafrost regions. FT - days identified as frozen without melting snow are used, FM - days with melting snow are considered as frozen ground~~in ASCAT~~, LM - model results.

	FM LM 2007	FM LM 2008	FT LM 2007	FT LM 2008	SSMI LM 2007	SMMI LM 2008
inside IPA	19.15	18.10	16.44	18.10	14.50	14.55
(% of IPA)	(94%)	(89%)	(81%)	(89%)	(71%)	(71%)
outside IPA	5.90	4.78	2.98	4.78	3.08	2.40
(% of IPA)	(-29%)	(-23%)	(-15%)	(-23%)	(-15%)	(-12%)

~~Maps of modelled mean annual ground temperature (top) and standard deviation of modelled MAGT (bottom) for ASCAT excluding melt days (FT; left) and SSM/I (right) based on all analyzed years. Source for permafrost extent classes: Brown et al. (1997).~~

Appendix A

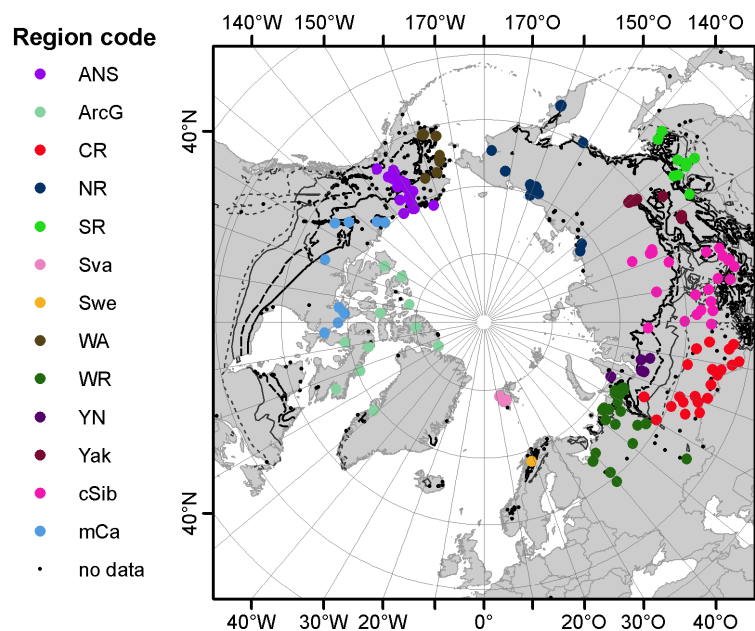



Figure A1. Map of used GTN-P boreholes with region class. 'No data' refers to sites without publicly available data or sites which failed the selection criteria. ANS - Alaskan Highway transect and North Slope, ArcG - Canadian High Arctic and Greenland, NR - Northern Russian Far East, SR - Southern Russian Far East, Sva - Svalbard, Swe - Sweden, WA - Western Alaska, WR - Western Russia, YN - Yamal-Nenets district, Yak - central and southern Yakutia, cSib - central Siberia, mCa - Mainland Canada. For legend of permafrost extent (line features) see Fig. 13.

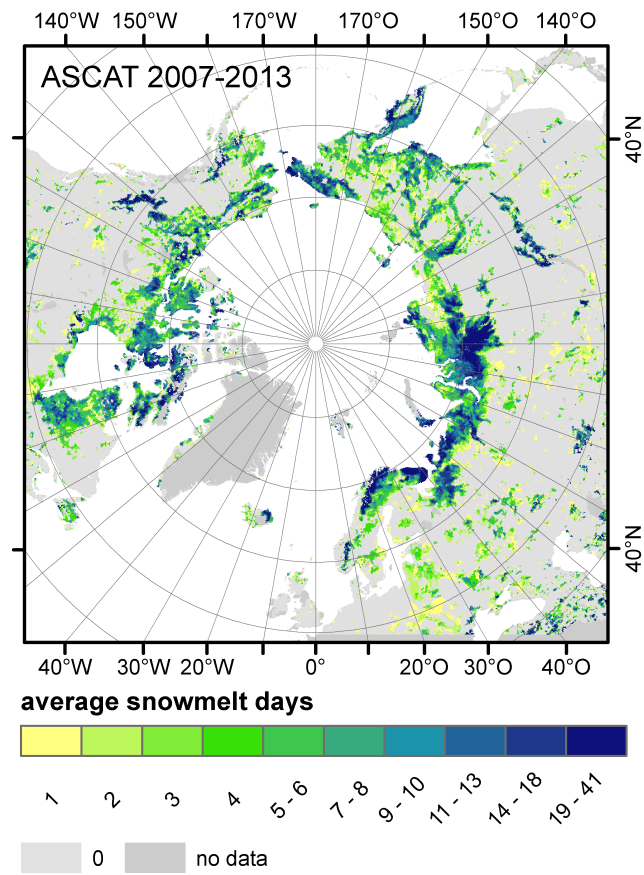



Figure A2. Average number of snow melting days per year from Metop ASCAT for 2007 – 2013 derived from Paulik et al. (2014).