CCI+ PHASE 1 – NEW ECVS
PERMAFROST

D2.5 PRODUCT VALIDATION PLAN (PVP)

VERSION 2.0

30 NOVEMBER 2019

PREPARED BY

Document Status Sheet

Issue	Date	Details	Authors
1.0	28.02.2019	Final version WP 210	BH, MW, CP, RD, CB, SW, AB
2.0	30.11.2019	Update in situ data to be used, summary of year 1 validation results Final version WP 220	BH, MW, AB

Author team

Birgit Heim, AWI

Mareike Wieczorek, AWI

Cécile Pellet, UNIFR

Reynald Delaloye, UNIFR

Chloé Barboux, UNIFR

Sebastian Westermann, GUIO

Annett Bartsch, B.GEOS

Tazio Strozzi, GAMMA

ESA Technical Officer:

Frank Martin Seifert

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The required parameters by the Global Climate Observing System (GCOS) for the Essential Climate Variable (ECV) Permafrost are permafrost temperature and active layer thickness. Since Earth Observation (EO)-based algorithms operate at a km-spatial scale, the products of Permafrost_cci must be set in context with the spatial variability of permafrost temperatures and active layer thickness. In many permafrost regions, these can have a high variability at spatial scales of meters, which is much finer than the footprint of EO sensors. An additional parameter is permafrost extent (fraction), the fraction within an area (pixel) at which the definition for the existence of permafrost (ground temperature <0 ºC for two consecutive years) is fulfilled. The variables to be retrieved and validated in Permafrost_cci comprise therefore: a) permafrost temperature, b) active layer thickness and c) permafrost fraction. The generation relies on a ground thermal model (CryoGrid CCI) that is forced by EOs Land Surface Temperature (LST) and Snow Water Equivalent (SWE) with boundary conditions of EO Land-Cover.

A critical step in the acceptance of the Permafrost_cci products is to provide reliable information on the product quality. The Committee on EO Satellites-Working Group on Calibration and Validation (CEOS-WGCV) defines validation as 'The process of assessing, by independent means, the quality of the data products derived from the system outputs' (http://lpvs.gsfc.nasa.gov/). The CEOS Quality Assurance framework for Earth Observation (QA4EO) provides guidelines, e.g. on Fiducial Reference Measurements (FRM) using meteorological standards. However, commonly, for geoscientific EO applications, accuracy can only be measured in terms of an agreement, or in terms of omission and commission errors. Therefore, if validation against precise FRM is not possible, validation against suitable in-situ reference data or against other sources using expert knowledge is feasible. According to QA4EO, the validation data need to be independent from the retrieval process of the product. In the QA4EO sense, suitable validation data are characterised by protocols and community-wide management practices, and published openly. The validation data shall be a part of a collaborative user environment within an international framework.

The World Meteorological Organisation (WMO) and GCOS delegated the global ground-based monitoring of the ECV Permafrost to the Global Terrestrial Network for Permafrost (GTN-P) managed by the International Permafrost Association (IPA). GTN-P/IPA established the Thermal State of Permafrost Monitoring (TSP) for permafrost temperature monitoring and the Circumpolar Active Layer Monitoring program (CALM) for active layer thickness monitoring. Both GTN-P monitoring programs, TSP and CALM, require standards for measurements and data collection and publish data sets (Biskaborn et al. 2015, 2019).

The validation in Permafrost_cci is fully independent as the validation team is independent of the algorithm development team and uses the global GTN-P data and monitoring networks such as the mountain permafrost monitoring program PERMOS in Switzerland and the meteorological monitoring network ROSHYDROMET in Russia. The characterisation of errors and uncertainties is carried out with comparative statistics, using bias, absolute error, relative percentage error and Root Mean Square Error (RMSE) as well as a qualitative assessment of permafrost properties and value ranges in specific permafrost landscapes. In situ data sets on ground temperature and active layer depth are available over longer time periods and for different permafrost conditions, land-cover, topography and climate. This will allow extending the validation to characteristics important for climate research, e.g. inter-annual variability.

1 INTRODUCTION

1.1 Purpose of the document

This document is the Product Validation Plan (PVP) version 2 (update of [RD-1]) of the ESA project Permafrost_cci. The PVP describes and defines the reference data sets, and the validation methods and strategies used for the validation of the Permafrost_cci product following CCI and QA4EO guidelines.

The variables of the Permafrost_cci product comprise the ECVs 'permafrost temperature' and 'active layer thickness'. Mean annual ground temperature (MAGT) forms also the basis for the variable 'permafrost extent (fraction)'. The generation of ground temperature and active layer thickness time-series relies on the ground thermal model CryoGrid CCI that is forced by EO-derived time series of LST and SWE with boundary conditions of EO land-cover [RD-2]. The Permafrost_cci PVP focuses on the validation of permafrost temperature and permafrost fraction in lowlands and mountain permafrost regions and active layer thickness in lowland permafrost regions in the Arctic and Asia, and mountain permafrost regions such as in the Alps and Central Asia. The project also undertakes validation experiments for mountain permafrost areas using rockglacier abundance and kinematics, and geophysical information or binary-based validation on permafrost abundance. Validation results according to the data and methods described in the PVP are presented in [RD-3]

1.2 Structure of the document

Chapter 2 provides information on how the validation follows the overall project guidelines of CCI. In chapter 3, unbiased validation, including independency of the validation from the algorithm development team and validation criteria are described. The validation activities are presented in chapter 4, giving an overview on the validation data from GTNP/TSP ground temperature and CALM active layer thickness time series in lowland permafrost areas and the Swiss mountain permafrost monitoring network PERMOS. Chapter 4 also describes the compilation of discrete and interpolated ground temperature-depth time series for validation, the collection of ambient metadata on vegetation, ground ice content and lithology, and the strategies for validation using match-up techniques and functional validation. Information on Permafrost_cci deliverables with validation activities and validation output is given in chapter 5.

1.3 Applicable Documents

[AD-1] ESA 2017: Climate Change Initiative Extension (CCI+) Phase 1 – New Essential Climate Variables - Statement of Work. ESA-CCI-PRGM-EOPS-SW-17-0032

[AD-2] Requirements for monitoring of permafrost in polar regions - A community white paper in response to the WMO Polar Space Task Group (PSTG), Version 4, 2014-10-09. Austrian Polar Research Institute, Vienna, Austria, 20 pp

[AD-3] ECV 9 Permafrost: assessment report on available methodological standards and guides, 1 Nov 2009, GTOS-62

[AD-4] GCOS-200, the Global Observing System for Climate: Implementation Needs (2016 GCOS Implementation Plan, 2015.

[AD-5] GEO/CEOS Quality Assurance framework for Earth Observation (QA4EO) protocols 3-4

[AD-6] ESA Climate Change Initiative. CCI Project Guidelines. EOP-DTEX-EOPS-SW-10-0002

1.4 Reference Documents

[RD-1] Bartsch, A., Heim, B., Wieczorek, M., Pellet, C., Barboux, C., Delaloye, R.,Strozzi, T. (2019): ESA CCI+ Product Validation Plan, v1.0

[RD-2] Bartsch, A.; Grosse, G.; Kääb, A.; Westermann, S.; Strozzi, T.; Wiesmann, A.; Duguay, C.; Seifert, F. M.; Obu, J.; Goler, R.: GlobPermafrost – How space-based earth observation supports understanding of permafrost. Proceedings of the ESA Living Planet Symposium, pp. 6.

[RD-3] Heim, B., Wieczorek, M., Pellet, C., Barboux, C., Delaloye, R., Bartsch, A., B. Kroisleitner, C., Strozzi, T. (2019): ESA CCI+ Product Validation and Intercomparison Report, v1.0

[RD-4] Bartsch, A., Matthes, H., Westermann, S., Heim, B., Pellet, C., Onacu, A., Kroisleitner, C., Strozzi, T.(2019): ESA CCI+ Permafrost User Requirements Document, v1.0

[RD-5] Bartsch, A., Westermann, Strozzi, T., Wiesmann, A., Kroisleitner, C. (2019): ESA CCI+ Permafrost Product Specifications Document, v1.0

[RD-6] IPA Action Group 'Specification of a Permafrost Reference Product in Succession of the IPA Map' (2016): Final report.
https://ipa.arcticportal.org/images/stories/AG_reports/IPA_AG_SucessorMap_Final_2016.pdf

[RD-7] Nitze, I., Grosse, G., Heim, B., Wieczorek, M., Matthes, H., Bartsch, A., Strozzi, T. (2019): ESA CCI+ Climate Assessment Report (CAR), v1.0

[RD-8] Bartsch, A., Westermann, S., Heim, B., Wieczorek, M., Pellet, C., Barboux, C., Kroisleitner, C., Strozzi, T. (2019): ESA CCI+ Permafrost Data Access Requirements Document, v1.0

[RD-9] Westermann, S., Bartsch, A., Strozzi, T. (2019): ESA CCI+ Product Validation and Assessment Report, v1.0

1.5 Bibliography

A complete bibliographic list that support arguments or statements made within the current document is provided in Section 6.1.

1.6 Acronyms

A list of acronyms is provided in section 6.2.

1.7 Glossary

The list below provides a selection of terms relevant for the parameters addressed in Permafrost_cci [RD-4]. A comprehensive glossary is available as part of the Product Specifications Document [RD-5].

active-layer thickness
The thickness of the layer of the ground that is subject to annual thawing and freezing in areas underlain by permafrost.
The thickness of the active layer depends on such factors as the ambient air temperature, vegetation, drainage, soil or rock type and total water con-tent, snow-cover, and degree and orientation of slope. As a rule, the active layer is thin in the High Arctic (it can be less than 15 cm) and becomes thicker farther south (1 m or more).
The thickness of the active layer can vary from year to year, primarily due to variations in the mean annual air temperature, distribution of soil moisture, and snow-cover.
The thickness of the active layer includes the uppermost part of the permafrost wherever either the salinity or clay content of the permafrost allows it to thaw and refreeze annually, even though the material remains cryotic ($T < 0$ °C).
Use of the term "depth to permafrost" as a synonym for the thickness of the active layer is misleading, especially in areas where the active layer is separated from the permafrost by a residual thaw layer, that is, by a thawed or non-cryotic ($T > 0$ °C) layer of ground.
REFERENCES: Muller, 1943; Williams, 1965; van Everdingen, 1985

continuous permafrost
Permafrost occurring everywhere beneath the exposed land surface throughout a geographic region with the exception of widely scattered sites, such as newly deposited unconsolidated sediments, where the climate has just begun to impose its influence on the thermal regime of the ground, causing the development of continuous permafrost.
For practical purposes, the existence of small taliks within continuous permafrost has to be recognised. The term, therefore, generally refers to areas where more than 90 percent of the ground surface is underlain by permafrost.
REFERENCE: Brown, 1970.

discontinuous permafrost
Permafrost occurring in some areas beneath the exposed land surface throughout a geographic region where other areas are free of permafrost.
Discontinuous permafrost occurs between the continuous permafrost zone and the southern latitudinal limit of permafrost in lowlands. Depending on the scale of mapping, several subzones can often be distinguished, based on the percentage (or fraction) of the land surface underlain by permafrost, as shown in the following table.

Permafrost	English usage	Russian Usage
Extensive	65-90 %	Massive Island
Intermediate	35-65 %	Island
Sporadic	10-35 %	Sporadic
Isolated Patches	0-10 %	-

SYNONYMS: (not recommended) insular permafrost; island permafrost; scattered permafrost.
REFERENCES: Brown, 1970; Kudryavtsev, 1978; Heginbottom, 1984; Heginbottom and Radburn, 1992; Brown et al., 1997.

mean annual ground temperature (MAGT)
Mean annual temperature of the ground at a particular depth.
The mean annual temperature of the ground usually increases with depth below the surface. In some northern areas, however, it is not un-common to find that the mean annual ground temperature decreases in the upper 50 to 100 metres below the ground surface as a result of past changes in surface and climate conditions. Below that depth, it will increase as a result of the geothermal heat flux from the interior of the earth. The mean annual ground temperature at the depth of zero annual amplitude is often used to assess the thermal regime of the ground at various locations. van Everdingen, 1998

permafrost
Ground (soil or rock and included ice and organic material) that remains at or below 0 °C for at least two consecutive years.
Permafrost is synonymous with perennially cryotic ground: it is defined on the basis of temperature. It is not necessarily frozen, because the freezing point of the included water may be depressed several degrees below 0 °C; moisture in the form of water or ice may or may not be present. In other words, whereas all perennially frozen ground is permafrost, not all permafrost is perennially frozen. Permafrost should not be regarded as permanent, because natural or man-made changes in the climate or terrain may cause the temperature of the ground to rise above 0 °C.
Permafrost includes perennial ground ice, but not glacier ice or icings, or bodies of surface water with temperatures perennially below 0 °C; it does include man-made perennially frozen ground around or below chilled pipelines, hockey arenas, etc.
Russian usage requires the continuous existence of temperatures below 0 °C for at least three years, and also the presence of at least some ice.
SYNONYMS: perennially frozen ground, perennially cryotic ground and (not recommended) biennially frozen ground, climafrost, cryic layer, permanently frozen ground.
REFERENCES: Muller, 1943; van Everdingen, 1976; Kudryavtsev, 1978.

2 CCI PROJECT GUIDELINES

A critical step in the acceptance of the CCI products by GCOS and Climate Science communities is providing confidence in the quality of each CCI product and its uncertainties through validation against independent data such as in situ reference measurements or alternate estimates. In general, CCI recommendations focus on validation of Level 2 products. Further discussions and developments are still needed regarding validation of Level 1, Level 3 and Level 4 products. Permafrost_cci will produce a Level 4 product simulated by a ground thermal model that is forced by EO products.

In response to the Permafrost_cci Statement of Work [AD-1] and CCI project guidelines [AD-6], we have summarised in Table 2.1 how the validation activities for each of the products will meet those requirements.

Table 2.1: *Validation in Permafrost_cci compared to the overall project guidelines [AD-5; V1-V6] and as outlined in the Statement of Work (SoW) [AD-1; R-5, TR]*

Recommendations for validation	**Permafrost Temperature (GCOS ECV parameter)**	**Active Layer Thickness (GCOS ECV parameter)**	**Permafrost Fraction (no GCOS-ECV; target parameter in WMO OSCAR RRR)**
CCI [V-1] All CCI projects should use the definition of validation approved by the CEOS-WGCV	Validation for all products is compliant with the definition.		
CCI [V-2] All CCI project Product Validation Plans (PVP) shall adhere to the described three requirements regarding independence.	All three rules for independence of validation are adopted for each product.		
CCI [V-3] The CCI consortia shall use established, community accepted, traceable validation protocols where they exist. If such protocols do not exist, then CCI projects may adapt existing protocols if appropriate and in any event shall offer their final protocol for future community acceptance.	GTN-P TSP offers protocols on in situ measurements, data processing and data publication. Accepted statistical measures will be applied (mean, std dev., RMSE) in the validation process.	GTN-P CALM offers protocols on in situ measurements, data processing and data publication. Accepted Statistical measures will be applied (mean, std dev., RMSE) in the validation process.	Experimental validation with a range of reference data sets with all levels of quality of documentation. The binary validation uses GTN-P TSP ground temperature data.

Table 2.1 cont.

Recommendations for validation	**Permafrost Temperature (GCOS ECV parameter)**	**Active Layer Thickness (GCOS ECV parameter)**	**Permafrost Fraction (no GCOS-ECV; target parameter in WMO OSCAR RRR)**
CCI [V-5] The CCI programme should hold a dedicated session (or workshop) on common validation infrastructure during a CCI co-location meeting.	This will be feasible when the results of the Round Robin (RR) and validation are available		
CCI [V-6], SoW [R-5], [TR-14] The PVP shall fully describe the validation process for each CCI project. An independent international review board of experts should be invited to review the PVP of each project team. Each CCI project should involve experts from the CRG throughout their validation activities. A CCI product will be deemed to be validated once all steps of the validation process documented in the PVP have been completed and documented accordingly.	Permafrost_cci is involving the permafrost research community, the IPA and stakeholders and the CRG to give feedback on the validation procedure and published validation protocols on the Permafrost_cci product. We will use dedicated workshops and outreach on relevant Polar community and Polar political conferences. Independent assessment is specifically sought from the IPA Permafrost Mapping Action Group. This collaboration will be documented in all versions of the Climate Assessment Report (CAR). Dr. Isabelle Gärtner-Roer, University of Zuerich, CH, Vice president of IPA, leader of the IPA Permafrost Mapping Action Group [RD-6] and Science Officer of the World Glacier Monitoring Service (WGMS), is stating in [RD-7] that a very profound validation is being performed in Permafrost_cci by using the in-situ data from the GTN-P repository and from PERMOS. The validation of the Permafrost_cci ground temperature product is considered as the most important as it builds the base for the other products, such as active layer thickness and permafrost extent. Members of the IPA will be engaged in the validation and assessment activities via the Visiting Scientist program. Peer-reviewed joint papers are planned to document the results. Specifically, the outcome of the collection of reference data may result in Earth System Science Data ESSD publications which is highly attractive due to its high impact factor (e.g. Scopus 9.74, accessed November 2019).		
[TR-21] An independent validation shall be performed against metrics pre-defined defined by the contractor and endorsed by the user community.	The validation is fully independent as the validation team is independent of the algorithm development team and uses the global GTN-P data and monitoring networks such as the meteorological monitoring program ROSHYDROMET in Russia. Mountain permafrost is addressed by PERMOS in Switzerland. The characterisation of the errors and uncertainties is carried out using conventional evaluation of bias, absolute error and Root Mean Square Error RMSE difference.		Permafrost abundance provided by users will be compared with CCI permafrost fraction. The regional assessment will be carried out specifically in interaction with the IPA Permafrost Mapping Action Group.

Table 2.1 cont.

[TR-28] A full validation of all permafrost ECV products produced shall be performed.	A full validation for all permafrost ECV products is being performed		
Recommendations for validation	**Permafrost Temperature** **(GCOS ECV parameter)**	**Active Layer Thickness** **(GCOS ECV parameter)**	**Permafrost Fraction** **(no GCOS-ECV; target parameter in WMO OSCAR RRR)**
[TR-29] Validation shall quantify the uncertainty of the permafrost ECV products as well as the quality of the product uncertainty estimates themselves.	Validation is quantifying the uncertainty (RMSE in °C) and relative percentage error for the ensemble mean of annual average ground temperature per depth as well as the quality of the product uncertainty estimates themselves. Mountain permafrost regions will be addressed separately as here the uncertainties are expected to be significantly higher.	Validation is quantifying the uncertainty (RMSE in cm) and relative percentage error for ensemble mean of active layer thickness as well as the quality of the product uncertainty estimates themselves.	Validation is quantifying the uncertainty of the permafrost fraction product Mountain permafrost regions is addressed separately as here the uncertainties are expected to be significantly
[TR-30] The long-term stability of the Permafrost_cci time series of delivered epochs shall be assessed	The reference data sets on ground temperature and active layer depth in lowland and mountain permafrost areas are available over longer time periods and for different permafrost conditions, land cover, topography and climate. This will allow to extent the validation to characteristics important for climate research.		

3 RULES FOR UNBIASED VALIDATION AND VALIDATION CRITERIA

3.1 Unbiased validation

The project team shall ensure independency for the validation implying that the assessment of the Permafrost_cci product, as well as its uncertainties, is established with independent data sets and suitable statistical approaches [AD-1, 5, 6]. The validation needs to be carried out by teams not involved in the final algorithm selection [AD-5, 6].

The validation in Permafrost_cci will be fully independent as the validation team is independent of the algorithm development team and uses fully independent validation data sets from the global GCOS Global Terrestrial Network for Permafrost (GTN-P) program and further national measurement networks such as PERMOS in Switzerland and ROSHYDROMET in Russia [AD-1, RD-8]. WMO and GCOS delegated the ground-based monitoring of the ECV Permafrost to the Global Terrestrial Network for Permafrost (GTN-P) managed by the International Permafrost Association (IPA). GTN-P/IPA established the Thermal State of Monitoring (TSP) and the Circumpolar Active Layer Monitoring program (CALM, Brown et al., 2000), including standards for measurements and data collection (Clow 2014). The need to develop a suitable benchmark dataset has been stressed by users, as it does not yet exist [RD-4, 2, 3]. To validate the Permafrost_cci products, we are thus optimising the GTN-P ground data. This comprehensive dataset includes variable timeframes from hourly over annually to sporadic measurements and covers a wide range of different vegetation and permafrost types, however with incomplete data as well as metadata coverage and a considerable large percentage of misplaced coordinates which mostly relates to older boreholes and depend on region/data author [RD-8].

The validation and evaluation efforts innovatively consider high-mountain permafrost regions, using in-situ observations of ground temperatures, and velocities of permafrost creep, provided by national data-services such as GTN-P PERMOS in Switzerland. In addition, the EO derived inventory on rock glacier abundance, extent, and creep, which was developed by the ESA Data User Element (DUE) GlobPermafrost program since 2016 and which is continued in Permafrost_cci, will be used. The PERMOS monitoring data and ESA GlobPermafrost rock glacier inventory will support the validation in mountain areas, where the Permafrost_cci products contain the highest uncertainties [RD-8].

Independent assessment is also sought from the IPA Permafrost mapping action group. This collaboration will be documented in all versions of the Climate Assessment Report (CAR). Members of the IPA will be engaged in the validation and assessment activities via the Visiting Scientist program. Dr. Isabelle Gärtner-Roer, University of Zuerich, CH, Vice president of IPA, leader of the IPA Permafrost mapping action group [RD-6] and Science Officer of the World Glacier Monitoring Service (WGMS), is stating in CAR v2 [RD-7] that a very profound validation is beeing performed in Permafrost_cci by using the in-situ data from the GTN-P repository and from PERMOS. Still, in-situ data are clustered in regions with active permafrost monitoring programs/projects, and that therefore some regions are underrepresented and validations are less detailed. The IPA states that they will continue to support the GTN-P repository and foster standardized compilation of in-situ data on permafrost temperatures and active layer thicknesses, especially in underrepresented regions. For the validation in Permafrost_cci, IPA further provides the recommendation that the validation of the

Permafrost_cci ground temperature product is the most important as it builds the base for the other products, such as active layer thickness and permafrost extent.

3.2 Validation criteria

The required parameters by GCOS for the Permafrost ECV are [AD-1,4]

a) permafrost temperature (K), and

b) active layer thickness (m).

The main focus of Permafrost_cci lies on the ECV permafrost temperature. Since EO-based algorithms operate at the spatial scale of pixels, the spatial resolution of the output must be set in context with the spatial variability of permafrost temperatures and active layer thickness. In many permafrost regions, these can display a strong variability at spatial scales of meters, which is much finer than the footprint of EO sensors. For this reason, it makes sense to add an additional variable,

c) permafrost extent (fraction),

as mapped permafrost variable, which is the fraction within an area (pixel) at which the definition for the existence of permafrost (ground temperature < 0 °C for two consecutive years) is fulfilled.

The assessment of the products is performed with the optimised and harmonised validation data sets that are independent from the production of the Permafrost_cci product using a range of statistical approaches [RD-9]. The characterisation of errors and uncertainties is carried out using conventional evaluation measures of mean bias, absolute error difference and Root Mean Square Error (RMSE).

The validation will be performed point-wise, grid-wise and functional [AD-1]. The validation in all versions is described in the Product Validation and Intercomparison report [RD-3].

- Point-wise site-specific match-up analyses per ground temperature profile per standardised depth and time stamp for annual time scales provide average bias, mean absolute error and RMSE. Metadata on ice content, vegetation and organic layer, and lithostratigraphy provide further information on temperature-influencing factors
- Functional validation on circum-Arctic scale will provide the relative MAGT error per pixel by classifying the site-specific match ups over different classes (land cover, lithology, topography). Coming from these classes to class specific relative errors per pixel, this method provides the relative error per pixel in the CCI product on a circum-Arctic scale.
- Point-wise and grid-wise regional comparisons of Permafrost abundance can be carried out specifically in interaction with the IPA Permafrost Mapping Action Group by users. The regional assessment will be carried out in regions with existing ground data availability and regional expertise provided by members of the user community, specifically through IPA.

4 VALIDATION ACTIVITIES

4.1 Validation Data Sets

Special emphasis in Permafrost_cci is placed on validation using data from international and national permafrost monitoring networks and in cooperation with the permafrost community [AD-1]. These available ground data sets and their characteristics and data availability (data access via data portals and program websites) are described in detail in the DARD [RD-8]. The data sets used for the first validation are described in the PVIR version 1 [RD-3]. In the following sub-chapters we provide more details on used and also on planned data sets.

4.1.1 Time series on ECV permafrost temperature

Data sets on permafrost temperature for the validation of terrestrial permafrost are managed and made publicly available at no costs via large-scale international and regional programs where several of the team members are in close cooperation with. GTN-P together with the Arctic Portal provides a dynamic GCOS GTN-P database for upload and download of data containing CALM and TSP data in the Arctic, Antarctic, Central Asia and mountain regions [RD-8]. The national monitoring networks also sustain national databases and portals for downloading data, such as GTN-P PERMOS in Switzerland.

Table 4.1: *Ground data for sub-ground thermal properties available for validation in Permafrost_cci.*

Region	Data on ground temperature, active layer	Contributor
Circumpolar Arctic and Antarctica	temperature data (boreholesoil,), active layer depths, several decades for some sites	CALM, dynamic GCOS GTN-P database
North American Permafrost Regions	temperature data (borehole, soil,), active layer depths, several decades for some sites	CALM, GIPL/UAF and NPS published in NSF Arctic Data Centre (Wang et al. 2018), dynamic GCOS GTN-P database, NORDICANA D
Siberian Permafrost Regions	temperature data (borehole, soil) active layer depths	dynamic GCOS GTN-P database, CALM ROSHYDROMET
European high-latitude Permafrost Regions	temperature data (borehole, soil,)	ROSHYDROMET, dynamic GCOS GTN-P database

For Canadian datasets, Nordicana D is the data repository of the Canadian Centre d'études Nordiques (CEN). Nordicana D curates long-term time series of permafrost borehole temperatures, and also hosts and publishes datasets of shallow ground and air temperature in high temporal resolution that are not in full extent part of the GTN-P database. ROSHYDROMET is the national Russian meteorological monitoring network providing long-term ground temperature records close to meteorological stations. Permafrost_cci is compiling these ground temperature records from GTN-P and ROSHYDROMET (see Table 4.1 to 4.1, [RD-8]), and further national permafrost data sets.

Table 4.2: *Overview on GCOS GTN-P ground temperature data collections.*

GCOS GTN-P data collections	**No. of boreholes**
GCOS GTN-P boreholes	1360
GCOS GTN-P boreholes containing data	485
GCOS GTN-P boreholes containing measured in situ temperature data	372
GCOS GTN-P boreholes containing measured in situ data without empty datasets	369
GCOS GTN-P boreholes containing measured in situ data without empty datasets in mountain permafrost regions	**35**
GCOS GTN-P boreholes not in mountain permafrost regions containing measured in situ data without empty datasets	334
GCOS GTN-P boreholes not located in the permafrost zone (continuous, discontinuous, patchy, isolated)	63
GCOS GTN-P boreholes containing measured in situ data without empty datasets within the permafrost zone (continuous, discontinuous, patchy, isolated, permafrost probability >0.1)	**270**

Table 4.3: *Ground temperature monitoring records from ROSHYDROMET by permafrost zone.*

Definition from	**Permafrost**	**Boreholes**
Permafrost Extent NSIDC	**C (=Continuous)**	40
	D (=Discontinuous)	25
	I (=Isolated patches)	54
	S (=Sporadic)	32
	NA	65
Permafrost Probability	**≤1**	46
	≤0.8	13
	≤0.6	18
	≤0.4	52
	≤0.2	73
	0	10
	NA	4

Table 4.4: *Overview on ground temperature monitoring sites and characteristics in the Swiss National GTN-P PERMOS monitoring program*

borehole	GTN-P	start	lat [DD]	long [DD]	elevation [m]	depth [m]	morphol-ogy	surface type	permafrost thickness
ATT_0108	CH 01	2008	46.09677	7.273075	2661	26	talus slope	coarse blocks	>24 m
ATT_0208	CH 02	2008	46.09675	7.273682	2689	21	talus slope	coarse blocks	>20 m
ATT_0308	CH 03	2008	46.0966	7.274924	2741	15	talus slope	coarse blocks	no permafrost
COR_0200	CH 14	2000	46.42853	9.82202	2672	63	rock glacier	coarse blocks	>62 m
COR_0287	CH 13	1987	46.42879	9.821836	2670	62	rock glacier	coarse blocks	>60 m
DRE_0104	CH 04	2004	46.27333	6.889508	1580	15	talus slope	coarse blocks	no permafrost
FLU_0102	CH 05	2002	46.74887	9.943555	2394	23	talus slope	debris	ca. 5 m
GEM_0106	CH 06	2006	46.60125	8.610422	2905	40	crest	bedrock	no permafrost
GEN_0102	CH 07	2002	46.08371	7.302472	2888	20	moraine	debris	>20 m
JFJ_0195	CH 31	1995	46.54611	7.973192	3590	21	crest	bedrock	
LAP_0198	CH 08	1998	46.10612	7.284349	2500	20	talus slope	coarse blocks	>20 m
LAP_1108	CH 32	2008	46.10623	7.284724	2500	40	talus slope	coarse blocks	ca. 15 m
LAP_1208	CH 33	2008	46.10564	7.283808	2535	35	talus slope	coarse blocks	ca. 20 m
MAT_0205	CH 09	2005	45.98232	7.676049	3295	53	crest	bedrock	>53 m
MBP_0196	CH 10	1996	46.4964	9.931076	2946	18	talus slope	debris	>18 m
MBP_0296	CH 11	1996	46.49657	9.93141	2942	18	talus slope	debris	>18 m
MUR_0199	CH 12	1999	46.50757	9.927823	2536	70	rock glacier	coarse blocks	no permafrost
MUR_0299	CH 34	1999	46.50723	9.927338	2539	64	rock glacier	coarse blocks	ca. 18 m
MUR_0499	CH 35	1999	46.50723	9.927703	2549	71	rock glacier	coarse blocks	>15 m
RIT_0102	CH 15	2002	46.17469	7.849835	2690	30	rock glacier	coarse blocks	>13 m
SBE_0190	CH 16	1990	46.49738	9.926302	2754	67	rock glacier	coarse blocks	>16 m
SBE_0290	CH 17	1990	46.4988	9.925215	2732	60	rock glacier	coarse blocks	>25 m
SCH_5000	CH 19	2000	46.55828	7.834426	2910	101	crest	debris	>100 m
SCH_5198	CH 18	1998	46.55828	7.834621	2910	14	crest	debris	>13 m
SCH_5200	CH 20	2000	46.55828	7.834426	2910	100	crest	debris	>100 m
STO_6000	CH 21	2000	45.98679	7.824201	3410	100	crest	debris	>100 m
STO_6100	CH 22	2000	45.98655	7.824057	3410	31	crest	debris	>17 m
TSA_0104	CH 23	2004	46.10905	7.548442	3040	20	crest	bedrock	>20 m

While all ROSHYDROMET-datasets contain temperature measurements down to a depth of maximum 3.2 m, the GTN-P dataset consists of a variety of deep boreholes and shallow profiles with 2-dimensional temperature measurements per specific depths in the sub-ground. About 35 % of the GTN-P data come from borehole measurements that are more shallow than 5 m depth (Figure 4.1). A majority of these shallow sites originates from Russia (containing some sites, which are in the ROSHYDROMET dataset as well) and from the United States (Table 4.5).

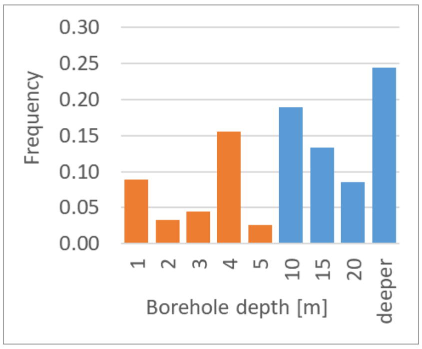

Figure 4.1: *Borehole-depth and their frequency within the GTN-P dataset, separated in shallow boreholes ≤5 m depth (orange bars) and deeper boreholes (blue bars).*

Table 4.5: *National origin of shallow (≤5 m) boreholes in GTN-P*

Country	Borehole<=5m
United States	30
Russia	56 (~30 of these in RHM)
Canada	5
Greenland	1
Antarctica	2

Ground temperature accuracy – estimated impact on ground temperature 0.1K. Ground Temperature in soil profiles or boreholes is measured either by lowering a calibrated thermistor into a borehole, or recorded using permanently installed multi-sensor cables Measurements are recorded either manually with a portable temperature system or by automated continuous data logging. The reported measurement accuracy of the temperature observations, including manual and automated logging systems, varied from ±0.01 to ±0.25 °C with a mean of ±0.08 °C. Previous tests have shown the comparability of different measurement techniques to have an overall accuracy of ±0.1 °C. Thermistors are the most commonly used sensors for borehole measurements. Their accuracy depends on (1) the materials and process used to construct the thermistor, (2) the circuitry used to measure the thermistor resistance, (3) the calibration and equation used to convert measured resistance to temperature, and (4) the aging and resulting drift of the sensor over time. Thermistors are typically calibrated to correct for variations due to (1) and (2).

About 10 to 20 % of the boreholes are visited once per year and measured using single thermistors and a data logger. In this case, the system is routinely validated in an ice-bath allowing correction for any calibration drift. The accuracy of an ice-bath is ~± 0.01 °C. Using the offset determined during this validation to correct the data greatly increases the measurement accuracy near 0 °C, an important reference point for permafrost. The remaining systems are permanently installed and typically ice-bath calibrated at 0 °C before deployment. The calibration drift is difficult to quantify as thermistor chains are not frequently removed for re-calibration or validation. In many cases, removal of thermistor chains becomes impossible some time after deployment, e.g. because of borehole shearing.

The drift rate among bead thermistors from different manufacturers was <0.01 °C per year during a 2-year experiment at 0, 30, and 60 °C. The calibration drift of glass bead thermistors was found to be 0.01 mK per year, at an ambient temperature of 20 °C. A single drifting thermistor in a chain is detectable through its anomalous temporal trend. Such data are excluded from final data sets.
The above discussion of accuracy relates to the absolute temperature values measured, but the detection of temperature change is more accurate because errors in calibration offset have no impact, sensor nonlinearities are generally small and not of concern. We therefore consider <0.1 °C a conservative average estimate of the accuracy of temperature change on an individual sensor basis.

Synthesized reference time series data

GTN-P and ROSHYDROMET time series provide climate research data sets, however no easy-to use time series depth data that are data fit for validation and RR exercises. The GCOS GTN-P data collection is accessible via http://gtnpdatabase.org/boreholes and comprises 1360 sites with metadata entries (Table 4.2) (Biskaborn et al., 2015). Within this active GCOS GTN-P data collection, only 485 sites contain -in addition to the metadata entries - also datasets on measurements to date, of which 113 have satellite-derived DUE Permafrost satellite products on LST and Surface Moisture and no in-situ ground temperature data. Several PIs, (e.g., China with a share of around 200 sites) have up to date not submitted GTN-P borehole data to the data portal. The GTN-P office is currently supporting borehole data intake into the GTN-P data portal. There are some issues with technical data output from the data portal with empty data sets leaving 369 GTN-P sites with measured in situ data. 270 sites with ground temperature datasets are placed in lowland regions within the permafrost zone.
After screening the GTN-P data collection, we found that ~40-50 % are usable for validation in Permafrost_cci. The ground temperature time series frequently contain large data gaps, data input errors or further artefacts. Also ROSHYDROMET data are known to have artefacts like wrong temperature data that were manually put in, and problems concerning geo-location, having only two decimal digits.

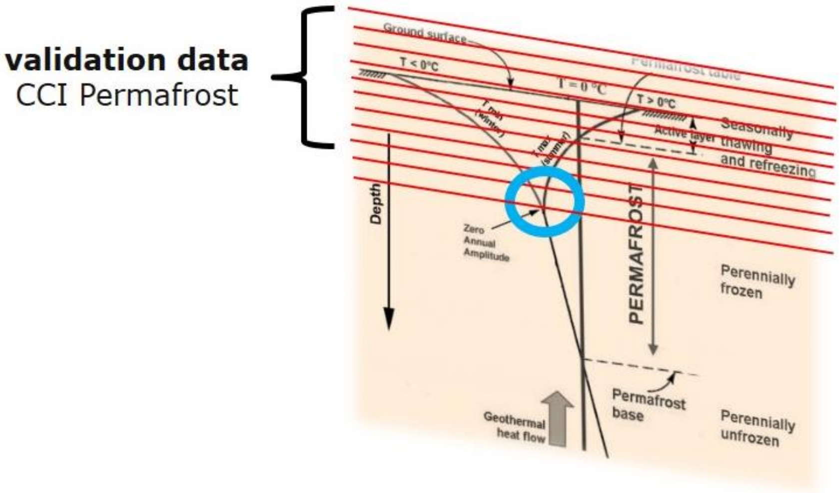

Figure 4.2: *Concept of Zero Annual Amplitude (ZAA) depth and Permafrost_cci depth stratification.*

Although the various datasets contain measured temperature at many different depths, some depths are available at a high frequency (Figs 4.3, 4.4). The ‘Alaska shallow’ ground temperature data collection (blue bars) contains already interpolated temperature data at four depths (0.25 m, 0.5 m, 0.75 m, 1 m) for many Alaskan ground temperature soil profiles (Wang, 2018; Wang et al., 2018).

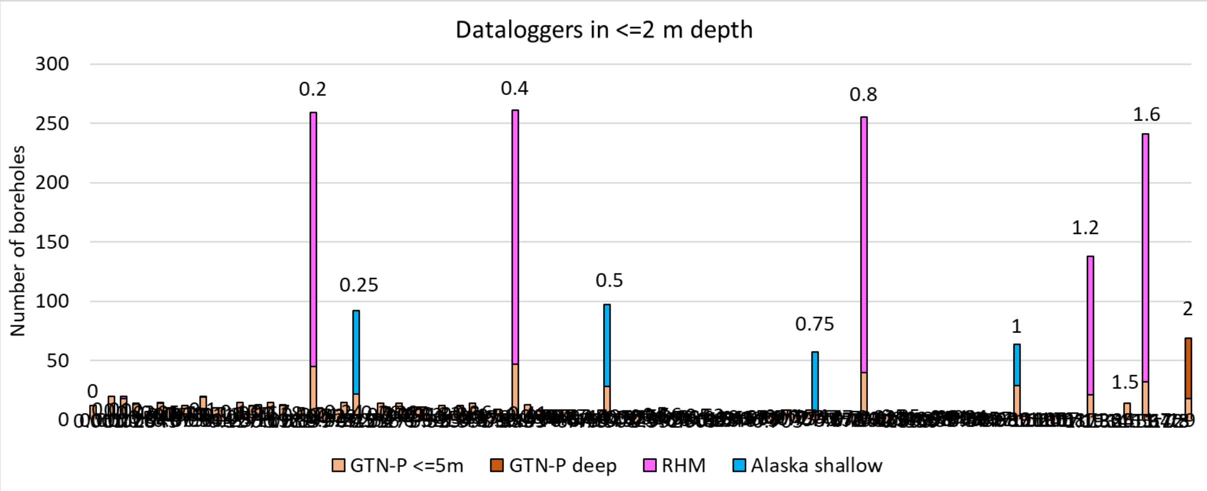

Figure 4.3: *Number of boreholes/sites (x-axis) with sensors at a given depth (in meter, y-axis). Only sensors in <= 2 m depth are given for GTN-P, ROSHYDROMET and the synthesized Alaska ground temperature data collection (Wang, 2018; Wang et al., 2018). Measurements from temperature sensors in <2 m depth are omitted for deep boreholes (5 m and deeper).*

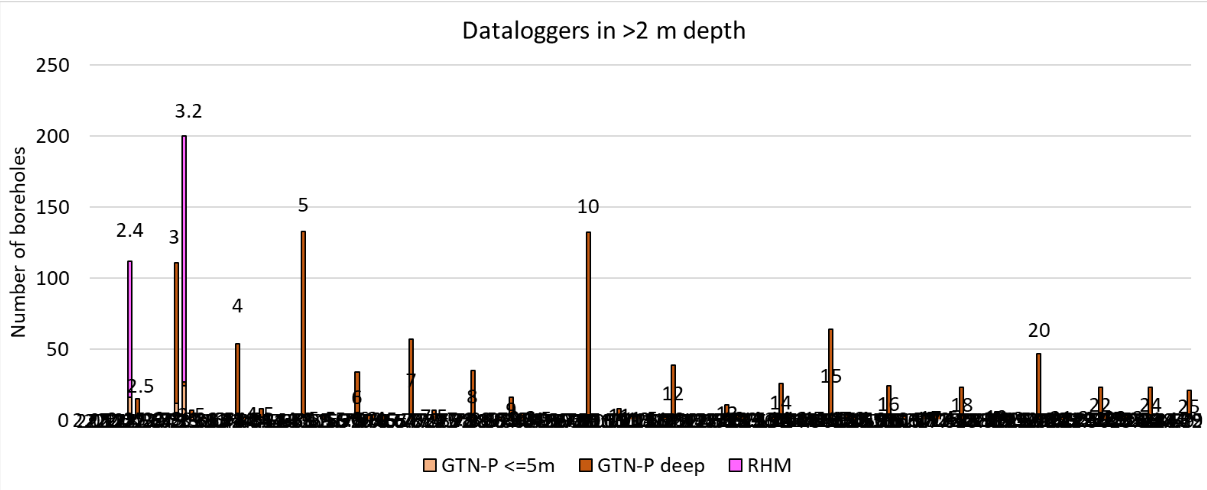

Figure 4.4: *Number of boreholes/sites (x-axis) with sensors at a given depth (in meter, y-axis). Only sensors in >2 m depth are given for GTN-P and ROSHYDROMET.*

Version 1 synthesised data set - discrete depths

In a first step, for Version 1, we created a dataset with discrete depths (Fig 4.5, [RD-3]). As we frequently detected moving sensors, which change position over time, we calculate discrete depths with some range around the focus depth. We will thus provide temperature data at (0, 0.2, 0.25, 0.4, 0.5, 0.75. 0.8, 1,0 1.2) m ± 3 cm, at (1.6, 2.0, 3.0, 3.2, 4) m ± 5 cm, at 5 m ± 10 cm and at (10, 20) m ± 20 cm. We distinguished between two borehole types. Group I contains ground temperature measurements down to 5 m depth (exclusive). In Group I, all discrete values are calculated. Group II contains ground temperature measurements of 5 m depth and deeper. Within Group II, we discard all data <2 m depth from deep boreholes as in boreholes with large diameters, there is either frequently artificial infilling material or air in the borehole. The air temperature can then influence measurements in shallow depths. At many boreholes, only annual measurements are available. Measurement depths above ZAA need to be discarded in case of annual measurement frequency that would refer for most boreholes to depths shallower than 10 m.

Ground temperature records are then processed to monthly and yearly mean/min/max, containing metadata information, which allows assessing the quality of each temperature value product (Table 4.6). These metadata comprise for monthly values the ratio of missing values per month and for yearly values the ratio of missing data per month/year (missing days per year/365) and the amount of completely missing months. Monthly mean/min/max are not calculated if >20 % of the monthly values are missing. Yearly mean/min/max are not calculated if >20 % of yearly values are missing OR if more than one complete month is missing.

Permafrost_cci match-up data set in phase 1, Version 1 used for the first validation[RD-3]) contains standardised ground temperature per depth GTD data with annual resolution from 2003 and 2017 with a circum-Arctic geographic coverage.

This annual GTD data set (mean, min, max) from 2003 to 2017 is compiled from all the discrete depths and time stamps and national and international programs available: depths are at 0, 0.2, 0.25, 0.4, 0.5, 0.6, 0.75. 0.8, 1.0 1.2, 1.6, 2.0, 2.4, 2.5, 3.0, 3.2, 4.0, 5.0, 10.0, 20.0 m. Permafrost_cci match-up data set in phase 1, Version 1 has been compiled from

- GTN-P (https://gtnp.arcticportal.org/),
- Roshydromet RHM (http://meteo.ru/data/164-soil-temperature),
- Nordicana-D (http://www.cen.ulaval.ca/nordicanad/dpage.aspx?doi=45291SL34F28A9491014AFD; Allard et al., 2016, CEN 2013),
- PANGAEA (https://doi.pangaea.de/10.1594/ PANGAEA.905233; Boike et. al. 2019),
- Arctic Data Center (https://arcticdata.io/catalog/ #view/doi:10.18739/A2KG55; Wang et al. 2018)

The Permafrost_cci GT data collection v1 contains data from 336 in situ measurement locations (GTN-P n = 82, RHM n = 169, Nordicana-D n = 12, PANGAEA n = 3, Arctic Data Center n = 70), with overall n = 10865 data points (Ø ~ 517 values per depth).

The results of the first validation activity [RD-3] show that simulated Permafrost_cci MAGT has large deviations in the ‘warm’ GT range in regions where in situ MAGT reaches temperatures >1 °C. 147 in situ measurement locations with 2841 match-up pairs in time and depth fall into the match-up group of mean annual GT <1 °C.

The match up results of the first depth- and time specific validation activities [RD-3] further revealed that match up results are more robust for all depths below 40 cm [RD-3]. 145 datasets with 2215 match-up pairs in time and depth fall into the match-up group of MAGT <1 °C and depths ≥40 cm.

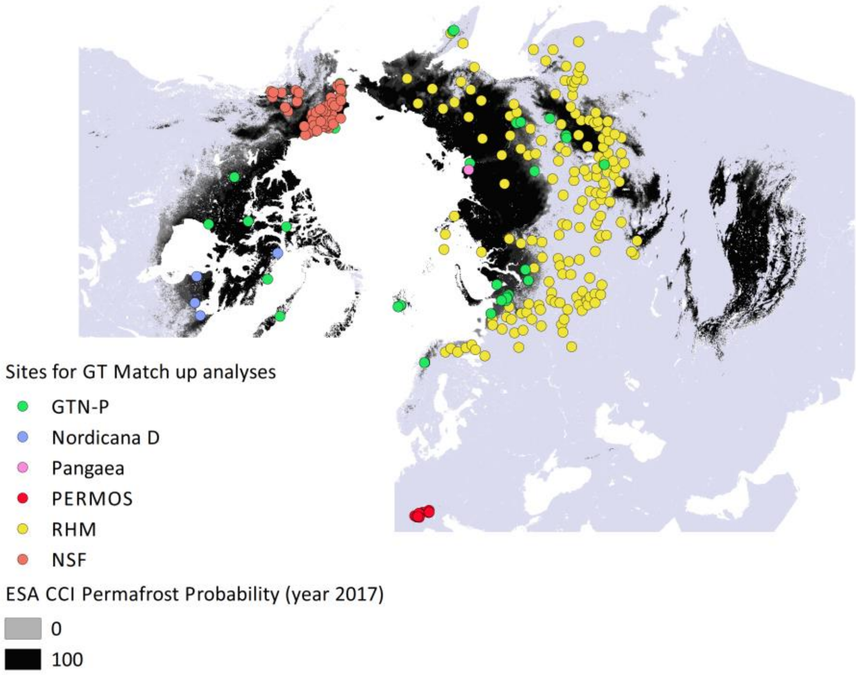

Figure 4.5. *Northern hemisphere Permafrost_cci permafrost probability and in situ ground temperature stations (grouped by data source), including PERMOS boreholes in the Swiss Alps.*

Version 2 synthesised data set - interpolated depths

In a next step, for Version 2, we will interpolate the at-sensor temperature measurements (Group III, Figure 4.6). As sensors in different boreholes and soil profiles are not always placed in the same depth, interpolation allows to create datasets with consistent depths, that are easier to process in further analyses. We will follow the approach by Wang et al. (2018) and do a stepwise interpolation from sensor to sensor. Interpolation will be conducted (i) on datasets with sufficient sensors and (ii) down to the deepest sensor.

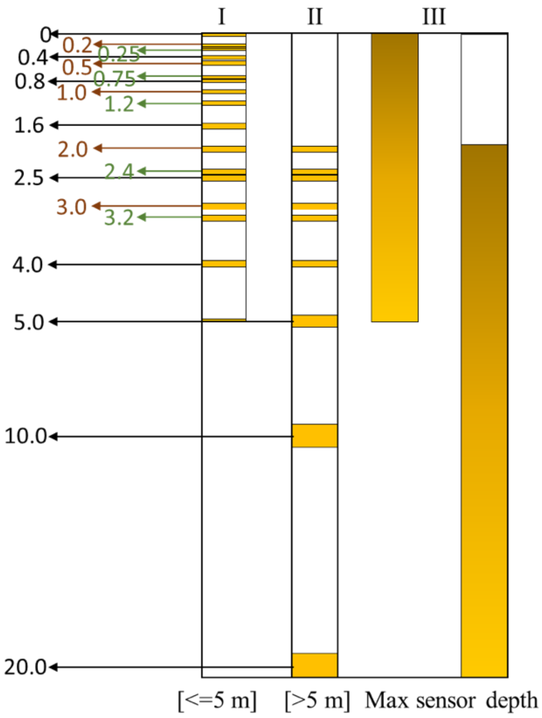

Figure 4.6: *Discrete depths in version 1 will be provided small-scaled in shallower depths and larger scaled in deeper depths. In deep boreholes (Group II, >5 m depth), <1.5-2 m measurements will be discarded due to their inaccuracy in large borehole set-ups. In step 2 (Group III), temperature data will be interpolated for shallow and deep temperature profiles down to the maximum sensor depth. Also here, the upper 1.5-2 m temperature measurements of deep boreholes will be discarded.*

Table 4.6: *Example of how the compiled dataset will provide yearly values. Mxx = Ratio of missing values per month/year at depth xx m. mMxx = Number of completely missing months per year at depth xx m.*

1	Site	Year	Type	M0	M0.2	M0.25	M0.4	M0.5	M0.75	M0.8	M1	mM0	mM0.2	mM0.2	mM0.4	mM0.5	mM0.7	mM0.8	mM1	0	0.2	0.25	0.4	0.5	0.75	0.8	1
2	FB_dry_l	2006	Mean	1	1	1	1	1	1	1	1	12	12	12	12	12	12	12	12	NA	NA	NA	NA	NA	NA	NA	NA
3	FB_dry_l	2006	Max	1	1	1	1	1	1	1	1	12	12	12	12	12	12	12	12	NA	NA	NA	NA	NA	NA	NA	NA
4	FB_dry_l	2006	Min	1	1	1	1	1	1	1	1	12	12	12	12	12	12	12	12	NA	NA	NA	NA	NA	NA	NA	NA
5	FB_wet_	2006	Mean	0.414	0.414	0.414	0.416	0.414	NA	NA	NA	5	5	5	5	0	NA	NA	NA	1.33	1.64	1.56	1.35	1.12	NA	NA	NA
6	FB_wet_	2006	Max	0.414	0.414	0.414	0.416	0.414	NA	NA	NA	5	5	5	5	0	NA	NA	NA	18.9	12.7	12	10.4	8.07	NA	NA	NA
7	FB_wet_	2006	Min	0.414	0.414	0.414	0.416	0.414	NA	NA	NA	5	5	5	5	0	NA	NA	NA	-19.1	-12	-11.5	-10.2	-8.95	NA	NA	NA
8	FB_dry_l	2007	Mean	0.581	0.581	0.581	0.581	0.581	1	0.586	0.699	7	7	7	7	7	12	7	8	-3.58	-2.65	-2.53	-2.38	-2.44	NA	-2.4	-2.59
9	FB_dry_l	2007	Max	0.581	0.581	0.581	0.581	0.581	1	0.586	0.699	7	7	7	7	7	12	7	8	13.6	10.4	9.31	8.01	4.87	NA	1.73	0.63
10	FB_dry_l	2007	Min	0.581	0.581	0.581	0.581	0.581	1	0.586	0.699	7	7	7	7	7	12	7	8	-21.9	-17.5	-16.9	-16	-14.6	NA	-11.9	-8.83
11	FB_wet_	2007	Mean	0	0	0	0	0	NA	NA	NA	0	0	0	0	0	NA	NA	NA	-5.99	-5.41	-5.62	-5.48	-5.63	NA	NA	NA
12	FB_wet_	2007	Max	0	0	0	0	0	NA	NA	NA	0	0	0	0	0	NA	NA	NA	17.8	15.2	11.7	10.6	7.49	NA	NA	NA
13	FB_wet	2007	Min	0	0	0	0	0	NA	NA	NA	0	0	0	0	0	NA	NA	NA	-30.2	-23.3	-22.7	-21.3	-20.3	NA	NA	NA

4.1.2 Time series on Active Layer Thickness

A comprehensive collection of active layer thickness time series is available from the Circumpolar Active Layer Monitoring Network (CALM, Table 4.7, Brown et al., 2000, Fagan and Nelson, 2017). The data are available for download on https://www2.gwu.edu/~calm/ - some as the active layer thickness in a specific year expressed as maximum thaw depth measured in late summer at the CALM long term monitoring grids. The single active layer thickness grid point measurements are either averaged per grid, or provided as single grid measurement values of thaw depths. Only few published CALM datasets have no observation data on thaw depth but contain metadata only. Round 60 % of the CALM data are also

available in the Arctic Portal GTN-P data collection for download (http://gtnpdatabase.org/activelayers) (Table 4.7).
About half of the CALM datasets consist of up to 10 measurements (though not necessarily in ten consecutive years), ~15 % of the data provide more than 20 years and up to a maximum of 29 years of observations on active layer thickness.
Permafrost_cci match up data set in phase 1, Version 1: standardised active layer thickness ALT data with annual resolution from 2003 and 2017 with a circum-Arctic geographic coverage. The collection contains data from 207 sites (China + Mongolia: 67, Greenland + Svalbard + Scandes: 11, Canada: 6, Russia: 57, USA: 207), with overall 1835 match-up pairs in time [RD-3]. Also, several users will provide single-point observations on thaw depth measurements that are standard measurements accompanying activities for soil profiles and permafrost cores. These thaw depth observations contain the point coordinates, the day of the measurement and the thaw depth.

Table 4.7: *Regional overview of active layer measurements from CALM website and GTN-P Arctic Portal respectively.*

CALM Website				**GTN-P Arctic Portal**		
Country	**No. Of sites**	**with Data**		**Country**	**No. Of sites**	**with Data**
Antarctica	27	15		Antarctica	12	1
South America	14	5				
Canada	31	31		Canada	31	9
China	11	7		China	11	0
Denmark (Greenland)	3	3		Greenland	3	2
Kazakstan	3	3		Kazakhstan	3	0
Mongolia	47	47		Mongolia	46	0
Norway/Svalbard	3	3		Norway	1	0
Poland/Svalbard	4	4				
Russia	68	68		Russia	68	65
				Svalbard	7	1
Sweden/Svalbard	2	2		Sweden	1	0
Switzerland	2	2		Switzerland	2	0
United States	67	66		United States	67	26
SUM	**283**	**256**			**252**	**104**

6 of these not on CALM-Website

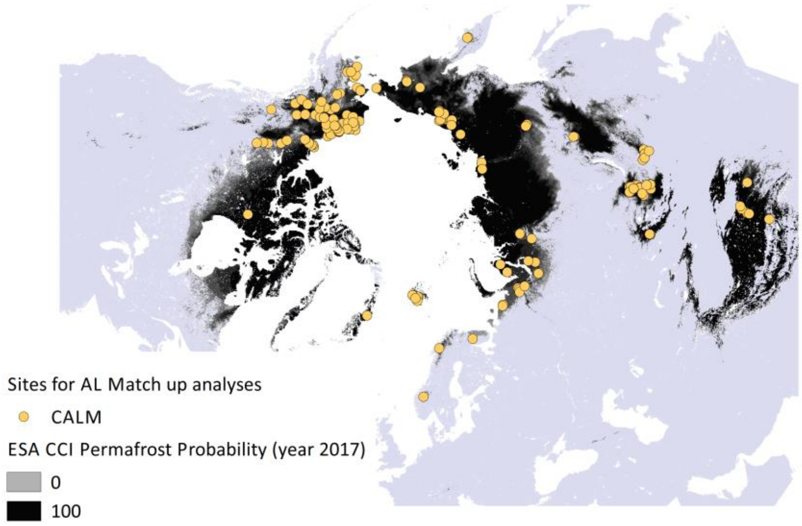

Figure 4.7. *Northern hemisphere Permafrost_cci permafrost probability and in situ sites of active layer depth ALT (GTN-P CALM programme)*

Thaw depth accuracy – estimated impact on thaw depth 0.02 m
[Thaw depth is an essentially instantaneous value that is always less than or equal to the thickness of the fully developed active layer. Probing of the active layer is performed mechanically with a graduated rod. The typical probe is a 1 m long stainless-steel rod. The probe rod is inserted into the ground to the point of resistance. A distinctive sound and feel is apparent when ice-rich frozen ground is encountered. At sites where thaw depth is very large (e.g., 1-3 m), it is very difficult, however, to extract a probe in deeply thawed soils, or stony soils. Optimally executers should have experience with this measurement and body strength].
Active layer thickness accuracy– estimated impact on active layer thickness 0.05 m
Nelson and Hinkel (2003, in "Methods for measuring active-layer thickness. In: A Handbook on Periglacial Field Methods") highlight that the term of thaw depth is distinct from the term of active layer thickness. The permafrost ECV active layer thickness is used in reference to the maximum development of the thawed layer, reached at the end of the warm season. This is distinct from the term active layer depth referring to the thickness of the thawed layer at any time during its development in summer.
Active Layer thickness is usually measured on grids of 10, 100 or 1000 m with evenly spaced nodes at 1, 10 or 100 m. Fagan and Nelson (2017) showed, that a systematic stratified unaligned design has advantages over a systematic design, but that the inaccuracy of a systematic design is only small in comparison stratified unaligned design. Active-layer thickness can vary substantially on an inter-annual basis. In general, it is greater in years with warmer summers and thinner in those with cooler temperatures (Brown et al., 2000).
For an estimation of the ECV active layer thickness it is relevant to measure active layer depths in the grid at the end of the thawing season (https://www2.gwu.edu/~calm/data/north.html). For some measurements in the CALM data collection, metadata information is provided if a value was measured earlier during a year. These measurements are discarded from the validation data set on Active Layer thickness.

4.2 Metadata relevant for permafrost modelling

Vincent et al. (2017) formulated the ‘3-layer permafrost Earth system approach’ integrating geosystem and resilience frameworks. Their definition goes beyond the classical 2-layer permafrost system with the permafrost overlain by the seasonally dynamic active layer. The geocryology or phase composition of soil/rock, ice, air, unfrozen water, organic content, and cryotexture, and cryostructure define all thermal sub-ground properties of the two layers in the geosystem. Vincent et al. (2017) formulated how in natural environments the 3rd layer, the buffer layer, consists of the above-ground vegetation, from polar desert soil crusts to tundra grasses, forbs, and lichens to shrubs and trees farther to the south (Figure 4.8). In engineered environments, the buffer layer includes the infrastructure. In both cases, this surface buffer layer strongly affects the transfer of heat between the atmosphere and the active layer. This effect is compounded by the accumulation of snow in the buffer layer, which is determined not only by the regional precipitation regime, but also by the snow-trapping efficiency of above-ground vegetation or engineered structures. Permafrost lands vary greatly in horizontal space, and the properties of each of the three layers and their interfaces can change over short length scales. In discontinuous permafrost, units are interspersed with non-permafrost units including lakes, bogs, rivers, and unfrozen ground.

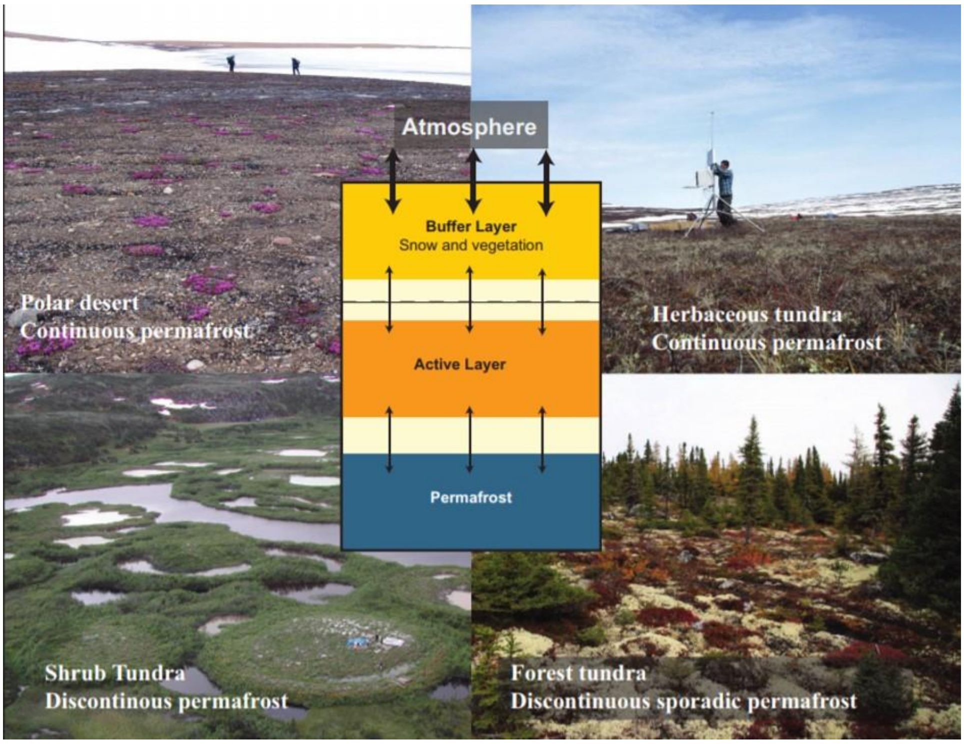

***Figure 4.8** (Fig. 4 in Vincent et al., 2017): The three-layer model applied to natural landscapes. As illustrated here, the buffer layer varies greatly in its geometry (thickness), albedo, and other physical properties, both within and between landscapes, and as a function of vegetation type and season. The arrows indicate exchanges of heat, water, and gases and the white bands indicate interface zones. Upper left: Ward Hunt Island, Nunavut; upper right: Daring Lake, Northwest Territories; lower left: BGR valley, Nunavik; lower right: Umiujaq region, Nunavik.*

Therefore, in Permafrost_cci, we will assemble

I) a data collection on 'ground temperature' and 'thaw depth' in time (with annual time stamp) and in space (geographic coordinates with decimal degree with minimum of 4 digits) and ground depth (single depth or depth interval)

II) a data collection on metadata (Figure 4.9)

a) on the 2-layer system permafrost and active layer (stratigraphy, organic layer (abundance, thickness), ground ice content, dominating lithology and texture)

b) on the buffer layer (vegetation composition, height of vegetation, infrastructure, surface habitus (boulders, gravel, …)).

We will use all published information available, the detailed CALM metadata if available, and will also retrieve this information from the PIs directly. To avoid having to discard older measurements, where meta-information is not available, we will provide best guesses based on surrounding measurements, field photos, remote sensing data, and expert knowledge. A quality index for these values will then help users working with the dataset, to assess the value of the data (Tables 4.8-4.11).

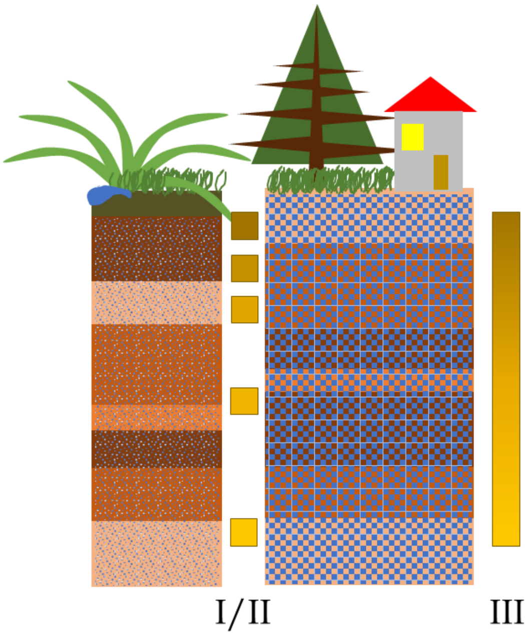

Figure 4.9: *Permafrost_cci metadata collection planned on stratigraphy, organic layer, ground ice content, dominating lithology and texture and on the buffer layer (vegetation composition, infrastructure,…) will be provided at the best quality available and including a quality index (cf. Tables 4.8 – 4.11).*

Table 4.8: *A metadata table on the information of thickness of the insulating organic layer could look like this, although final decisions on the quality indexes for the different sources are still pending*

Organic-Layer (O-L)	O-L detail	O-L origin	O-L quality index
thickness: 12 cm; type: moss layer	quantitative, qualitative	PI, CALM metadata, publication	1 = best quality
’thick moss layer’ >5 cm thickness or ’thin moss layer’ <5 cm thickness or no moss layer	qualitative	Indirect retrieval of thickness of the insulating organic layer published Site Pictures, CALM landscape description	2
’thick moss layer’ >5 cm thickness or ’thin moss layer’ < 5cm thickness or no moss layer	qualitative	Information on comparable locations close by (e.g. same landscape type)	3
’thick moss layer’ >5 cm thickness or ’thin moss layer’ <5 cm thickness or no moss layer	qualitative	high spatial resolution satellite data, other sources	4
’thick moss layer’ >5 cm thickness or ’thin moss layer’ <5 cm thickness or no moss layer	qualitative	best guess with few information available	5 = worst quality

Table 4.9: *A metadata table on the information of vegetation cover could look like this, although final decisions on the quality indexes for the different sources are still pending*

Vegetation-Cover (V-C)	V-C detail	V-C origin	V-C quality index
Forest tundra (tree height 5 m, 25 % coverage) with dwarf-shrubs (15 %), moss layer layer (75 %)	quantitative, qualitative	PI, CALM metadata, publication	1 = best quality
‘Forest tundra’ or ‘Polygonal tundra’ or ‘Tundra’ or […]	qualitative	Indirect retrieval of vegetation cover (in classes?), from published Site Pictures, CALM landscape description	2
‘Forest tundra’ or ‘Polygonal tundra’ or ‘Tundra’ or […]	qualitative	Information on comparable locations close by (e.g. same landscape type)	3
‘Forest tundra’ or ‘Polygonal tundra’ or ‘Tundra’ or […]	qualitative	high spatial resolution satellite data, other sources	4
‘Forest tundra’ or ‘Polygonal tundra’ or ‘Tundra’ or […]	qualitative	best guess with few information available	5 = worst quality

Table 4.10: *A metadata table on the information of ice content could look like this, although final decisions on the quality indexes for the different sources are still pending*

Ice Content (I-C)	I-C detail	I-C origin	I-C quality index
'40-60 %'	quantitative	PI, CALM metadata, publication	1 = best quality
'High Ice content' or 'Medium Ice content' or 'Low Ice content'	qualitative	Indirect retrieval of ice content (in classes?), published Site Pictures, CALM landscape description	2
'High Ice content' or 'Medium Ice content' or 'Low Ice content'	qualitative	Information on comparable locations close by (e.g. same landscape type)	3
'High Ice content' or 'Medium Ice content' or 'Low Ice content'	qualitative	high spatial resolution satellite data, other sources	4
'High Ice content' or 'Medium Ice content' or 'Low Ice content'	qualitative	best guess with few information available	5 = worst quality

Table 4.11: *A metadata table on the information of lithostratigraphy could look like this, although final decisions on the quality indexes for the different sources are still pending*

Lithology and Texture (L-T)	L-T detail	L-T origin	L-T quality index
'60 %' silt [0 – 1 m] '40 %' sand [0 – 1 m] … [1 – 2 m]	quantitative	PI, publication	1 = best quality
'silt-dominance' or 'sand-dominance' or 'clay-dominance'	qualitative	Indirect retrieval of dominance of lithography sand, silt, clay content published Site Pictures, CALM landscape description	2
'silt-dominance' or 'sand-dominance' or 'clay-dominance'	qualitative	Information on comparable locations close by (e.g. same landscape type)	3
'silt-dominance' or 'sand-dominance' or 'clay-dominance'	qualitative	?	4
'silt-dominance' or 'sand-dominance' or 'clay-dominance'	qualitative	best guess with few information available	5 = worst quality

4.3 Validation Strategies

4.3.1 Point-wise validation of the ECVs permafrost temperature and active layer thickness

The match-up process can be compared to the meteorological validation using FRM in the QA4EO sense that a match-up represents a measurement of a traceable variable in space and time that can adequately be matched by another measurement of the same variable if it is sufficiently close in space and time. The match-up is carried out pairwise time- and depth-specific. A direct comparison between the 'match-ups' at the individual sites still suffers from the scale incompatibility between the local representativeness of the reference measurement and the km-scale of the EO-derived Permafrost_cci product.

Validation using ground temperature time series: In year 1, we conducted point-wise site-specific match-up analyses per borehole or soil temperature profile-derived mean, minimum and maximum temperature per standardised depth and per year versus Permafrost_cci ensemble derived mean, minimum and maximum ground temperature per depth and year, providing average bias, absolute error, standard deviation and RMSE and more statistical metrics [RD-3].
We focus on match-up datasets down to 10 m depth for i) gaining the widest value range between minimum and maximum ground temperature that towards deeper depths gets reduced until zero variability at the depth of zero annual amplitude ZAA, ii) enabling RR for climate and land surface models that do not contain adequate parameterisation of the deeper sub-ground and where the simulation skills deteriorate towards deeper depths iii) enabling RR for the EO microwave-derived product that contains signal information on the shallow subsurface. The results of the first depth- and time specific validation activities [RD-3] revealed that match ups are more robust for depths of 40 cm and deeper. For mountain permafrost, the surface temperature at 0 m was valid for assessing the dynamics in time, but deeper ground temperature time series did not correspond to the temporal dynamics [RD-3]. We therefore exclude the shallow depths down to 40 cm from the validation data set for lowland permafrost landscapes but keep the shallow depth-time series for assessing mountain permafrost GT:
In case of well-equipped GTN-P measurement fields (e.g., Nadym (RU), Svalbard, and others and PERMOS measurement fields with multiple temperature measurements) it will be interesting to investigate summer versus winter periods and monthly averages.

Validation using active layer thickness time series and thaw depth measurements: We conducted point-wise site-specific match-up analyses per CALM measurement grid with time series on active layer thickness – the mean grid value of maximum thaw depth per year (measured grid-wise on 10 m x 10 m, 100 m x 100 m or 1000 m x 1000 m fields). In year 1, these CALM active layer thickness grid data per year were matched with Permafrost_cci ensemble derived active layer thickness per year providing average bias, standard deviation and RMSE [RD-3]. In Permafrost_cci year 2, we will conduct point-wise site-specific match-up analyses per site of a thaw depth measurement at a specific measurement date versus Permafrost_cci ensemble means-derived thaw depth at a specific date.

4.3.2 Functional validation of the ECVs permafrost temperature and active layer thickness

Since the spatial variation of ground temperature and thaw depth is related to surface and sub-ground characteristics such as lithology, moisture or vegetation cover that may be highly heterogeneous, site

observations may not represent the Permafrost_cci km-scale output. There are inherent inconsistencies in making the direct comparison between the EO observation at km scale and a point value.
In year 2 und 3, we will undertake match-up analyses using single Permafrost_cci product endmembers to find the best performing model endmember per site, or per lithostratigraphy, vegetation cover, ground ice content.
This prepares for the functional validation on circum-Arctic scale providing relative error per pixel by classifying the site-specific match ups over different classes (land cover, soil, topography). Coming from these classes to class specific relative errors per pixel, this method can provide the relative error per pixel in the CCI product on a circum-Arctic scale.

4.3.3 Binary point- and grid-wise validation of Permafrost abundance

Validation using binary point- and grid-wise observation of permafrost abundance
Binary point- and grid-wise observation of permafrost abundance from geophysical transects, geomorphology (e.g. patterned ground, rock glacier, pingo, other surface feature) and ground temperature provided by users will be compared with the Permafrost_cci product of permafrost fraction. Geophysical measurements of the sub-ground are the standard method for permafrost detection in discontinuous permafrost areas and are routinely carried out for permafrost change detection enabling a temporal validation of the Permafrost_cci product. The regional assessment will be carried out in regions with existing ground data availability and regional expertise provided by members of the Permafrost_cci user community, specifically through the IPA Permafrost Mapping Action Group [RD-6].
In [RD-3] we conducted a first preliminary binary point-wise and time-wise match-up assessment of Permafrost Extent. We allowed a small variability around MAGT 0 °C not setting “permafrost” strictly as in situ MAGT <0 °C in 2 consecutive years. We compared simulated Permafrost_cci MAGT to in situ MAGT at all depths down to 240 cm, analysing the amount of simulated and measured temperatures being both ≤0.5 °C (“permafrost”) or both >0.5 (“no permafrost”). We analysed the bulk data set and additionally the “warm” (MAGT >0 °C) and the “cold” (MAGT <1 °C) temperature groups. As Permafrost_cci MAGT is simulated ‘too cold’ for the ‘warm temperature range of MAGT >0 °C, the match-up of Permafrost extent showed the lowest accuracy performance. Also with shallow depths <40 cm excluded, data characteristics of in situ MAGT data set do not match with any of the five Permafrost_cci MAGT ensembles. The binary match-up is characterised by a relatively high uncertainty of around 65 % (within the 5 to 95 % Quantile).

4.3.4 Potential of temporal dynamics for validation experiments

In general, ground temperatures in shallow depths are characterised by a wider spread between minimum and maximum annual ground temperature and high inter-annual variability (Figure 4.10, example from PERMOS). Also, CALM time series of active layer thickness show high inter-annual variability depending on the annual air temperature and precipitation regimes.
Permafrost lands vary greatly in horizontal space and the properties of each of the three conceptual layers (permafrost, active layer and buffer layer) and their interfaces can also change over short length scales (Vincent et al. 2017). Ground temperature distribution may be heterogeneous due to e.g. variability in ground ice content, lithology and moisture conditions, like in wet versus drier regimes. Some PIs in the permafrost research community specifically set up permafrost measurement fields on heterogeneous terrains spanning dry to wet regimes to optimally investigate functional relationships

between the three layers (permafrost, active layer, buffer layer) and the atmosphere. Examples are, for example GIPL permafrost measurement fields on the North Slope, Alaska such as Franklin Bluff, or in Russia the Nadym and Vaskiny Dachy permafrost measurement fields and the AWI permafrost measurement fields in Svalbard (Boike et al. 2018) and in the Lena River Delta (Boike et al. 2019). A first analysis of the GIPL Franklin Bluff ground temperature records shows the difference in shallow ground temperatures between a wet and a dry site that are located close together and their high inter-annual variability (Figure 4.11). Within the validation approach, we will identify, if Permafrost_cci will be able to reproduce warm and cold MAGT or deep or shallow active layer thickness years.

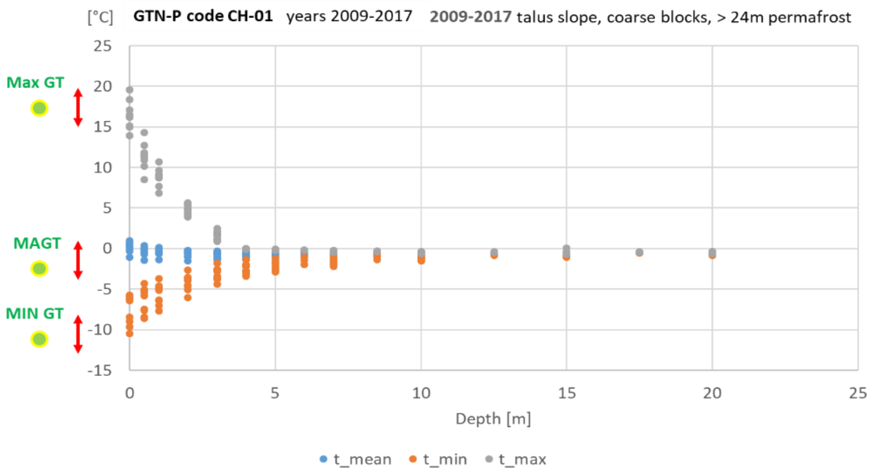

Figure 4.10: *High inter-annual variability of MAGT, min annual GT and max annual GT in mountain permafrost*

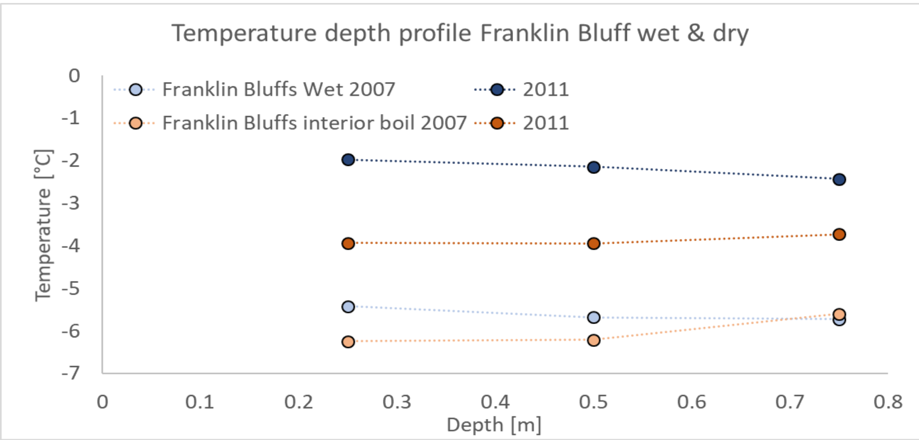

Figure 4.11: *High inter-annual and spatial variability of MAGT at Franklin Bluff (data from Wang 2018)*

4.3.5 Data filtering

In the course of our validation activity, we will identify characteristics of in situ sites and conduct the following steps:

Finished:

- Exclude warm sites with MAGT >1 °C
- Exclude shallow depths <40 cm

In process:

Gather information on sites, including

- Lithology
- Topography (mountain or lowland)
- Vegetation cover
- Nearby settlements, industry…
-

In case of

- Sites nearby rivers or coast
- Problematic coordinates

we will check the plausibility of the location
(e.g. a borehole located in a river sandbank is less plausible than to shift it to the nearby airport, as especially old Russian boreholes are often located near weather stations of airports) and the surrounding Permafrost_cci variable grid cells (if a grid cell deviates from the surrounding landscape and the associated borehole is biased, than it might be appropriate, to shift the location to an adjacent grid-cell.

We will furthermore group sites according to their characteristics to deeply investigate deviations of the Permafrost_cci variables. It might also be necessary to remove strongly influenced sites, e.g. situated close to a big river or lake and thus not representing the landscape scale temperature, from the in situ dataset.

4.4 First Validation Results

4.4.1 Permafrost Temperature

The match-up of sites with MAGT <1 °C and depths ≥ 40 cm gave a model bias of ~ −0.12 °C for the pairwise depth- and time-specific calculation and an RMSE of 1.72 °C. The relative percentage error is lower than 5 % (within the 5 to 95 % Quantile), absolute percentage error ~ 60 %.

Permafrost_cci ground temperatures are also colder than MAGT retrievals from satellite derived surface state (FT2T model). The average difference (all depths and sites) is -1.7 °C. The direct comparison reveals regional differences. They are largest for northern central Siberia and the Canadian High Arctic. Warmer in situ regions (MAGT>1 °C) are not well captured by the simulated temperature range (Figure 4.12) and thus not incorporated in validation procedures. Furthermore, depth <40 cm are discarded in the validation procedure, as the topographic precision of shallow depths cannot be given.

Despite the overall good results of MAGT match-up in depths ≥40 cm and sites with <1 °C in situ GT, several outliers can still be detected (Figure 4.13, 4.14). Out of the 145 datasets, 30 % have a bias and SD of bias < 1 °C.

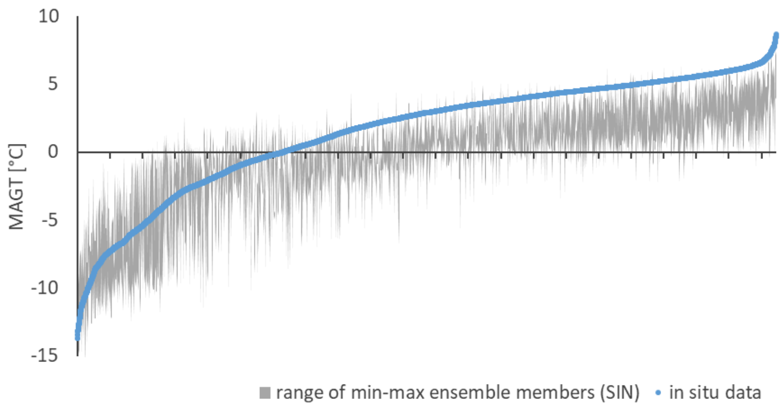

Figure 4.12 *Match up data set characteristics: in situ data MAGT (blue points) and Permafrost_cci MAGT min to max temperature from 5 ensemble members (grey-shaded)*

4.4.2 Active layer thickness

Permafrost_cci ALT match-up shows a moderate positive model bias of ~ 0.35 cm and RMSE of ~ 1 m if calculated for the bulk data collection, a relative percentage error of ~ 76 % (within the 5 to 95 % Quantile), absolute percentage error ~ 90 %.

There is one type of overestimation linked to either Arctic rock and stone desserts in Svalbard and Greenland or to valley bottoms in mountain regions with shallow in situ ALT measured due to the fine-grained lithology but relatively warm in situ MAGT temperatures across all latitudes (with exception of the southern latitudes). The Permafrost_cci ALT underestimation is mainly linked to the discontinuous (e.g., Norway, Spitsbergen) and southern boundaries (e.g., China) of permafrost.

4.4.3 Permafrost Probability

The binary match-up of “permafrost” versus “no permafrost” for Permafrost_cci PFR permafrost probability versus in situ MAGT ranges shows that permafrost probability in the grid cell is overestimated compared to in situ-derived “no permafrost” and MAGT ≤0.5 °C. Permafrost_cci PFR permafrost probability in the grid cell >0 % occurs together with a wide range of “warm” in situ MAGT >0 °C. A large fraction of Permafrost_cci permafrost probability grid cells >60 % occurs together with an in situ MAGT range from 0 to 5 °C occurring at regional scales that are already independent from pixel-scale heterogeneity.

Less permafrost area than in the GlobPermafrost extent product is estimated in all classes except the discontinuous zone where Permafrost_cci extent based on temperature in two meter depth assigns nearly 1 Million km² more area.

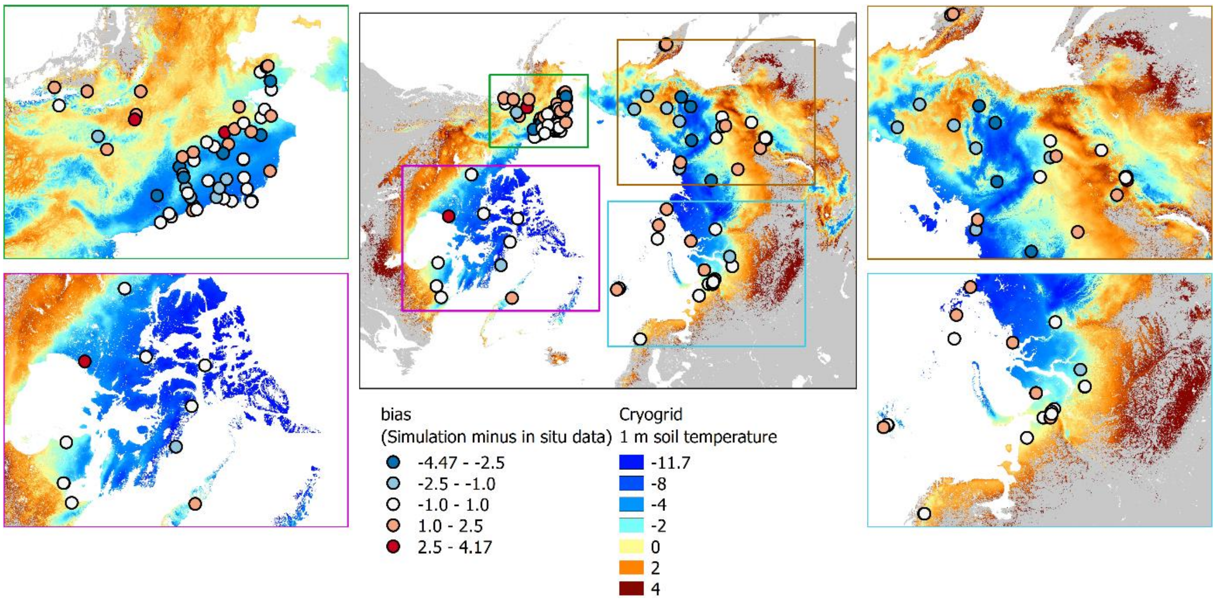

Figure 4.13 *Overview of sites with an in situ MAGT <1 °C and depths ≥40 cm. Map background Permafrost_cci MAGT in 1 m depth for the year 2017. Coloured dots depict the range of bias (Permafrost_cci MAGT minus in situ data) over all years and all depths.*

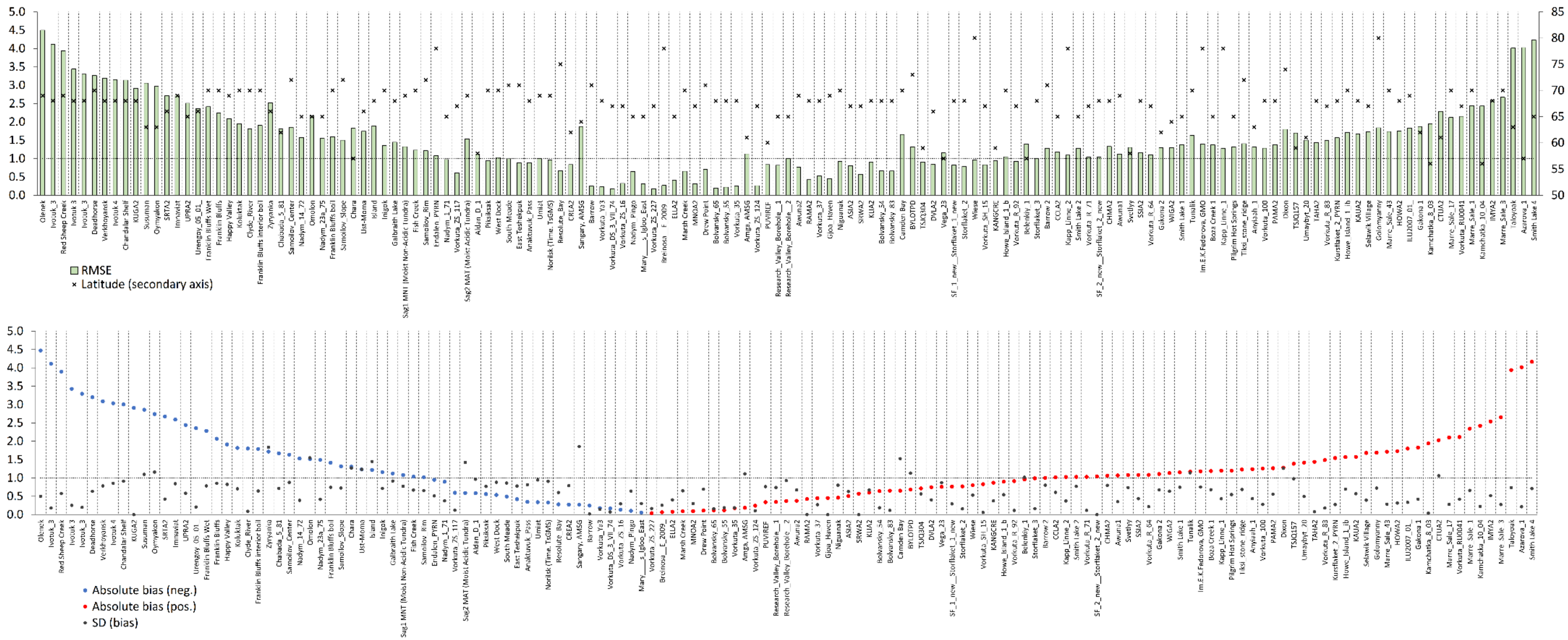

***Figure 4.14** Various statistics between Permafrost_cci MAGT mean and in situ MAGT over all years (2003-2017) and all depths ≥40 cm, calculated for sites with an in situ MAGT <1 °C. Upper panel: Residual mean square error (RMSE, bars) and latitude (black stars, on the secondary axis). Lower panel: Absolute bias and SD of bias (grey dots). Negative bias (Permafrost_cci MAGT colder than in situ MAGT) are given in blue, positive bias (Permafrost_cci mean MAGT warmer than in situ MAGT) in red dots. The black horizontal line indicates 1 °C.*

While no latitudinal trend is detectable, Permafrost_cci ALT shows a large positive bias at high and also low latitude sites. Permafrost_cci MAGT shows a negative bias and/ or a high standard deviation (i.e. a high yearly variability of the bias) frequently at sites close to surface waters. Other sites have only a small bias in MAGT, but with a high standard deviation (e.g. site Sangary AMSG, Figure 4.15)

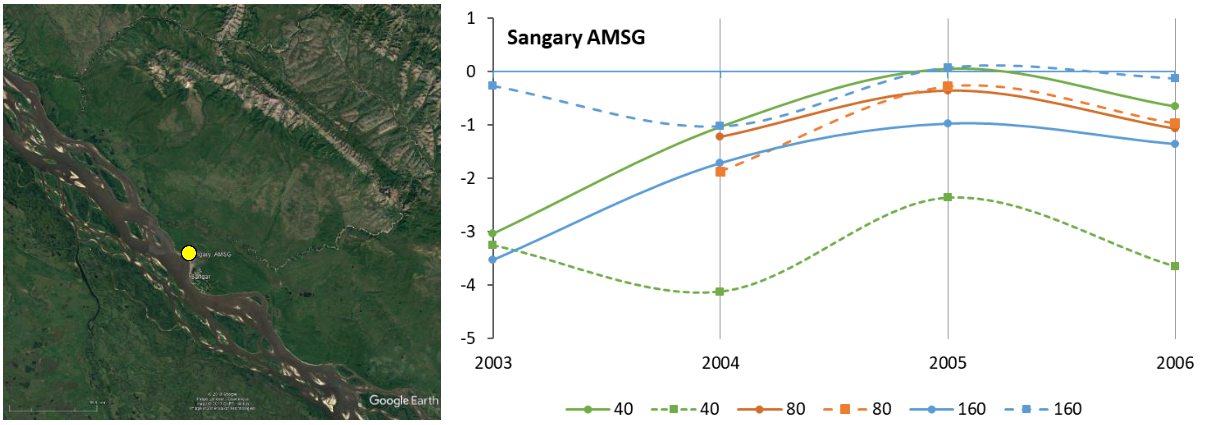

Figure 4.15 *Left Panel: Yellow point=Locality of Sangary AMSG (Dataset 24652 from RHM, Long 127.46, Lat 63.96),Map source: Google Earth Pro ©, Version 7.3.2. Right panel: MAGT of site Sangary AMSG for depths of 40, 80 and 160 cm. Solid lines= in situ mean MAGT [°C], broken lines= Permafrost_cci mean MAGT [°C]. The deviation of Permafrost_cci and in situ MAGT data varies both, between depths and between years (mean bias -0.27 °C, SD of bias 1.86 °C).*

In multi-site permafrost networks, nearby sites can show very different magnitudes of bias (Figure 4.16). The reason for such spatial small-scale variations can be found in small scale differences of topography, soil or vegetation cover in comparison to the 1 km² grid cell size.

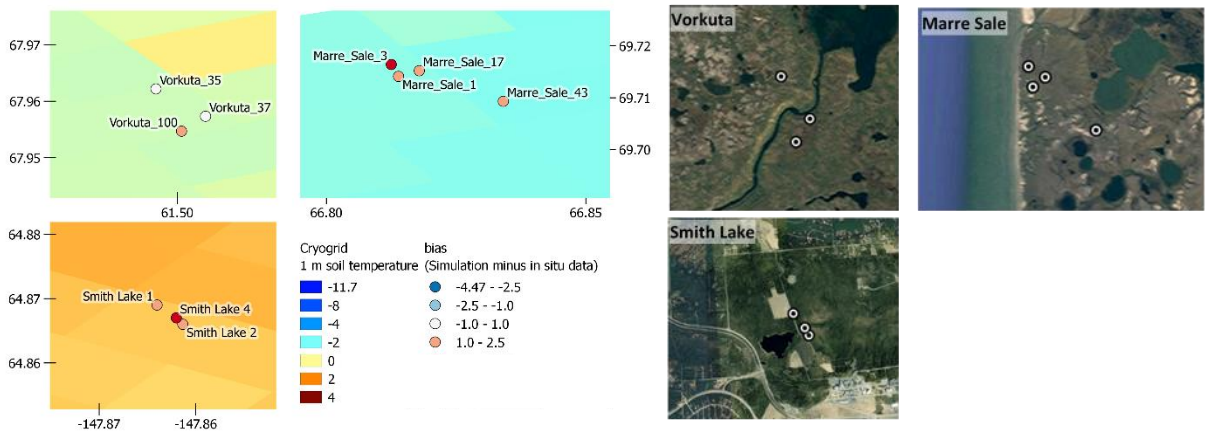

Figure 4.16 *Overview of sites with an in situ MAGT <1 °C and depths ≥40 cm. Map background gives Permafrost_cci MAGT in 1 m depth for the year 2017. Coloured dots depict the range of bias (Permafrost_cci MAGT minus in situ data) over all years and all depths. Right: satellite imagery from Google Earth Pro ©. Version 7.3.2.*

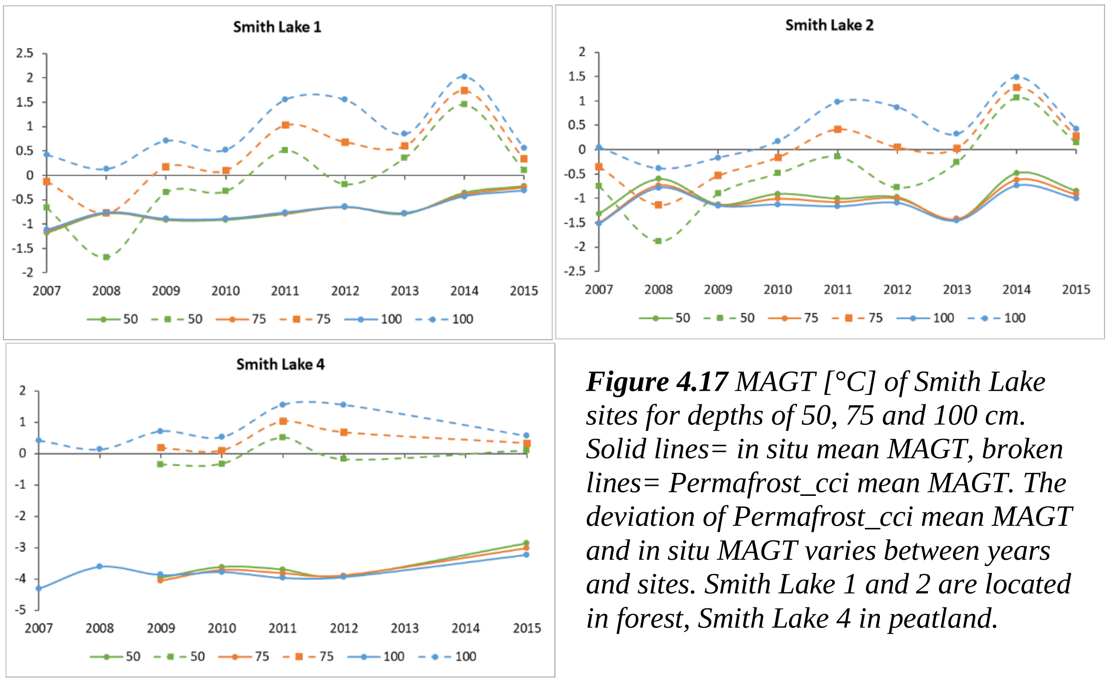

Figure 4.17 *MAGT [°C] of Smith Lake sites for depths of 50, 75 and 100 cm. Solid lines= in situ mean MAGT, broken lines= Permafrost_cci mean MAGT. The deviation of Permafrost_cci mean MAGT and in situ MAGT varies between years and sites. Smith Lake 1 and 2 are located in forest, Smith Lake 4 in peatland.*

4.5 Validation experiment in Mountain Permafrost

The validation and evaluation efforts are also carried out in high-mountain permafrost regions. Binary point and grid-wise regional comparison to ground temperature measurements, geophysical transects and regional inventories of rock glaciers including the kinematic state (or active rate) are performed. In addition to the PERMOS borehole temperature data, the EO derived inventory on rock glacier abundance, extent, and creep, which was developed by the ESA DUE GlobPermafrost program since 2016 and continues in Permafrost_cci is used to validate the binary permafrost extent product in mountain areas. The GTN-P PERMOS monitoring data and the EO derived rock glacier inventory supports the validation of Permafrost_cci products in mountain areas, where the Permafrost_cci products contain the highest uncertainties.

4.4.1 PERMOS mountain permafrost network

Amongst the 35 GTN-P mountain permafrost boreholes, 27 belong to the Swiss permafrost monitoring network PERMOS. The PERMOS boreholes cover the whole range of typical mountain permafrost landforms (i.e. talus slope, rock glacier, rock walls, mountain crest and summit) (Table 4.4) and are spatially distributed over the different geographical region within the Swiss Alps (Figure 4.18).

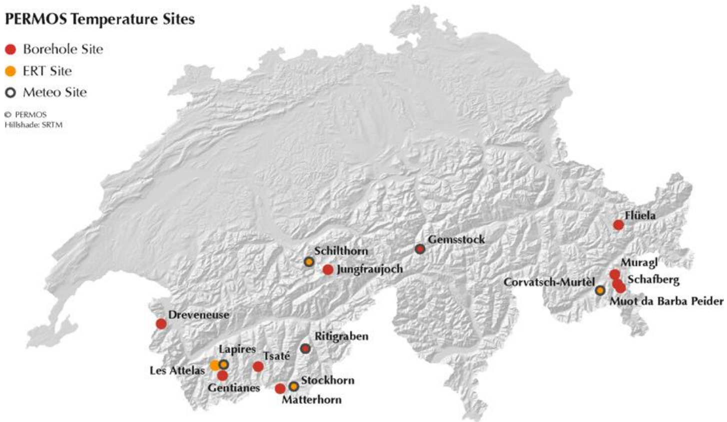

Figure 4.18: Overview on PERMOS borehole and geophysical and meteorological measurement sites.

The longest record totalises more than 30 years of observation, whereas the majority of the PERMOS boreholes has between 10 and 23 years of observation. In addition to borehole temperatures, the PERMOS network also collects long-term observations of ground surface temperature, permafrost creep velocities, permafrost resistivities and meteorological data in the Swiss Alps.

4.4.2 Binary point- and grid-wise validation of permafrost abundance

The binary validation approach using active rock glacier abundance (Figure 11) is solely based on remote sensing products (e.g. optical images or InSAR) and thus well suited for regional validation in any remote mountain area. However, such inventories as is are not usable for temporal validation since active rock glaciers will be on the same place for decades. To improve this point, it is suggested to develop regional indices of kinematic evolution based on velocity changes observed at large scale using EO InSAR data. The rock glacier creep rate (kinematics) being dependent on the permafrost temperatures, this approach would enable region-wide temporal model validation in mountain permafrost (Figure. 12).

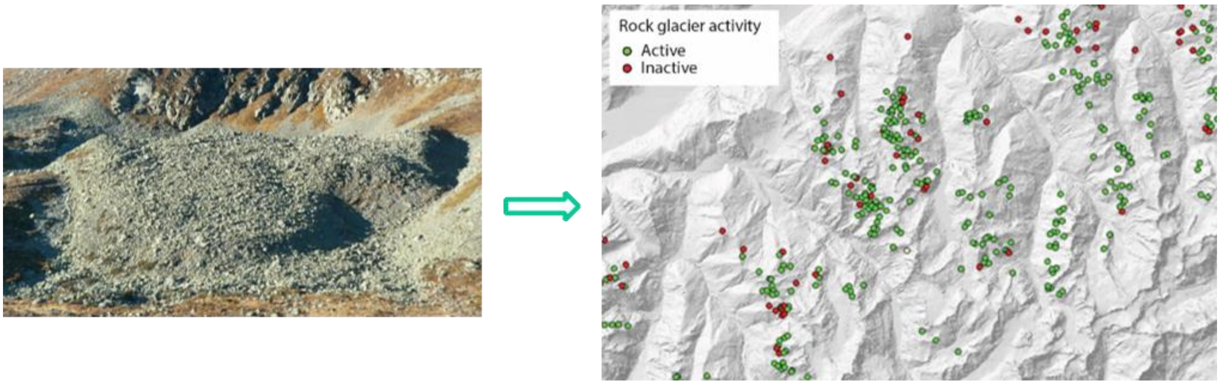

Figure 4.19: *Example of typical rock glacier (left panel) and rock glacier inventory in the Swiss Alps, which includes the state of activity of each landform (right panel).*

The regional assessment will be carried out specifically in regions with existing ground data or EO data availability as well as regional expertise by PERMOS and the Permafrost_cci Mountain Permafrost team. This work will be undertaken in close collaboration with the ongoing **IPA-funded Action Group on Rock Glacier Inventories and Kinematics** (2018-2020), which aims at i) defining widely accepted standard guidelines for inventorying rock glaciers in mountain permafrost regions, including information on the activity rate and ii) promoting the use of satellite SAR interferometry, e.g. Sentinel data, for monitoring the rock glacier activity at a regional scale. The latter objective also entails to set up standard guidelines for selecting an appropriate number of rock glaciers per region that can be used to assess temporal trends with decadal to intra-decadal time steps.

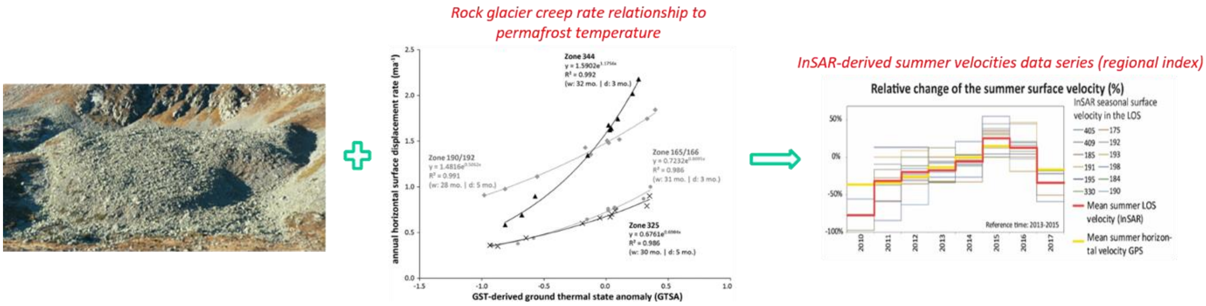

Figure 4.20: *Schematic representation of the proposed regional kinematic indices computation.*

5 VALIDATION DOCUMENTS AND ENDORSEMENTS

Table 5.1 provides an overview on deliverables with information on product validation and the results of the algorithm selection. Apart from those that are already part of the project deliverables, we also seek for documenting the results in additional publications, such as a peer-reviewed paper in a scientific journal (validation and Round Robin RR results of all products). Whereas the former will be prepared by the Permafrost_cci consortium, the latter will be prepared together with the interested community and PI data providers of reference data sets. We will seek for an open review process of all results achieved by informing the respective group of scientists and stakeholders such as the IPA, specifically involvement of the **IPA Permafrost Mapping Action Group**, **IPA Rock Glacier Inventory and Kinematics Action Group**, the RR participants, and the CRG, when these documents are accessible. We will also seek outreach via the international community relevant mailings lists Permalist and Cryolist to participate in the validation activities. If the results of the validation and RR activities of the individual Permafrost_cci products can be presented in form of publications and data publications, the largest possible endorsement is achieved.

Table 5.1: *Documents related to validation of the Permafrost_cci product.*

Deliv. No.	Name	Date	Comment
D1.3	DARD	January 2019 November 2020	describes data accessibility
D2.1	PVASR	February 2019 November 2019 May 2021	summarises Round Robin (RR) experiments and algorithm selection
D2.3	E3UB	February 2019 November 2019 March 2021	defines sources of errors and uncertainties
D2.5	PVP	February 2019 November 2019 November 2020	outlines planned validation strategies
D4.1	PVIR	September 2019 September 2020 May 2021	provides a summary on quality and uncertainty of ECV products
D4.2	CRDP	May 2020 February 2021	describes the Climate Research Data Package, a fully uncertainty characterised, long time series of Permafrost_cci products in compliance with CCI Data standards. The validation reference data will be part of the CRDP.
D4.3	PUG	August 2019 August 2020 February 2021	describes delivered Permafrost_cci products in the CRDP
D5.2	CAR	October 2019 October 2020 May 2021	describes the Climate Science study cases using the CCI products and the user's feedback. Validation, specifically the validation and upscaling experiments in lowland permafrost and mountain permafrost, will be part of the Climate Science studies.

6 REFERENCES

6.1 Bibliography

Allard, M., Sarrazin, D., L'Hérault, E., 2016: Borehole and near-surface ground temperatures in northeastern Canada, v. 1.4 (1988-2016). Nordicana D8, doi: 10.5885/45291SL-34F28A9491014AFD.

Biskaborn, B.K.; Smith, S.L.; Noetzli, J.; Matthes, H.; Vieira, G.; Streletskiy, D.A.; Schoeneich, P.; Romanovsky, V.E.; Lewkowicz, A.G.; Abramov, A.; Allard, M.; Boike, J.; Cable, W.L.; Christiansen, H.H.; Delaloye, R.; Diekmann, B.; Drozdov, D.; Etzelmüller, B.; Grosse, G.; Guglielmin, M.; Ingeman-Nielsen, T.; Isaksen, K.; Ishikawa, M.; Johansson, M.; Johannsson, H.; Joo, A.; Kaverin, D.; Kholodov, A.; Konstantinov, P.; Kröger, T.; Lambiel, C.; Lanckman, J.-P.; Luo, D.; Malkova, G.; Meiklejohn, I.; Moskalenko, N.; Oliva, M.; Phillips, M.; Ramos, M.; Sannel, A.B. K.; Sergeev, D.; Seybold, C.; Skryabin, P.; Vasiliev, A.; Wu, Q.; Yoshikawa, K.; Zheleznyak, M., Lantuit, H., 2019, Permafrost is warming at a global scale. Nature Communications, 10, 264

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

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

Boike, J., Juszak, I., Lange, S., Chadburn, S., Burke, E., Overduin, P.P., Roth, K., Ippisch, O., Bornemann, N., Stern, L., Gouttevin, I., Hauber, E., Westermann, S., 2018, A 20-year record (1998–2017) of permafrost, active layer and meteorological conditions at a high Arctic permafrost research site (Bayelva, Spitsbergen), Earth Syst. Sci. Data, 10, 355-390, https://doi.org/10.5194/essd-10-355-2018, 2018.

Brown, J., Hinkel, K.M., Nelson, F.E., 2000, The Circumpolar Active Layer Monitoring (Calm) Program: Research Designs and Initial Results. Polar Geography 24, 166–258. https://doi.org/10.1080/10889370009377698

Brown, J., Ferrians, O.J., Heginbottom, J.A., Melnikov, E.S., 1997, Circum-arctic map of permafrost and ground ice conditions. International Permafrost Association, US Geological Survey.

Brown J., 1970, Permafrost in Canada: Its influence on northern development. University of Toronto Press, Toronto 234 p.

Center for Northern Studies in Canada, CEN, 2013: Environmental data from Northern Ellesmere Island in Nunavut, Canada, v. 1.0 (2002-2012). Nordicana D1, doi.org/10.5885/44985SL-8F203FD3ACCD4138.

Clow, G.D. 2014, Temperature data acquired from the DOI/GTN-P Deep Borehole Array on the Arctic Slope of Alaska, 1973–2013, Earth Syst. Sci. Data, 6, 201–218.

Fagan, J.D. and Nelson, F.E., 2017, Spatial Sampling Design in the Circumpolar Active Layer Monitoring Programme. Permafrost and Periglacial Processes 28, 42–51. https://doi.org/10.1002/ppp.1904

Heginbottom, J.A., 1984, The mapping of permafrost. Canadian Geographer, 28, 1, 78-83.

Heginbottom, J.A., and Radburn, L.K., 1992, Permafrost and Ground Ice Conditions of Northwestern Canada (Mackenzie Region). National Snow and Ice Data Center, Boulder, CO, USA.

Muller, S.W, 1943: Permafrost or permanently frozen ground and related engineering problems. U.S. Engineers Office, Strategic Engineering Study, Special Report No. 62. 136p. (Reprinted in 1947, J. W. Edwards, Ann Arbor, Michigan, 231p.)

Nelson, F.E. and Hinkel, K.M., 2003, Methods for measuring active-layer thickness. in: O. Humlum and N. Matsuoka, eds., a Handbook on Periglacial Field Methods. Longyearbyen, Norway: University of the North in Svalbard

van Everdingen, R.O. (1976) Canadian Journal of Earth Sciences, 13:862-867, https://doi.org/10.1139/e76-089

van Everdingen R.O., 1985, Unfrozen permafrost and other taliks. Workshop on Permafrost Geophysics, Golden, Colorado, October 1984 (J. Brown, M.C. Metz, P. Hoekstra, Editors). U.S. Army, C.R.R.E.L., Hanover, New Hampshire, Special Report 85-5, pp.101-105

van Everdingen, Robert, ed. 1998 revised May 2005. Multi-language glossary of permafrost and related ground-ice terms. Boulder, CO: National Snow and Ice Data Center/World Data Center for Glaciology. (http://nsidc.org/fgdc/glossary/; accessed 23.09.2009)

Vincent, W., Lemay, F., and Allard, M. 2017, Arctic Permafrost Landscapes in Transition: Towards an Integrated Earth System Approach. Arctic Science 3(2): 39–64. https://doi.org/10.1139/as-2016-0027.

Wang, K., 2018. A Synthesis Dataset of Near-Surface Permafrost Conditions for Alaska, 1997-2016. https://doi.org/10.18739/a2kg55

Wang, K., Jafarov, E., Overeem, I., Romanovsky, V., Schaefer, K., Clow, G., Urban, F., Cable, W.L., Piper, M., Schwalm, C., Zhang, T., Kholodov, A., Sousanes, P., Loso, M., Hill, K., 2018, A Synthesis Dataset of Permafrost-Affected Soil Thermal Conditions For Alaska, USA. Earth Syst. Sci. Data, 10, 2311–2328. https://doi.org/10.5194/essd-10-2311-2018

Williams, J.R., 1965, Ground water in permafrost regions: An annotated bibliography. U.S. Geological Survey, Professional Paper 696, 83p.

6.2 Acronyms

AD	Applicable Document
AWI	Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research
B.GEOS	b.geos GmbH
CALM	Circumpolar Active Layer Monitoring
CCI	Climate Change Initiative
CEOS	Committee on Earth Observing Satellites
CEN	Canadian Centre d'études Nordiques
CMUG	Climate Modelling User Group
DUE	Data User Element
ECV	Essential Climate Variable
EO	Earth Observation
ESA	European Space Agency
FRM	Fiducial Reference Measurement
GAMMA	Gamma Remote Sensing
GCOS	Global Climate Observing System
GIPL	Geophysical Institute Permafrost Laboratory
GT	Ground Temperature
GTN-P	Global Terrestrial Network for Permafrost
GUIO	Department of Geosciences University of Oslo
InSAR	Interferometric Synthetic Aperture Radar
IPA	International Permafrost Association
LST	Land Surface Temperature
MAGT	Mean Annual Ground Temperature
maxAGT	maximum Annual Ground Temperature
minAGT	minimum Annual Ground Temperature
NSF	National Science Foundation
NSIDC	National Snow and Ice Data Center
OSCAR	Observing Systems Capability Analysis and Review Tool
PERMOS	Swiss Permafrost Monitoring Network
PVP	Product Validation Plan
R	Requirement
RD	Reference Document
RHM	Roshydromet
RMSE	Root Mean Square Error
RR	Round Robin
RRR	Rolling Requirements Review
SI	International System of Units
SWE	Snow Water Equivalent
std dev.	Standard Deviation
TSP	Thermal State of Permafrost
TR	Technical Requirement
UAF	University of Alaska, Fairbanks

UNIFR	Department of Geosciences University of Fribourg
QA4EO	Quality Assurance framework for Earth Observation
WGCV	Working Group on Calibration and Validation
WMO	World Meteorological Organization