# CCI+ PHASE 2– NEW ECVS

# Permafrost

# D2.5 Product Validation Plan (PVP)

**Version 4.0**

**31 May 2023**

**Prepared by**

# Document status sheet

| Issue | Date | Details | Authors |
|---|---|---|---|
| 1.0 | 28.02.2019 | Overview on in situ data to be used and methods planned for validation for lowland and mountain permafrost with focus on permafrost temperature | BH, MW, CP, RD, CB, SW, AB |
| 2.0 | 30.11.2019 | Update on in situ data to be used and success of methods being applied in the year 1 validation for lowland and mountain permafrost with focus on permafrost temperature | BH, MW, AB |
| 3.0 | 22.12.2020 | Update on in situ data to be used and success of methods being applied in the year 2 validation for lowland and mountain permafrost with focus on permafrost temperature, planned method of temporal stability | BH, MW, AB |
| 4.0 | 31.05.2023 | Update on in situ data to be used for phase II and review of methods applied in the phase I year 3 validation for lowland and mountain permafrost with focus on permafrost temperature, application of temporal stability | BH, MW, CP, 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

EUROPEAN SPACE AGENCY CONTRACT REPORT

The work described in this report was done under ESA contract. Responsibility for the contents resides in the authors or organizations that prepared it.

# TABLE OF CONTENT

# EXECUTIVE SUMMARY

This document presents the Product Validation Plan (PVP) of the European Space Agency (ESA) Climate Change Initiative (CCI) Permafrost project (Permafrost_cci). CCI represents ESA's global monitoring program with a major aim to provide Earth Observation (EO)-based Essential Climate Variable (ECV) time series serving the climate modeling and climate user communities. Permafrost_cci was part of phase I of CCI+ (2018-2021) and has been selected for phase II (2022-2025) with the production of ECVs, set by the Global Climate Observing System (GCOS)/World Meteorological Organisation (WMO) for permafrost. The PVP describes the reference data and validation strategies for the Permafrost_cci products: i) permafrost temperature expressed as Ground Temperature per Depth (GTD) [C°] ii) Active Layer Thickness (ALT) [cm] and iii) permafrost extent expressed as Permafrost FRaction (PFR) [%].

The Committee on EO Satellites (CEOS) Working Group on Calibration and Validation (WGCV) defines validation as 'the process of assessing, by independent means, the quality of the data products derived from the system outputs' (*lpvs.gsfc.nasa.gov*). According to the CEOS Quality Assurance framework for Earth Observation (QA4EO) and ESA CCI guidelines, the validation data need to be independent from the product generation. In the QA4EO sense, suitable reference data are characterized by protocols and community-wide management practices, and published openly. In Permafrost_cci accordingly, assessments of the Permafrost_cci products are carried out independently from the algorithm development team using in situ data from the WMO/GCOS Global Terrestrial Network for Permafrost (GTN-P) managed by the International Permafrost Association (IPA). Within the GTN-P/IPA framework, the Thermal State of Permafrost Monitoring (TSP) program is managing the temperature monitoring via borehole temperature profiles and shallow ground temperature profiles, whereas the Circumpolar Active Layer Monitoring program (CALM) is providing global monitoring for ALT via standardized measurement grids. Both GTN-P monitoring programs, TSP and CALM, fulfill QA4EO criteria by their standards for measurements, data collection and open data publication practices. Permafrost_cci also specifically involves the mountain permafrost monitoring program GTN-P/PERMOS in Switzerland to cope with the challenge of validation of the Permafrost_cci products in mountainous regions providing PERMOS permafrost monitoring data at highest quality levels. In addition, we incorporated in situ data collections from individual Principal Investigators (PIs) and additional national ground monitoring programs in the Permafrost_cci reference data set.

The PVP describes available in situ data collections and measurement techniques and the Permafrost_cci reference data set preparation. The methods for processing the Permafrost_cci reference data were already successful in Permafrost_cci phase I and are being extended and continued in phase II. As a standard, quality assessments of Permafrost_cci products are carried out by point-wise match-ups of location, measurement depth and equivalent year using standard statistics (such as bias, absolute error, relative percentage error, root mean square error). Permafrost_cci is also innovatively undertaking assessments in comparing Permafrost_cci GTD with EO-derived Freeze-Thaw to Temperature (FT2T) and for mountain permafrost areas using the EO-derived inventories on rock glacier occurrence, which was developed by the ESA Data User Element (DUE) GlobPermafrost team since 2016 and which is continued in Permafrost_cci Phase I and worldwide in 12 mountain regions in phase II.

# 1 INTRODUCTION

## 1.1 Purpose of the document

This document is the Product Validation Plan (PVP) version 4 (update of [RD-1]) of the ESA CCI+ project Permafrost_cci. The PVP describes and defines the reference data, and validation methods and strategies used for quality assessments of the Permafrost_cci Climate Research Data Packages (CRDP), following CCI and CEOS Quality Assurance framework for Earth Observation (QA4EO) guidelines [AD-1, RD-2].

Besides the required Permafrost ECVs i) permafrost temperature, and ii) active layer thickness, Permafrost_cci provides iii) permafrost extent (permafrost fraction within a pixel), as an additional variable derived from permafrost temperature: the areal fraction within the grid cell that fulfills the definition for the existence of permafrost (ground temperature <0 ºC for two consecutive years).

The generation of Permafrost_cci CRDP depth-specific ground temperature, GTD, ALT and Permafrost FRaction (PFR) time series relies on the ground thermal model Permafrost_cci CryoGrid, that is forced by EO time series of Land Surface Temperature (LST) and Snow Water Equivalent (SWE) with boundary conditions of EO-derived Land Cover [RD-3].

The Permafrost_cci CRDPv3 will be an update of CRDPv2, released in 2021 [RD-3], that included three time series covering the Northern Hemisphere north of 30° N:

• simulated EO-forced **mean annual Ground Temperature per Depth (GTD) in five discrete depths** (0 m, 1 m, 2 m, 5 m, 10 m) from 1997 to 2019 - will be updated to 2021

• simulated EO-forced **annual Active Layer Thickness (ALT)** from 1997 to 2019 - will be updated to 2021

• **annual Permafrost FRaction (PFR)** derived from GTD from 1997 to 2019 - will be updated to 2021

In Permafrost_cci phase II we will continue validation experiments for mountain permafrost areas using rock glacier abundance and binary-based validation on permafrost abundance similar to validation of mountain permafrost in phase I [RD-1,2].

## 1.2 Structure of the document

Chapter 1 contains introduction and overview on applicable documents, as well as the community glossary for Permafrost. Chapter 2 provides information on how the validation follows the overall project guidelines of CCI. In Chapter 3, the unbiased validation in independence from the algorithm development team is described. Chapter 4 provides the overview on the validation data from GTNP/TSP ground temperature and CALM active layer thickness and from additional PIs and national measurement programs. Chapter 4 further describes the compilation and standardization strategies 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 match-up techniques. Chapter 4 also contains the specific validation required for mountain permafrost, carried out by the Swiss mountain permafrost monitoring network GTN-P/PERMOS. Information on Permafrost_cci validation documents and their endorsement is given in chapter 5.

## 1.3 Applicable documents

[AD-1] GEO/CEOS Quality Assurance framework for Earth Observation (QA4EO) protocols 3-4

[AD-2] ESA 2017: Climate Change Initiative Extension (CCI+) Phase 1 – New Essential Climate Variables - Statement of Work. ESA-CCI-PRGM-EOPS-SW-17-0032

[AD-3] ESA Climate Change Initiative. CCI Project Guidelines. EOP-DTEX-EOPS-SW-10-0002

[AD-4] ECV 9 Permafrost: Assessment report on available methodological standards and guides, 1 Nov 2009, GTOS-62

[AD-5] 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.

## 1.4 Reference documents

[RD-1] Heim, B., Wieczorek, M., Pellet, C., Delaloye, R., Barboux, C., Westermann, S., Strozzi, T. (2020): ESA CCI+ Product Validation Plan, v3.0

[RD-2] Heim, B., Lisovski, S., Wieczorek, M., Pellet, C., Delaloye, R., Bartsch, A., Jakober, D., Pointner, G., Strozzi, T. (2020): ESA CCI+ Product Validation and Intercomparison Report, v3.0

[RD-3] Bartsch, A., Westermann, Strozzi, T., Wiesmann, A., Kroisleitner, C., Wieczorek, M., Heim, B. (2023): ESA CCI+ Permafrost Product Specifications Document, v4.0

[RD-4] 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)

[RD-5] 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-6] Nitze, I., Grosse, G., Heim, B., Wieczorek, M., Matthes, H., Bartsch, A., Strozzi, T. (2019): ESA CCI+ Climate Assessment Report, v1.0

[RD-7] 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-8] Bartsch, A., Matthes, H., Westermann, S., Heim, B., Pellet, C., Onacu, A., Kroisleitner, C., Strozzi, T. (2023): ESA CCI+ Permafrost User Requirements Document, v3.0

[RD-9] 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-10] Heim, B., Wieczorek, M., Pellet, C., Delaloye, R., Bartsch, A., Jakober, D., Pointner, G., Strozzi, T. (2020): ESA CCI+ Product Validation and Intercomparison Report, v2.0

[RD-11] Bartsch, A., Westermann, S., Heim, B., Wieczorek, M., Pellet, C., Barboux, C., Delaloye, R., Kroisleitner, C., Strozzi, T. (2020): ESA CCI+ Permafrost Data Access Requirements Document, v2.0

[RD-12] Rouyet, L., Schmid, L., Pellet, C., Delaloye, R., Onaca, A., Sirbu, F., Poncos, V., Kääb, A., Strozzi, T., Jones, N., Bartsch, A. (2023): CCN4 Mountain Permafrost: Rock Glacier Inventories (ROGI) and Rock Glacier Velocity (RGV) Products Product Specification Document, v1.0

## 1.5 Bibliography

A complete bibliographic list that supports 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 glossary below based on [RD-4] provides a selection of terms relevant for Permafrost_cci [AD-2]. A comprehensive glossary is available as part of the Product Specifications Document [RD-3,4].

**active-layer thickness**

The thickness of the ground layer that is subject to annual thawing and freezing above permafrost.

The thickness of the active layer depends on factors such as the ambient air temperature, vegetation, drainage, soil or rock type and total water content, snowcover, 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 snowcover.

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 noncryotic (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 recognized. 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 uncommon to find that the mean annual ground temperature decreases in the upper 50 to 100 meters 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.

**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 pipe-lines, 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 WMO/GCOS and the science communities is providing confidence in the quality of each CCI product and thoroughly investigating its uncertainties through assessments against independent data such as in situ reference measurements or alternate estimates. In response to the Permafrost_cci Statement of Work [AD-2] and CCI project guidelines [AD-3], we have summarized in Table 2.1 how the validation activities for each of the Permafrost_cci products meet those requirements. In addition, a Rolling Review of Requirements (RRR) process is defined by WMO/GCOS: User requirements for the observing geophysical variables in support of the WMO activities are collated in a comprehensive, technology free and quantitative way in the WMO Observing Systems Capability Analysis and Review (OSCAR) requirements database.

***Table 2.1:*** *Validation in Permafrost_cci compared to the overall project guidelines [AD-3; CCI V1-V6] and as outlined in the Statement of Work (SoW) [AD-2; RD-6, Technical Requirements (TR)]*

| **Recommendations for validation** | **Permafrost Temperature** GCOS ECV | **Active Layer Thickness** GCOS ECV | **Permafrost Fraction** WMO/ OSCAR RRR |
|---|---|---|---|
| CCI [V-1] All CCI projects should use the definition of validation approved by the CEOS-WGCV | Validation for all Permafrost_cci 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 Permafrost_cci 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.<br>Accepted statistical measures are applied (mean, std dev., RMSE) in the validation process. | GTN-P CALM offers protocols on in situ measurements, data processing and data publication.<br>Accepted statistical measures are 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 the assembled Permafrost_cci MAGT and ALT reference data sets. |
| CCI [V-5] The CCI programme should hold a dedicated session (or workshop) on common validation infrastructure during a CCI co-location meeting | A workshop could be held at the International Conference on Permafrost ICOP, Whitehorse, Yukon, Canada in June 2024. However, the number of conference attendees could be limited at the Whitehorse ICOP conference event, and further options can be at the European Conference on Permafrost EUCOP 2025 and AGU conferences in 2024, 2025. | | |

***Table 2.1 cont.***

| **Recommendations for validation** | **Permafrost Temperature** GCOS ECV | **Active Layer Thickness** GCOS ECV | **Permafrost Fraction** WMO/ OSCAR RRR |
|---|---|---|---|
| CCI [V-6], SoW [RD-6], [TR-14]<br><br>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 involves 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 undertook outreach on relevant Polar community and Polar political conferences. Independent assessment was specifically sought from the IPA Permafrost Mapping Action Group in Permafrost_cci phase 1. This collaboration was also documented in the Climate Assessment Report (CAR) [RD-6] where Dr. Isabelle Gärtner-Roer, University of Zurich, CH, at that time Vice president of IPA and leader of the IPA Permafrost Mapping Action Group [RD-5], and Science Officer of the World Glacier Monitoring Service (WGMS), was confirming 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. She considered the validation of Permafrost_cci GTD as the most important as it builds the base for the other products, such as ALT and permafrost extent. Members of the IPA will be engaged in the validation and assessment activities via the Visiting Scientist program in phase 2.<br><br>Peer-reviewed joint papers are planned to document the results. Specifically, the outcome of the collection of reference data is planned to be published in Earth System Science Data ESSD which is highly attractive due to its high impact factor (e.g. Scopus 13.8, accessed April 2023). | | |
| [TR-21] An independent validation shall be performed against metrics 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 (public dissemination stopped) and other international and national networks and data sources (e.g. NASA Above, Canadian Monitoring, Helmholtz MOSES). Mountain permafrost is addressed by PERMOS in Switzerland. The characterization of the errors and uncertainties is carried out using conventional evaluation of bias, absolute error and Root Mean Square Error RMSE difference. | | ALD measurements provided by users were compared with CCI permafrost fraction. In addition, Permafrost abundance by other means (e.g., geophysical measurements) provided by users will be compared with CCI permafrost fraction. |
| [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 | | |

***Table 2.1 cont.***

| **Recommendations for validation** | **Permafrost Temperature** GCOS ECV | **Active Layer Thickness** GCOS ECV | **Permafrost Fraction** 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 Permafrost_cci GTD as well as the quality of the product uncertainty estimates themselves.<br><br>Mountain permafrost regions are addressed separately as here the uncertainties are significantly higher. | Validation is quantifying the uncertainty (RMSE in cm) and relative percentage error for Permafrost_cci ALT as well as the quality of the product uncertainty estimates themselves. | Validation is quantifying the uncertainty of the permafrost fraction product.<br><br>Mountain permafrost regions are addressed separately as here the uncertainties are significantly higher. |
| [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 thickness in lowland and mountain permafrost areas are available over long time periods and for different permafrost conditions, land cover types, topographies and climates. This allows extending the validation to investigate the temporal stability of the Permafrost_cci products that is a characteristic important for climate research. We will apply two approaches to investigate the temporal stability: i) g-score and ii) bias stability, | | |

# 3 RULES FOR UNBIASED VALIDATION AND VALIDATION CRITERIA

## 3.1 Unbiased validation

The CCI project team shall ensure independence 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-1,2,3]: the validation needs to be carried out by team members not involved in the final algorithm selection [AD-1,2].

The validation in Permafrost_cci is 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 additional national measurement networks such as PERMOS in Switzerland and national monitoring programs in Russia, Canada and United States, as well as data sets from individual PIs [AD-4, RD-1,2,6]. WMO/GCOS GTN-P managed by the International Permafrost Association (IPA) provides in situ measurements for the Permafrost ECVs from the Thermal State of Monitoring (TSP) and the Circumpolar Active Layer Monitoring program (CALM), including community standards for measurements and data collection (Brown et al., 2000, Clow, 2014). Specifically initiated by the International Polar Year (IPY 2007-2008), GTN-P established a temperature reference baseline for permafrost and could considerably increase the number of observations. Using this extended monitoring, the permafrost research community could for example demonstrate that during the IPY reference decade (2007 to 2016/2017) permafrost temperature at depths of the Zero Annual Amplitude (ZAA) increased globally by 0.29 °C (Biskaborn et al., 2019).

In addition to the community permafrost temperature data collection at ZAA (GTN-P, 2018, 2021), there is an obvious need for a standardized ground temperature benchmark dataset across all different depths, specifically also standardizing shallow depth measurements, as has been stressed by user communities, as it does not yet exist [AD-5, RD-7,8]. Profoundly, land surface and climate models lack standardized data on shallow ground temperature for a scientific evaluation of their simulated ground thermal conditions and permafrost states. Land surface and climate models are parameterized down to depths of 3 m or 5 m depths only, not reaching the deeper ZAA depths in continuous permafrost.

To validate the Permafrost_cci products, the team in Permafrost_cci responsible for validation has been thus compiling, checking and standardizing all available communities' ground temperature (GT) and ALT data [RD1,2]. The majority of the in-situ data collection is contributed from GTN-P/IPA and its individual Principal Investigators (PIs) and for the Eurasian Permafrost region from the Russian meteorological monitoring network ROSHYDROMET (RHM) program, in addition with contributions from GTN-P PIs, and datasets from Nordicana D for Canada, and NASA Arctic-Boreal Vulnerability Experiment ABoVE datasets for Alaska (United States) were additionally collected. GTN-P and RHM time series and the data collections from additional networks and PIs provide a large data collection of in situ measured reference datasets.

All these data are no easy-to-use or readily available time-series data that are data-fit for validation and round robin exercises. For example, the data collection of ground-temperature time series is a highly complex and heterogeneous data set including variable timeframes from hourly over annually to sporadic measurements, in different depths and not consistent over time.

In addition, all the available in situ data sets, despite being produced according to community standards and published, contain a large number of caveats, including erroneous or imprecise coordinate locations and non-corrected measurement errors, depending on region and PI [RD-1,11]. Within Permafrost_cci, these pre-existing community in situ data collections have been error-checked, corrected, homogenized, filtered and standardized. The newly compiled, harmonized Permafrost_cci GT depth-time series provides the first consistent reference data set covering all measurement depths for the circum-Arctic: it covers all permafrost zones from continuous to discontinuous, sporadic and isolated of the Northern Hemisphere with all available measurement depths down to 20 m [RD-1,2,11].

The validation and evaluation efforts also consider high-mountain permafrost regions, using in-situ observations of surface and ground temperatures provided by GTN-P PERMOS in Switzerland. In addition, the EO-derived inventories on rock glacier occurence, which was developed by the ESA Data User Element (DUE) GlobPermafrost team since 2016 and which is continued in Permafrost_cci phase I and II, are innovatively used for assessments of the Permafrost_cci products. The PERMOS monitoring data and the rock glacier inventories compiled in 12 regions around the globe in the framework of Permafrost_cci [RD-12] supports the validation in mountain areas, where the Permafrost_cci products contain the highest uncertainties [RD-1,2].

The IPA Permafrost mapping action group contributed in its active action group phase as an important collaborator for validation in Permafrost_cci phase I [RD-6]. Dr. Isabelle Gärtner-Roer, University of Zurich, CH, former vice president of IPA and former leader of the IPA Permafrost mapping action group [RD-5], and Science Officer of the World Glacier Monitoring Service (WGMS), was stating in [RD-6] that a very profound validation is being performed in Permafrost_cci by using the in-situ data from GTN-P and from PERMOS. IPA agrees on the fact that in-situ data are clustered in regions with active permafrost monitoring programs/projects, and that therefore some regions are underrepresented. For the validation in Permafrost_cci, IPA further provided 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 Permafrost ECVs by WMO/GCOS for Permafrost are [AD-2,6]

i) permafrost ground temperature and ii) active layer thickness.

Permafrost_cci added iii) permafrost extent (permafrost fraction), as a mapped permafrost variable, which is 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 main focus of Permafrost_cci lies on the ECV permafrost ground temperature as its derivation also forms the base for the derivation of active layer thickness and permafrost fraction. We performed the assessment of the products by compiling a reference data set in cooperation with the involved research communities and political entities and using common statistical approaches: the characterization of errors and uncertainties is carried out using conventional evaluation measures of mean bias, absolute error difference and Root Mean Square Error (RMSE). The assessment of CRDPv2 released in 2021 is in detail described in [RD-2,9,10].

In Permafrost_cci phase II we will continue

- point-wise site-specific match-up analyses per ground temperature profile per standardized depth and annual time stamps providing averaged bias, mean absolute error and RMSE for the Permafrost_cci GTD time series, with in-depth analyses of performance depending on measurement depths, permafrost landscape type and land cover.
- point-wise site-specific match-up analyses per averaged ALT per grid for annual time stamps providing averaged bias, mean absolute error and RMSE for the Permafrost_cci ALT time series, with in-depth analyses of performance depending on permafrost landscape type and land cover.
- point-wise site-specific match-up analyses per in situ permafrost temperature in the first two meters and per ALT for annual time stamps provide averaged bias, mean absolute error and RMSE for the Permafrost_cci PFR time series, with in-depth analyses of performance depending on permafrost landscape type and land cover.

In Permafrost_cci phase II we plan

- functional validation providing the relative error per pixel by classifying the site-specific match-ups related to CCI-Landcover. Coming from these classes to class-specific relative errors per pixel, this method can provide the relative error per pixel in the Permafrost_cci products on a circum-Arctic scale.

# 4 VALIDATION ACTIVITIES

## 4.1 Reference 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 science communities, specifically with the permafrost community [RD-1,2,6,9,10,11]. Available in-situ data sets and their characteristics and data availability (data access via data portals, program websites and Principal Investigators PIs) are described in [RD-11] and the specific data sets used in the validation rounds are described in [RD-2,9,10]. In the following sub-chapters, we provide more details on successfully compiled data and also on planned updates. For a cross-product assessment, we additionally apply the Freeze-Thaw to Temperature (FT2T) product, an EO-derived GT product, for comparison with the Permafrost_cci GTD product.

***Table 4.1:*** *In situ data of thermal properties and active layer depth measurements available for validation in Permafrost_cci.*

| Region | thermal properties and active layer depth & thickness | Contributor |
|---|---|---|
| Circumpolar Arctic and Antarctica | temperature data (borehole, soil, rock), active layer depths & thickness | GCOS GTN-P database, CALM, individual datasets from PI's |
| North American Permafrost Regions | temperature data (borehole, soil, rock), active layer depths & thickness | GCOS GTN-P database, CALM, NSF Arctic Data Centre (Wang, 2018), NASA ABOVE, NORDICANA D, individual datasets from PI's |
| Siberian Permafrost Regions | temperature data (borehole, soil, rock) active layer depths & thickness | ROSHYDROMET, CALM, GCOS GTN-P database, individual datasets from PI's |
| European high-latitude Permafrost Regions | temperature data (borehole, soil, rock) | GCOS GTN-P database, ROSHYDROMET |

### *4.1.1 ECV permafrost temperature*

GTN-P/IPA and other national permafrost monitoring programs established temperature depth profiles (either measured continuously or sporadically) in deeper boreholes down to several tens of meters (e.g., from exploration, or specifically installed for permafrost monitoring), and in the form of continuously measuring temperature sensors directly installed in the ground (soil, sediment, rock) covering shallow depth profiles down to a few meters. GTN-P also defined the quality guidelines for the instrumentation of the depth profiles (Burgess et al., 2000, Smith et al., 2009, Streletskiy et al., 2017). Specifically, PERMOS developed and optimized high-quality permafrost temperature measurements for mountain permafrost regions (Nötzli et al., 2021), Nötzli et al. (2021) describe in their best practices for permafrost boreholes the concept of the optimal depths of permafrost boreholes and sensors as diurnal temperature fluctuations occur in the uppermost 1–3 meters also at higher elevations and higher latitudes, whereas the ground temperatures at depths below the ZAA reveal the multi-annual trends. Romanovsky et al. (2001) describe that due to no seasonal dynamics at and below ZAA, permafrost temperature at ZAA can be measured only yearly in deep permafrost boreholes. The depths in-between the shallow ground and ZAA are still subject to seasonal temperature variations and can reflect effects of extreme weather periods, i.e., in case of yearly measurements only the ZAA depth measurement is reliable for estimating long-year permafrost temperature dynamics. In many cases, boreholes are equipped by permanently installed multi-sensor cables in depth profiles with automated continuous data logging. Nötzli et al. (2021) describe that typically in boreholes, a spacing of 0.2–1.0 m is applied in the uppermost 5–10 meters. Below 10 m depth, the spacing increases to 2–5 m down to ca. 30 meters. At larger depths, temperature is measured in much longer depth intervals, e.g., of 10 meters. A dense sensor installation in 0.5 to 1 m depth in equipped boreholes with large diameters is not considered scientifically appropriate because of the influence of the artificial or air infillings of the boreholes near the surface. Ideally, an additional shallow borehole or a shallow soil temperature profile is installed in close range (Harris et al., 2001, Isaksen et al., 2001).

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. Tests have shown the comparability of different measurement techniques to have an overall accuracy of ±0.1 °C. About 10 to 20 % of the GTN-P 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 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 need to be excluded by PIs from the final data sets. The discussed accuracy relates to absolute temperature values per sensor measured, but sensor nonlinearities are generally small. The community therefore considers <0.1° K a conservative averaged estimate of the accuracy of temperature on an individual sensor basis.

These data sets on permafrost temperature are managed and made publicly available via large-scale international and regional programs where several of the Permafrost_cci team members are in close cooperation with. The core data of the Permafrost_cci reference data collection come from the GTN-P program. One available source is the 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-11]. The GTN-P data collection that we use in the Permafrost_cci validation were downloaded from the database in late 2018 and with updates up to late 2022. After investigation of the GCOS GTN-P temperature data collection, we found that only ~40-50 % are usable for validation in Permafrost_cci.

The ground temperature time series frequently contained large data gaps due to the challenging logistics in the Arctic, but also, regrettably, several of the extracted temperature data sets from the GCOS/GTN-P database contained large data errors and further artifacts. We could retrieve published standardized ground temperature data sets from GTN-P PIs from the PANGAEA repository: GTN-P (2018), now updated to GTN-P (2021), Boike et al. (2018), Bergstedt & Bartsch (2020a). In addition, we received several individual ground temperature time series, not yet available in the GCOS GTN-P database, from PI's directly. We will receive further updated ground temperature data records directly from PIs of GTN-P via B.GEOS (EU Arctic Passion).

Amongst the 35 GTN-P mountain permafrost boreholes, 27 belong to the Swiss permafrost monitoring network PERMOS [RD-11]. 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 4.2) and are spatially distributed over the different geographical region within the Swiss Alps. The longest record totalizes more than 30 years of observation, whereas the majority of the PERMOS boreholes has between 10 and 23 years of observation. The depth of the boreholes spans from 14 m to 100 m depth.

***Table 4.2:*** *Borehole ground temperature monitoring in the Swiss National GTN-P PERMOS monitoring*

| PERMOS code | GTN-P code | start | lat [dd] | long [dd] | height m a.sl. | borehole depth m | morphology | 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 |
| 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 |
| 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 |
| 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 |
| MUR_0299 | CH 34 | 1999 | 46.50723 | 9.927338 | 2539 | 64 | rock glacier | coarse blocks | ca. 18 m |

| PERMOS code | GTN-P code | start | lat [dd] | long [dd] | height m a.sl. | borehole depth m | morphology | surface type | permafrost thickness |
|---|---|---|---|---|---|---|---|---|---|
| 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 |

For Canadian datasets, Nordicana D is the data repository of the Canadian centre d'études nordiques (CEN) curating long-term time series of permafrost borehole temperatures, and also shallow ground and air temperature in high temporal resolution that are not in the dynamic GCOS GTN-P database. Status for Permafrost_cci Phase II: in Nordicana D, we found 31 additional or updated datasets, which we will check for integration into our validation dataset. Additionally, we found Canadian data on the National Snow and Ice Data Center repository (NSIDC), which will be checked for possibility of integration for validation (https://nsidc.org/data/ggd503/versions/1).

For validation of CRDPv2 in Permafrost_cci phase I, we also used data from the NASA Arctic-Boreal Vulnerability Experiment ABoVE (https://above.nasa.gov/). ABoVE is a NASA Terrestrial Ecology Program field campaign that is conducted in Alaska and Western Canada. Three soil temperature compilations (Nicolsky et al., 2019, 2020, Rey et al., 2020) from 37 sites in Alaska have been collected from https://above.nasa.gov/profiles_/above_projects.html and further processed and standardized. Most of the data sets do not comprise a full year (and were not updated yet in the ABoVE data repository as of May 2023). In Permafrost_cci phase II we will contact PIs for updated data.

ROSHYDROMET (RHM) is the national Russian meteorological monitoring network that provided publicly available long-term ground temperature records close to meteorological stations. RHM data are known to have artifacts like wrong temperature data that were manually put in, and problems concerning geo-location, having only two decimal digits. Within Permafrost_cci phase I we corrected coordinates and artifacts and compiled all robust RHM ground temperature records. RHM data distribution has stopped and there are no further updates relevant for Permafrost_cci phase II.

*Permafrost_cci standardized and synthesized MAGT time series*

For validation in Permafrost_cci phase I we compiled a Northern Hemisphere standardized mean annual ground temperature (MAGT) data collection spanning from 1997 to 2019 encompassing shallow to deep going permafrost. In this ground temperature time series data collection, we specifically optimized the shallow ground temperature records.

The Permafrost_cci assembled MAGT per depth data collection comprehensively covers the wide range of different Northern Hemisphere permafrost landscapes and landcover providing ground temperature records from surface down to 20 m depths if available. This Permafrost_cci MAGT data collection will be updated to 2021 encompassing all available new sites, e.g. from new PIs and measurement programs.

**Permafrost_cci phase I** data compilations from:

**GTN-P** most datasets are from **GCOS/ GTN-P (https://gtnp.arcticportal.org/)** [dynamic GTN-P data base of the global GTN-P monitoring program]

**plus, individual GTN-P PI data collections in form of DOI-referenced data publication**
**PANGAEA** [National World Data Repository for environmental data, Germany]
https://doi.org/10.1594/PANGAEA.882061, Boike et al., 2017
https://doi.pangaea.de/10.1594/PANGAEA.905233, Boike et al., 2019
https://doi.pangaea.de/10.1594/PANGAEA.884711, GTN-P, 2018
https://doi.pangaea.de/10.1594/PANGAEA.912482, Bergstedt & Bartsch, 2020a

**Arctic Data Center** [National World Data Repository for Polar research, United States]
https://arcticdata.io/catalog/ #view/doi:10.18739/A2KG55; Wang, 2018 [includes also additional USGS data]

**plus, individual GTN-P PI data in form of direct contributions to Permafrost_cci**
Romanovsky & Kholodov University of Alaska Fairbanks (UAF), United States
Ulrich (University of Leipzig, Germany in cooperation with AWI)

**Updates planned in Permafrost_cci phase II:**

**GCOS/ GTN-P** will be regularly checked for updates in communication with the GTN-P director A. Irrgang, AWI)
**plus, individual GTN-P PI data in form of direct contributions to AWI and B.GEOS**

**Roshydromet RHM** national meteorological monitoring program
**Permafrost_cci phase I: http**://meteo.ru/data/164-soil-temperature [national monitoring program, Russia]
**Permafrost_cci phase II:** no updates available

**Nordicana-D** [National World Data Repository for Polar research, Canada]
Permafrost_cci phase I:
http://www.cen.ulaval.ca/nordicanad/dpage.aspx?doi=45291SL34F28A9491014AFD; Allard et al., 2016, CEN, 2013
**Permafrost_cci phase II: Nordicana-D** will be regularly checked for updates in communication with Nordicana-D data repository curators and directors

**NASA ABoVE** [National Terrestrial Ecology Program field campaign program for N-America (including United States and Canada]
**Permafrost_cci phase I:** https://above.nasa.gov/profiles_/above_projects.html
**Permafrost_cci phase II**: NASA ABoVE will be regularly checked for updates in communication with PIs

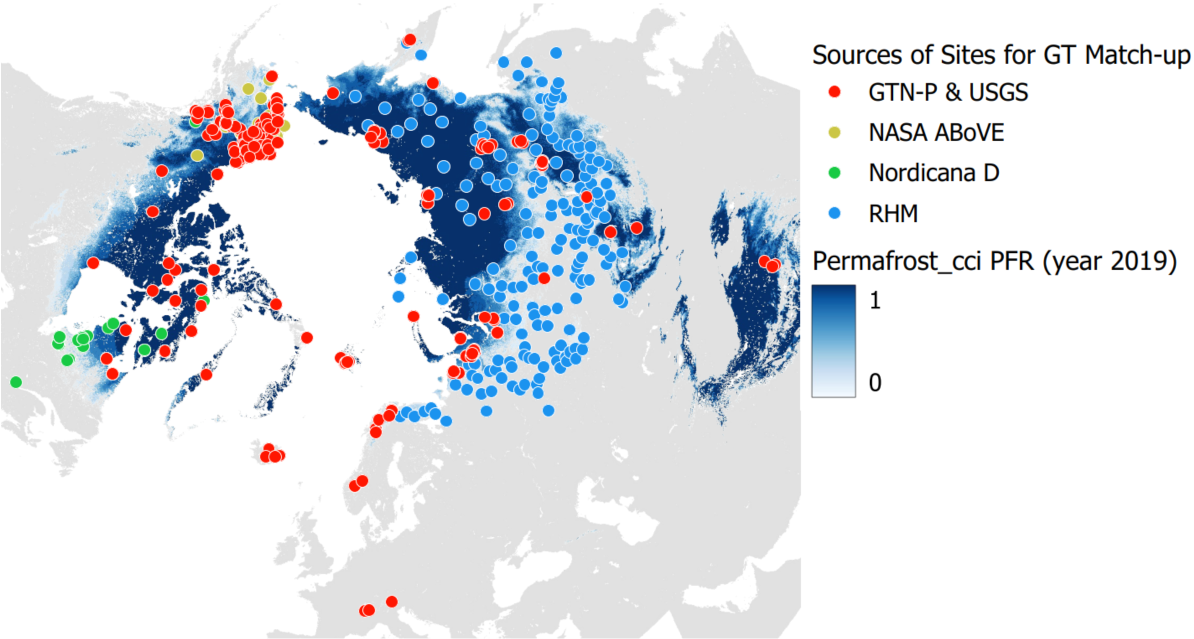


***Figure 4.1:*** *Northern hemisphere Permafrost_cci Permafrost FRaction PFR and in situ ground temperature (GT) sites grouped by data source.*

***Table 4.3:*** *Origin and number of datasets which can be used for validation with status of April 2023. C = Continuous Permafrost, D = Discontinuous Permafrost, S = Sporadic Permafrost, I = Isolated Permafrost; classification according to NSIDC map of Brown et al., 2002.*

| | No. | C | D | S | I | NA |
|---|---|---|---|---|---|---|
| GTN-P + USGS | 338 | 164 | 98 | 33 | 4 | 39 |
| NASA ABoVE | 37 | 25 | 12 | 0 | 0 | 0 |
| Nordicana D | 40 | 10 | 2 | 10 | 12 | 6 |
| RHM | 191 | 37 | 23 | 33 | 50 | 48 |

Several original site coordinates from various sources have been identified to be erroneously located in rivers, lakes or the ocean and not on the land surface. Checking each site manually and if necessary correcting coordinates was necessary. This inaccuracy is for a large part due to PI entries of decimal degree coordinates with only two decimal digits and also due to erroneously switched numbers. We contacted PIs and NSIDC about the correct geolocation and corrected geolocations with their support or based on our best guesses guided by high resolution Google Earth background images and data description and reports or related publications. Metadata on these corrections, the reason and method, the original coordinates and the best estimate are all part of the reference data collection.

Please note that we excluded all sites that are not representative of the landscape-scale of in-situ measurements from the match-up (RD-2) but keep them in the reference data collection with indicative metadata information. Excluded sites in the validation process are mountain sites that are specifically assessed by PERMOS, small-scale landscape anomalies such as very local peatland patches or in-situ temperature measurements in pingos (ice hills) that are considered landscape anomalies in respect to a 1 km grid cell resolution that is represented by the Permafrost_cci.

One focus of GCOS/GTN-P is to manage and provide temperature measurements at ZAA, frequently at depths of 10 m or deeper. In addition, about 35 % of the GCOS/GTN-P data collection contains temperature data that are in lower than 5 m depths. In general, the WMO suggests the depths of 0.05, 0.1, 0.2, 0.5, and 1 m as standard depths for soil measurements (WMO, 2021). Similarly, the assembled and interpolated ground temperature data collection from Wang (2018) integrating GTN-P sites from the University of Alaska Fairbanks (UAF) PIs (Romanovsky and his Laboratory) and USGS-managed sites provides interpolated temperature data at the four depths at 0.25, 0.5, 0.75 and 1 m for Alaskan ground temperature soil profiles (Wang et al., 2018). In contrast, RHM temperature measurements are set at up to seven standardized depths at 0.2, 0.4, 0.8, 1.2, 1.6, 2.4, and down to a depth of maximum 3.2 m.

In summary, the assembled Permafrost_cci Ground Temperature (GT) data collections from GCOS GTN-P, RHM, Nordicana D, NASA ABoVE and individual PI's cover various datasets measured at different depths, however with several depths available at a high frequency (e.g., Figure 4.2) and with a wide range of time steps, from hourly to yearly measurements. We compiled all available depths from all data sources and selected the most representative measurements that are set 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, 15.0 and 20.0 m depths (in detail described in the documents of RD-1,2,11 and visualized in Figure 4.2). As not all temperature sensors are placed at exact depths, we also included sensor measurements, with ±3 cm for depths <5 m, ±5 cm for depths from 5 to <10 m, ±10 cm for depths at 10 m and ±20 cm for depths from 10 to 20 m. Data from sensors <2 m depth were discarded if they represented wide-diameter boreholes (in general filled with air or artificial filling), except if confirmed by PI as reliable.

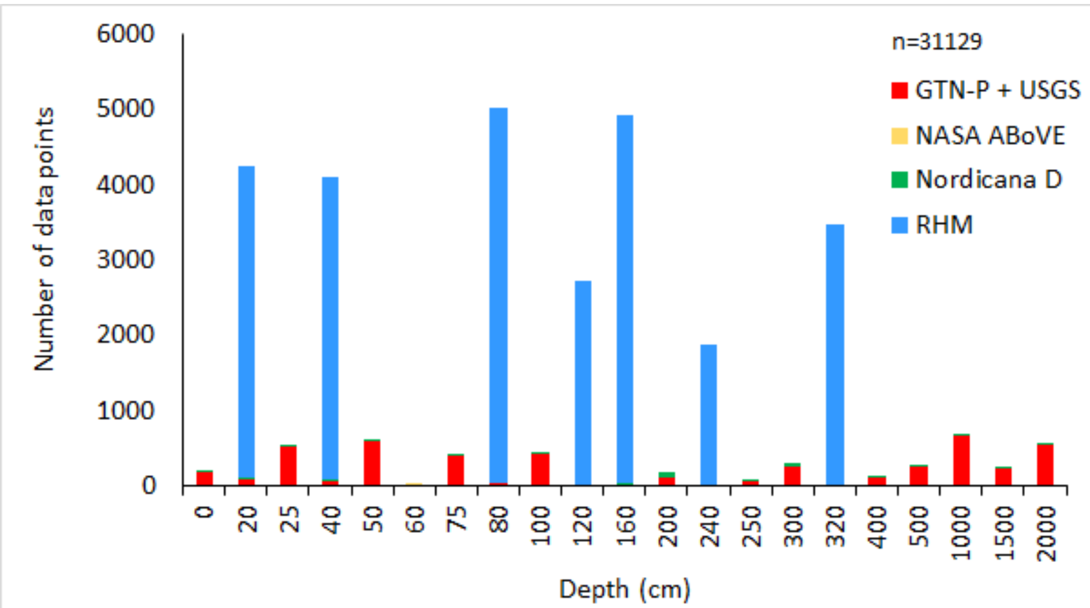


***Figure 4.2:*** *Number of Ground Temperature (GT) measurements at a given depth (in centimeter, y-axis) which can be used for validation in Permafrost_cci. Values include measurements over all years from 1980-2022 with n being the overall amount of available GT data per depth.*

In addition, as especially RHM does not provide GT measurements at 1 or 2 m depth, we interpolated temperature values accordingly to Permafrost_cci focus depths. We only undertook interpolation if the depth profiles contained a minimum of three sensors down to 1.20 m. Interpolation was conducted by linear regression between two single measurement depths, resulting in separate equations for each sensor-pair and year. Like this, Permafrost_cci MAGT time series per depth are also enriched with temperature data interpolated for shallow and deep temperature profiles down to the maximum sensor depth. These additions are coded so that the MAGT dataset can be used with discrete measurements or in addition with the enrichment of the interpolated depth data.

The GT data from all time steps are processed to one value per year: the standardized mean annual Ground Temperature per Depth (GTD). Yearly means were not calculated if >20% of yearly values are not available or if more than one month of data is missing. An exception is made for data at the depths of ZAA that are fulfilled by a single measurement in the year due to no seasonal temperature variations. We also attributed this in the form of metadata information, which allows assessing the quality of each temperature value (Table 4.4). These metadata comprise for yearly values the ratio of missing data per month/year (missing days per year/365) and the amount of completely missing months.

***Table 4.4:*** *Example of how the compiled data set provides metadata information of MAGT across depths. Mxx = ratio of missing values per month/year at depth xx m. mMxx = number of 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 |

The Permafrost_cci reference data consists of MAGT from 1980 to 2019 with product depths 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, 15.0, 20.0 m. The depth stratification provides high depth resolution in shallower depths and is larger scaled in deeper depths. The concept of ZAA and Permafrost_cci depth stratification is visualized in Figure 4.3 and 4.4.

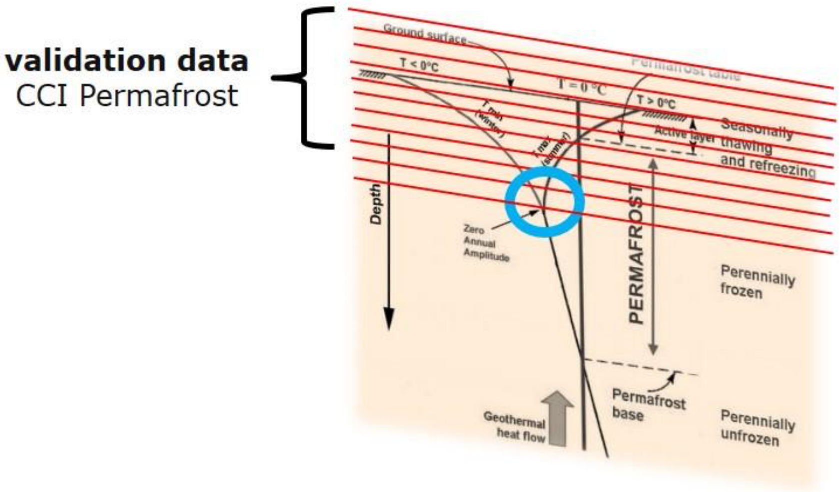


***Figure 4.3:*** *Concept of Zero Annual Amplitude (ZAA) depth (blue circle) and Permafrost_cci depth stratification.*

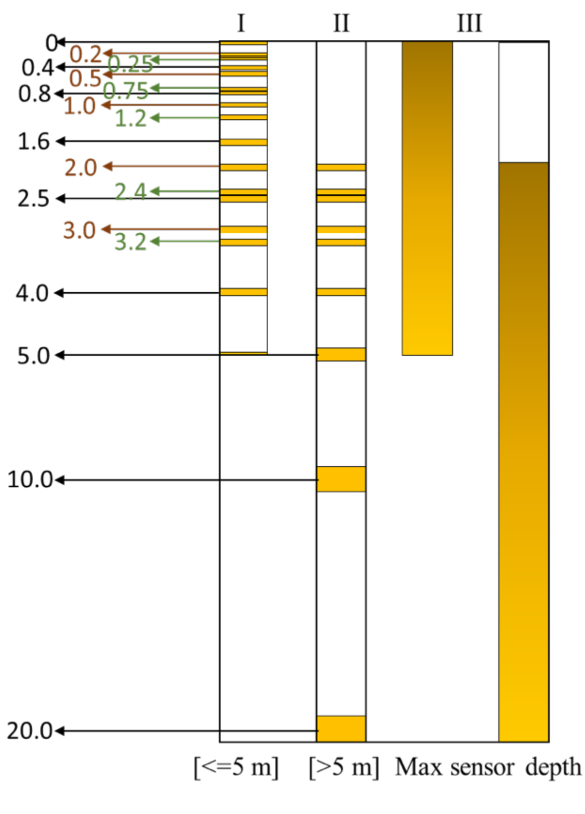


***Figure 4.4:*** *Discrete depths are provided small-scaled in shallower depths and larger scaled in deeper depths. In deep boreholes (Group II, >5 m depth), <1.5-2 m measurements were discarded due to their inaccuracy due to large borehole diameter and air or artificial infillings. In Group III, temperature data are 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 are discarded.*

*4.1.2 Satellite derived Freeze/Thaw Surface Status GT*

The Freeze-Thaw to Temperature (FT2T) model is an empirical model, based on a linear regression analysis between the annual sum of frozen days, derived from microwave EO sensors, and in situ GT measurements (Kroisleitner et al., 2018). It was initially developed from GTN-P data, spanning the years 2007-2013 available from Paulik et al. (2014). The method by Naeimi et al. (2012), which forms the basis for the 2007-2013 record of Paulik et al. (2014), has been applied to further records, extending the dataset to 2018. A Metop ASCAT global gridded data set available from EUMETSAT (SOMO12) has been used for this purpose. The method and set parameters were evaluated by in situ records and C-band SAR data (Sentinel-1; Bergstedt et al., 2020b). FT2T has been further developed for Permafrost_cci to represent the depths of the CRDPv2 and calendar years. With respect to in situ data availability for the model calibration, only in situ data in 1 m depth were considered. Further improvements have been made regarding bias correction for lake fraction using Sentinel-1 (Bergstedt et al., 2020c). These apply to lake rich permafrost regions. Records have been extracted for selected site comparisons and for regions in addition to the circumpolar comparison presented in [RD-6].

*4.1.3 ECV Active Layer Thickness*

The permafrost ECV Active Layer Thickness (ALT) is measured at the end of the thaw season in late summer on grids of 10, 100 or 1000 m with evenly spaced nodes at 1, 10 or 100 m. ALT can vary substantially on an inter-annual basis. In general, ALT is greater in years with warmer summers and shallower in those with cooler summer temperatures (Brown et al., 2000) and can by this serve as ECV indicative for permafrost development. Nelson and Hinkel (2003, in “Methods for measuring active-layer thickness. In: A Handbook on Periglacial Field Methods”) describe in detail how the term of thaw depth or active layer depth (ALD) is distinct from the term of the ECV in the form of ALT. For an estimation of the ECV ALT it is relevant to measure the thaw depth in the grid at the end of the thawing season in late summer (https://www2.gwu.edu/~calm/data/north.html). This is distinct to measurements of ALD, at any time during its development in summer that is also a standard procedure in permafrost research. ALD is an instantaneous value that is always less than or equal to the thickness of the fully developed active layer ALT at the end of the thaw season. A typical probe to measure ALD 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 are apparent when ice-rich frozen ground is encountered. At sites where thaw depth is very large (e.g., deeper than 1 m, such as in deeply thawed or stony soils), it is very difficult, however, to extract a probe. Optimally, ALD executors should have experience with this measurement and body strength. If executors are experienced, the accuracy of ALD is estimated to be around 2 cm, of an ALT grid measurement, estimations give around 5 cm accuracy (Fagan and Nelson, 2017).

A comprehensive collection of ALT time series from grids (various sizes) is available from the Circumpolar Active Layer Monitoring Network (CALM), Brown et al., 2000, Fagan and Nelson, 2017). The data are available for download on https://www2.gwu.edu/~calm/ as the ALT in a specific year expressed as maximum thaw depth measured in late summer at the CALM long-term monitoring grids. The single ALT 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.

About half of the CALM datasets consist of up to ten ALT grid measurements made in late summer (though not necessarily in ten consecutive years), ~20 % of the data provide more than 20 years of observations on ALT. We compiled a standardized, error-corrected (e.g., for incorrect coordinates, value typos, etc.) ALT data set with annual resolution with a circum-Arctic geographic coverage (Table 4.5., Figure 4.5). Please note that we excluded in Permafrost_cci phase I all sites in Mongolia, Central Asia, and on the Tibetan Plateau when we undertook validation as these consisted of very deep ALTs of several meters measured with temperature sensors with low depth resolution and come with a high inaccuracy.

***Table 4.5:*** *Regional overview of CALM active layer measurements, showing number of sites available in the dataset downloaded in 2018 for Permafrost_cci phase 1, plus the number of new sites in the CALM dataset downloaded 2023 and in parentheses the number of sites with new data after 2018 for old and new sites respectively.*

| | CALM 2018 | new in CALM 2023 |
|---|---|---|
| Canada | 31 (2) | 4 (0) |
| China | 11 (0) | 0 (0) |
| Denmark (Greenland) | 3 (2) | 0 (0) |
| Kazakstan | 3 (0) | 0 (0) |
| Mongolia | 47 (0) | 0 (0) |
| Russia | 66 (46) | 9 (9) |
| Svalbard (Italy) | 0 (0) | 1 (1) |
| Svalbard (Norway) | 3 (3) | 1 (1) |
| Svalbard (Poland) | 4 (0) | 0 (0) |
| Svalbard (Sweden) | 2 (1) | 0 (0) |
| Switzerland | 2 (2) | 10 (6) |
| United States (Alaska) | 67 (41) | 1 (1) |
| **SUM** | **239 (97)** | **26 (18)** |

For some measurements in the GTN-P CALM data collection, metadata provided information that this value represents ALD measured earlier than ALT in a specific year. In few cases, if the local processes were well known, the PIs provided an interpolation towards ALT. In addition, individual PIs were contributing a considerable data collection on ALD measurements during summer (measured and collected by B.GEOS) that we used in Permafrost_cci phase I for binary validation of permafrost extent (i.e., if there is an ALD measurement, it is in permafrost).

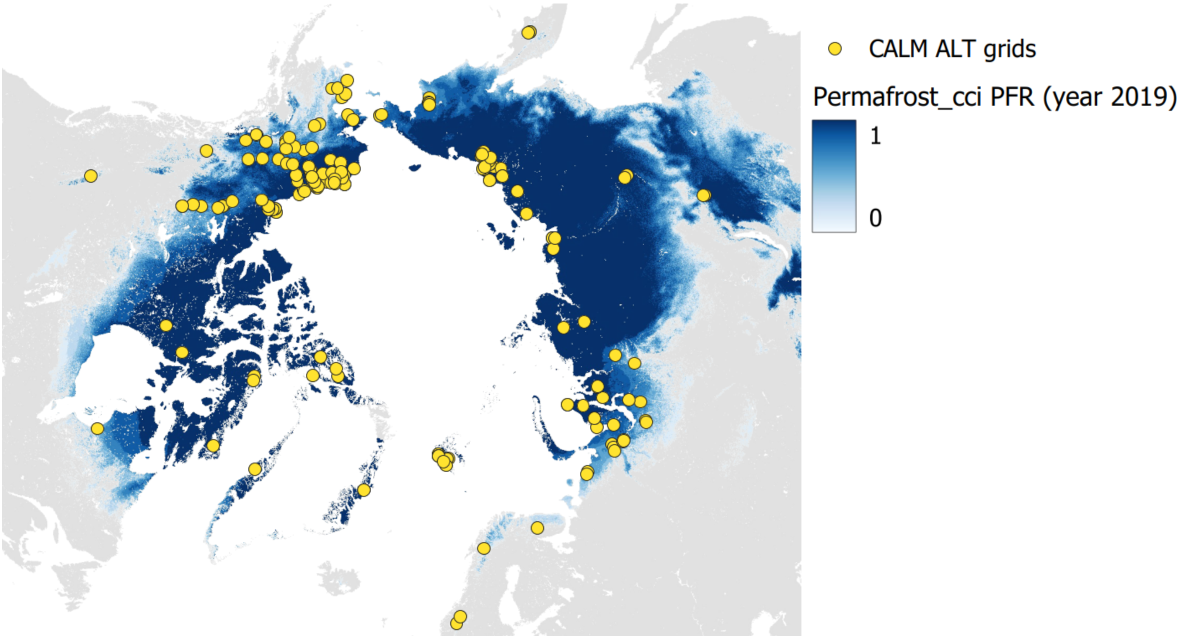


***Figure 4.5:*** *Northern hemisphere Permafrost_cci Permafrost FRaction PFR and in situ GTN-P CALM grid sites for ALT measurements (without sites with deep ALT in rocky terrain, China, Mongolia, Kazakhstan and in mountain permafrost, e.g., Switzerland).*

**Permafrost_cci phase I** data compilations from

**GTN-P / CALM** https://www2.gwu.edu/~calm/ [global GTN-P monitoring program]

**plus individual GTN-P PI data in form of direct contributions to B.GEOS**

**Updates planned in Permafrost_cci phase II:**

**GTN-P / CALM** will be regularly checked for updates in exchange with CALM PIs D. Streletskiy and N. Shiklomanov, Columbian College of Arts and Science, United States

**plus individual GTN-P PI data in form of direct contributions to AWI and B.GEOS**

### 4.2 Contextual data relevant for permafrost-related applications

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 defined by permafrost overlain by the seasonally dynamic active layer. The composition of soil/rock, ice, air, unfrozen water, organic content, it's cryotexture and -structure define the thermal sub-ground properties of the two layers permafrost and active layer. Vincent et al. (2017) added a 3rd layer, the buffer layer, consisting of the above-ground vegetation (in the form of polar desert soil crusts to tundra grasses, forbs, and lichens to shrubs and trees farther to the south) (as example Figure 4.6). In engineered environments, the buffer layer includes the infrastructure. In both cases, natural and engineered environments, the surface buffer layer strongly affects the transfer of heat between the atmosphere and the active layer. This effect is compounded by 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 landscapes vary greatly in horizontal space, and the properties of each of the three layers and their interfaces can change over short length scales. Units can be interspersed with non-permafrost units including lakes, bogs, rivers, and in discontinuous permafrost with unfrozen ground units.

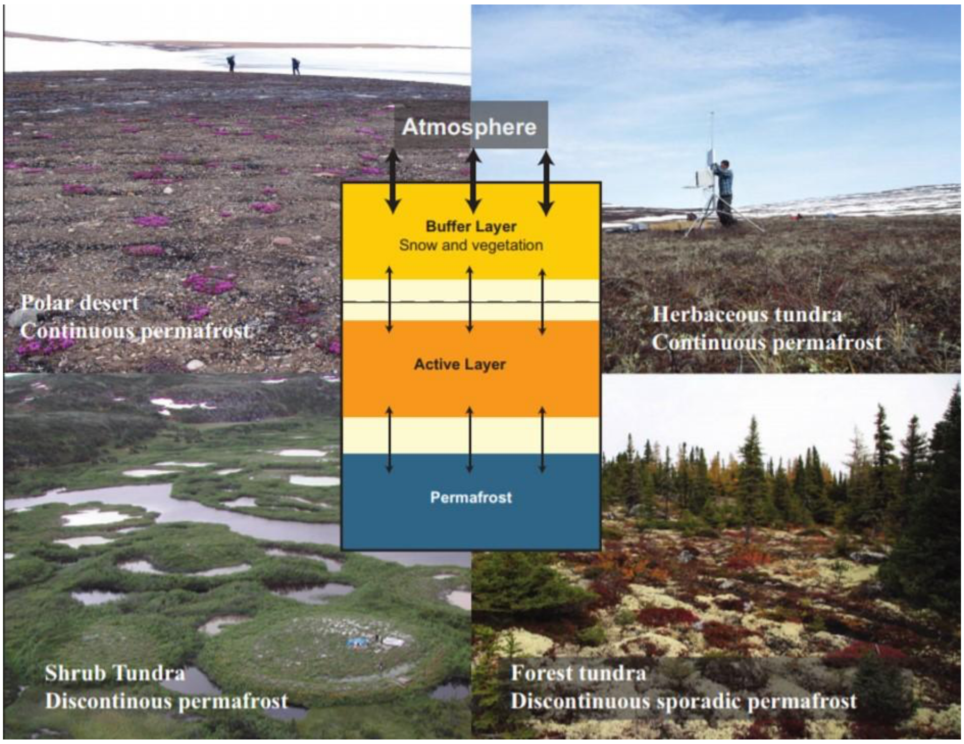


***Figure 4.6:*** *(Fig. 4 in Vincent et al., 2017): The three-layer model applied to natural landscapes. As illustrated here, the buffer layer varies greatly in its geometry (thickness), albedo, and other physical properties, both within and between landscapes, and as a function of vegetation type and 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 Phase II, we plan to assemble in addition to the data collection on GTD and ALT a data collection on structural and environmental data on the 2-layer system permafrost and active layer (stratigraphy, organic layer (abundance, thickness), ground ice content, dominating lithology and texture) and on the buffer layer properties (vegetation composition, height of vegetation, infrastructure, surface habitus (boulders, gravel, …)) as it is visualized in Figure 4.7.

We will use all published information available, the detailed GTN-P CALM metadata and contextual data 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 4.6-4.9).

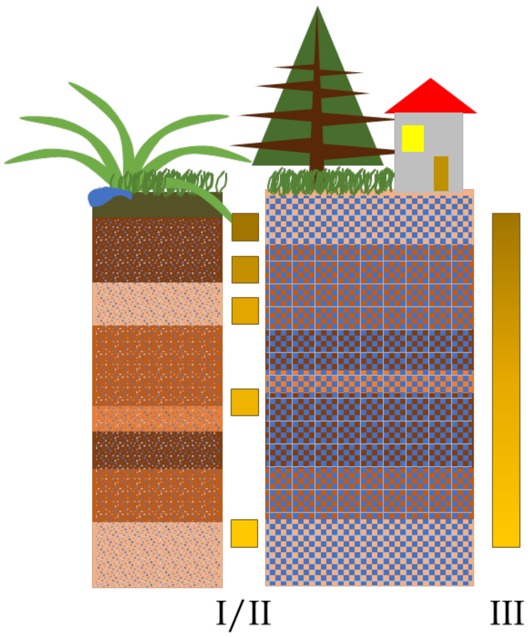


***Figure 4.7:*** *Permafrost_cci metadata collection planned on stratigraphy, organic layer, ground ice content, dominating lithology and texture and on the buffer layer (vegetation composition, infrastructure,…) will be provided at the best quality available and including a quality index (cf. Tables 4.6 – 4.9).*

***Table 4.6:*** *table on '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 from 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 4.7:*** *A table on 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 (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 4.8**: A table on 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 4.9:** A table on 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]<br>'40 %' sand [ 0 – 1 m]<br>… [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*

The Permafrost_cci match-up process is comparable to a validation using Fiducial Reference Measurements (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 in Permafrost_cci is carried out pairwise time- and depth-specific. For the direct comparison between the match-ups at the individual sites we still need to carefully consider the different spatial scales between the local representativeness of the reference measurement and the km-scale of the EO-derived Permafrost_cci product.

**MAGT:** In Permafrost_cci phase 1, we conducted point-wise site-specific match-up analyses for MAGT per standardized depth and per year providing average bias, absolute error, standard deviation, RMSE and several more statistical metrics [RD-2]. In Permafrost_cci phase 2, we plan to undertake the point-wise specific match-up analyses with the extended data sets.

Permafrost_cci GTD CRDPv2 match-up evaluation between simulated Permafrost_cci and in situ MAGT time series per site and specific depth showed the following characteristics: overall, Permafrost_cci GTD had a median MAGT bias of -1.12 °C (95% CI: -4.19 to 2.51 °C). Match-up pairs refined to in situ MAGT < 1°C and thus from permafrost sites showed a better performance and a median bias of 0.2°C (95% CI: -4.09 to 3.18 °C). Therefore, considering the map product Permafrost_cci GTD < 1°C covering the permafrost regions showed good performance across the Northern hemisphere, with a median bias of 0.2°C for all depths. Geographically, the extreme residuals showed no obvious spatial clusters. However, a relatively large proportion of residuals >95% quantile were located across Alaska, specifically in the boreal region. Residuals across depths varied around zero, without notable differences in the median bias [RD-2].

By specifically providing MAGT time series from shallow depths for match-up datasets i) we gain the widest value range between minimum and maximum ground temperature for validation experiments if we also use winter and summer temperature separately in addition to MAGT (that towards deeper depths gets reduced until zero variability at the depth of zero annual amplitude ZAA), ii) we enable Round Robin (RR) experiments and validation for climate and land surface models that do not contain adequate parameterization of deeper depths iii) we enable RR experiments and assessments for EO microwave-derived products that contains signal information on the shallow subsurface.

Regional comparisons with FT2T have been made in Permafrost_cci phase 1 for CRDPv2. The calibration of FT2T has been revised and extended to include 1 m depth borehole data (North America) and 80 cm depth data (Russian Arctic) in order to avoid a regional (and temperature range) bias. The comparison of Permafrost_cci MAGT at 0 and 2 m depth with FT2T derived ground temperatures for selected locations demonstrated the expected higher variability of surface state from year to year, but also agreement of the different data sources regarding temperature level. The strongest deviations were located in the transition zone (temperatures around 0° C) in Alaska as well as Russia [RD-2].

PERMOS investigations in the Swiss Alps showed in contrast a warm model bias of Permafrost_cci MAGT ranging from +1.22 °C at the surface to +1.81 °C at 10 m depth with the vast majority of inventoried ESA GlobPermafrost slope movement products located outside of the simulated Permafrost_cci permafrost extent area (Permafrost_cci PFR).

We thus consider Permafrost_cci CRDPv2 GTD for the Northern hemisphere to be reliable in the permafrost temperature range with GTD < 1°C.

**ALT:** In Permafrost_cci phase 1, we conducted point-wise site-specific match-up analyses per CALM ALT measurement grid value providing average bias, absolute error, standard deviation, RMSE and several more statistical metrics [RD-2]. In Permafrost_cci phase 2, we plan to undertake the point-wise specific match-up analyses with the extended data sets.

Because CRDPv2 Permafrost_cci GTD >1° C of the sporadic and non-permafrost zones is characterized by a cold median MAGT bias of -1.47 °C, ALT in turn turned out too shallow around the lower 60 ° Latitudes. We thus consider Permafrost_cci CRDPv2 ALT for the Northern hemisphere to be reliable in the permafrost temperature range with GTD < 1°C.

**Permafrost Probability**: In Permafrost_cci phase I, we conducted a binary point-wise and time-wise match-up assessment of the Permafrost_cci Permafrost Extent (Permafrost FRaction PFR) using in situ MAGT and ALT and ALD. We allowed a small variability around Permafrost_cci MAGT 0 °C not setting "permafrost" strictly as in situ MAGT <0 °C in two consecutive years. We compared Permafrost_cci MAGT CRDPv2 to in situ MAGT at all depths down to 240 cm, analyzing the amount of simulated and measured temperatures being both ≤0.5 °C ("permafrost") or both >0.5 ("no permafrost"). Additionally, we included ALT and ALD measurements into our PFR analyses, using all sites/years with an active layer thickness and active layer depth measurements as PFR=100%. In Permafrost_cci phase 2, we plan to undertake the point-wise specific match-up analyses with the extended data sets.

Because CRDPv2 Permafrost_cci GTD >1° C of the sporadic and non-permafrost zones is characterized by a cold median MAGT bias of -1.47 °C, PFR in turn turned out as an overestimation of the areal extent of permafrost PFR <50 % at the southern boundaries of Permafrost in discontinuous, and sporadic permafrost regions along the southern boundary of permafrost in Eurasia. We thus consider Permafrost_cci CRDPv2 PFR products for the Northern hemisphere to be reliable in the permafrost temperature range with GTD < 1°C and in PFR >50% as well as PFR <14% is reliable as non-permafrost.

### *4.3.2 Validation of Temporal Stability*

According to the validation recommendations (Table 2.1), the long-term stability of the Permafrost_cci time series of delivered epochs shall be assessed [TR-30]. In general, GT in shallow depths are frequently characterized by a wide spread between minimum and maximum annual GT and high inter-annual variability (Figure 4.8, example from PERMOS). Also, ALT time series show high inter-annual variability depending on the annual air temperature and precipitation regimes.

In Permafrost_cci phase I, we assessed the Permafrost_cci CRDPv2 with the g-score approach to assess the stability of the Permafrost_cci product time series throughout time. We derived in how many cases Permafrost_cci MAGT and Permafrost_cci ALT followed the same year-to-year trend as the Permafrost_cci reference data set of in-situ measurements. If within both - the Permafrost_cci product time series and the in-situ measurement time series - the slope decreases/increases simultaneously in the same direction (positive or negative), the value of 1 is given for this year match-up. If the slopes develop in different directions for a year, the value 0 is given, and if one slope changes direction while the other slope is constant in a year, the value of 0.5 is given.

The mean value of these year-to-year trend-values represents the fraction of synchronized curve development. This g-score approach gives an assessment on how well the Permafrost_cci variable follows the in-situ measured MAGT and ALT trends, respectively.

The trends in CRDPv2 Permafrost_cci GTD over years generally matched well between the in-situ measurements and the Permafrost_cci product (75%). In some cases (24%) the trends changed from a significant trend (p>0.05) to no trend or vice versa. However, the range of differences in slopes within this group remained low (0.001, 95% CI: -0.22 to 0.29). In mountain permafrost areas, although the absolute values were significantly different, both, the measured and the simulated MAGT, show a warming trend over the period 1997-2019. Differences in trends over time for CRDPv2 Permafrost_cci ALT are larger compared to the GTD product (note ALT match-up sample size is also considerably smaller). Still, the majority (58%) of Permafrost_cci ALT trends over time matched the in-situ trends. Except for one match-up pair, the remaining trends (40%) differed and switched between no trend to significant trend or vice versa (95% interval of slope differences; -8.57 to 2.57).

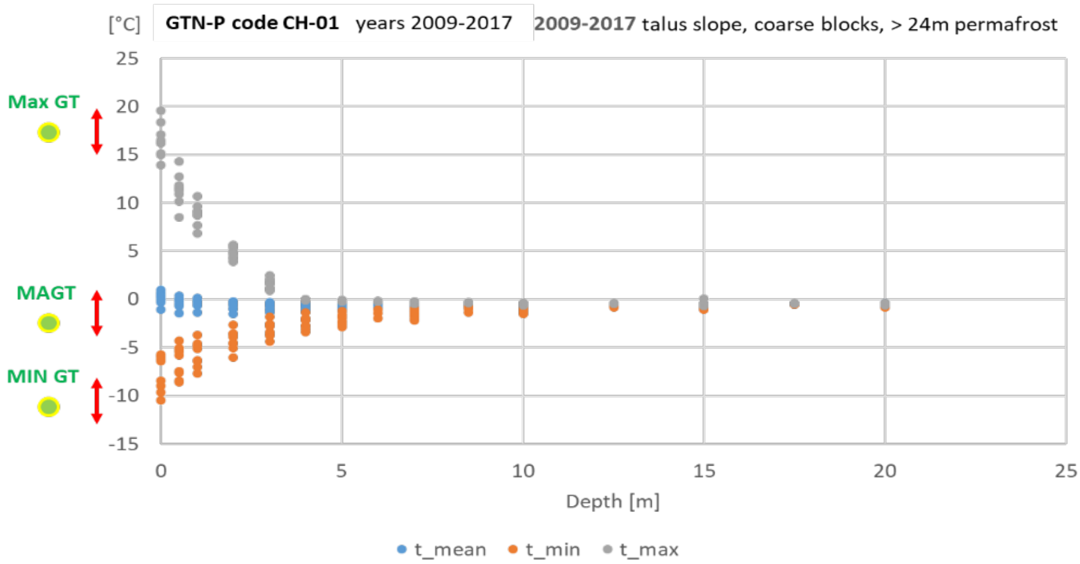


***Figure 4.8***: *High inter-annual variability of MAGT, min annual GT and max annual GT in mountain permafrost*

In Permafrost_cci phase II, we plan to undertake the g-score approach analyses with the extended data sets. In Permafrost_cci phase II we also will undertake the bias stability approach to investigate the magnitude of the interannual variability of the bias. The assumption is that the bias should not change in magnitude from one year to the next. We thus will calculate temporal stability by

$$ts = \frac{bias_j - bias_i}{year_j - year_i}$$

with i being the current year/bias and j being the previous year/bias. The difference is only calculated on a year-to-year basis and rejected, for every missing year at a specific site/depth.

### 4.4 Validation experiments for 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 inventories on rock glacier abundance, extent, and creep, which was developed by the ESA DUE GlobPermafrost program since 2016 and continues in Permafrost_cci phase 1 and phase II are used to validate the binary permafrost extent product in the 12 mountain regions of ESA Permafrost_cci CCN4 (RD-12). The GTN-P PERMOS monitoring data and the EO derived rock glacier inventories support 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 4.2) and are spatially distributed over the different geographical region within the Swiss Alps (Figure 4.9).

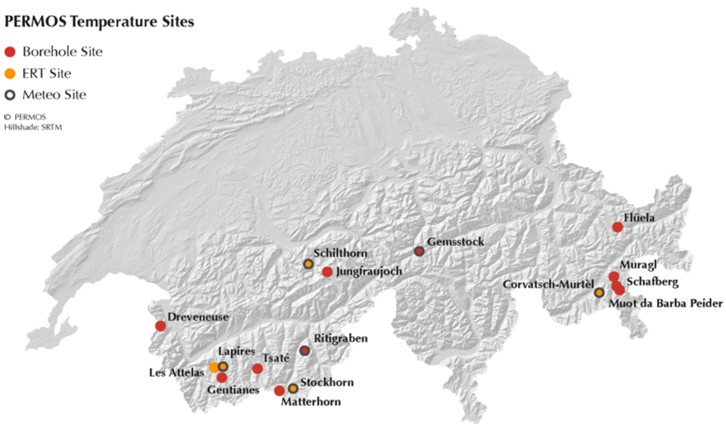


***Figure 4.9:*** *Overview on PERMOS borehole and geophysical and meteorological measurement sites.*

The longest record totalizes more than 30 years of observation, whereas the majority of the PERMOS boreholes has between 10 and 25 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 4.10) 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. We compare Permafrost_cci PFR with the Permafrost_cci rock glacier inventories compiled in 12 regions worldwide (RD-12). In addition, for the Swiss Alps, MAGT of the PERMOS boreholes is compared to the Permafrost_cci PFR product. Within the Permafrost_cci inventories, we selected only the landforms classified as rock glaciers, push moraines or a complex combination of the two, since they are the ones representative of permafrost occurrence. However, such inventories are not usable for temporal validation since active rock glaciers will be in 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 will enable region-wide temporal model validation in mountain permafrost (RD-12).

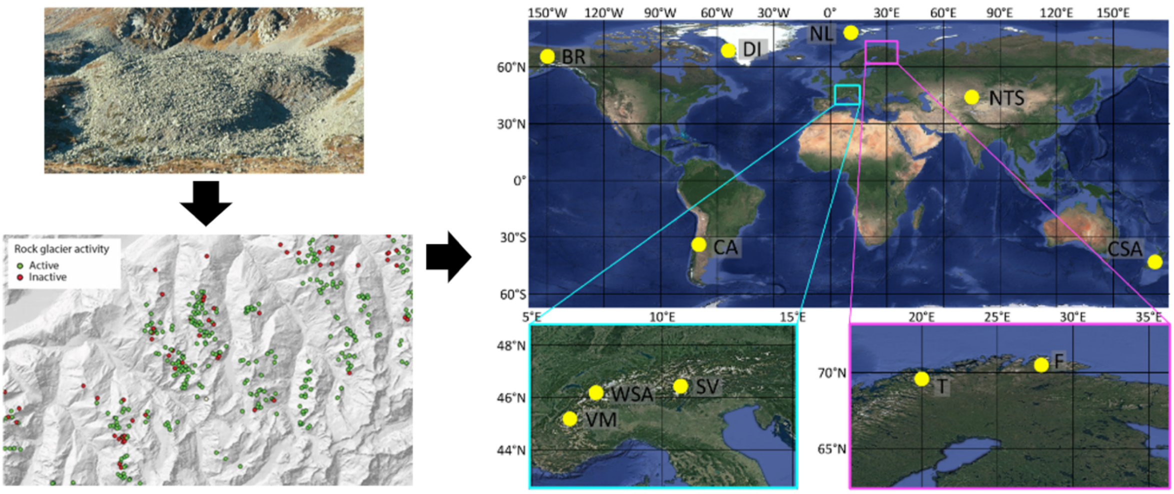


***Figure 4.10*** *Example of typical rock glacier (upper left panel), and rock glacier inventory in the Swiss Alps, which includes the state of activity of each landform (right lower left panel) and location of the Permafrost_cci inventories (right panel from Bertone et al. 2022).*

The regional assessment in mountainous areas is 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 is undertaken in close collaboration with the ongoing **IPA-funded Action Group on Rock Glacier inventories and kinematics** (2018-2023), 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-1 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.

# 5 VALIDATION DOCUMETS 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. Whereas the former is 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 are seeking 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 Rock Glacier Inventory and Kinematics Action Group**, and the CRG. If the results of the validation and round robin activities of the individual Permafrost_cci products can be presented in the form of publications and data publications, the largest possible endorsement is achieved.

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

| **Deliv.** | **Name** | **Date** | **Comment** |
|---|---|---|---|
| D1.3 | **DARD** | **phase I**<br>January 2019, December 2020 | describes data accessibility |
| D2.1 | **PVASR** | **phase I**<br>February 2019, November 2019, May 2021<br>**phase II**<br>May 2023 | summarizes algorithm selection |
| D2.3 | **E3UB** | **phase I**<br>February 2019, November 2019, March 2021<br>**phase II**<br>May 2023 | defines sources of errors and uncertainties |
| D2.5 | **PVP** | **phase I**<br>February 2019, November 2019, December 2020<br>**phase II**<br>May 2023 | outlines planned validation strategies |
| D4.1 | **PVIR** | **phase I**<br>September 2019, 2020, 2021 | provides a summary on quality and uncertainty of ECV products |
| D4.2 | **CRDP** | **phase I**<br>May 2020, February 2021 | describes the Climate Research Data Package, CRDP, a fully uncertainty characterized, long time series of Permafrost_cci products in compliance with CCI Data standards. The match-up data will be part of the CRDP in phase 2. |
| D4.3 | **PUG** | **phase I**<br>August 2019, 2020, February 2021 | describes delivered Permafrost_cci products in the CRDP |
| D5.2 | **CAR** | **phase I**<br>October 2019, 2020, September 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. and L'Hérault, E., 2016, Borehole and near-surface ground temperatures in northeastern Canada, v.1.4 (1988-2016). Nordicana D8, https://doi: 10.5885/45291SL-34F28A9491014AFD

Bergstedt, H. and Bartsch, A., 2020a, Near surface ground temperature, soil moisture and snow depth measurements in the Kaldoaivi Wilderness Area, for 2016-2018. PANGAEA, https://doi.org/10.1594/PANGAEA.912482

Bergstedt, H., Bartsch, A., Neureiter, A., Hofler, A., Widhalm, B., Pepin, N. and Hjort, J., 2020b, Deriving a frozen area fraction from Metop ASCAT backscatter based on Sentinel-1. IEEE Transactions On Geoscience And Remote Sensing, 58(9), 6008-6019. https://doi.org/10.1109/tgrs.2020.2967364

Bergstedt, H., Bartsch, A., Duguay, C and Jones, B., 2020c, Influence of surface water on coarse resolution C-band backscatter: Implications for freeze/thaw retrieval from scatterometer data. Remote Sensing of Environment, 247. https://doi.org/10.1016/j.rse.2020.111911

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. and Lantuit, H., 2019, Permafrost is warming at a global scale. Nature Communications, 10, 264. https://doi.org/10.1038/s41467-018-08240-4

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 System Science Data, 7, 245-259. https://doi.org/10.5194/essd-7-245-2015

Boike, J., Nitzbon, J., Anders, K., Grigoriev, M.N, Bolshiyanov, D.Y., Langer, M., Lange, S., Bornemann, N., Morgenstern, A., Schreiber, P., Wille, C., Chadburn, S., Gouttevin, I. and Kutzbach, L., 2019, Soil data at station Samoylov (2002-2018, level 1, version 201908). PANGAEA, https://doi.org/10.1594/PANGAEA.905233

Boike, J., Juszak, I., Lange, S., Chadburn, S., Burke, E.J., Overduin, P.P., Roth, K., Ippisch, O., Bornemann, N., Stern, L., Gouttevin, I. Hauber, E. and Westermann, S., 2017, Soil data at station Bayelva (1998-2017, level 2, version 1). PANGAEA, https://doi.org/10.1594/PANGAEA.882061

Brown, J., Ferrians, O., Heginbottom, J.A. and Melnikov, E., 2002. Circum-Arctic Map of Permafrost and Ground-Ice Conditions, Version 2. [Indicate subset used: permaice.shp]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center, https://doi.org/10.7265/skbg-kf16 [Date Accessed: 11.10.2018].

Brown, J., Hinkel, K.M. and 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. and 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, 234p.

Burgess, M.M., Smith, S.L., Brown, J., Romanovsky, V.E. and Hinkel, K., 2000, The Global Terrestrial Network for Permafrost (GTNet-P): permafrost monitoring contributing to global climate observations, Geological Survey of Canada.

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

Clow, G.D., 2014, Temperature data acquired from the GTN-P Deep Borehole Array on the Arctic Slope of Alaska, 1973–2013, Earth System Science Data, 6, 201-218. https://doi.org/10.5194/essd-6-201-2014

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

GTN-P, 2021, Long-term mean annual ground temperature data for permafrost. PANGAEA, https://doi.org/10.1594/PANGAEA.930669

GTN-P, 2018, GTN-P global mean annual ground temperature data for permafrost near the depth of zero annual amplitude (2007-2016). PANGAEA, https://doi.org/10.1594/PANGAEA.88471

Harris, C., Haeberli, W., Vonder Mühll, D. and King, L., 2001, Permafrost monitoring in the high mountains of Europe: the PACE Project in its global context. Permafrost and Periglacial Processes, 12, 3-11. https://doi: 10.1002/ppp.377

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.

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

Isaksen, K., Vonder Mühll, D., Gubler, H., Kohl, T. and Sollid, J.L., 2000, Ground surface-temperature reconstruction based on data from a deep borehole in permafrost at Janssonhaugen, Svalbard. Ann. Glaciol. 31, 287-294. https://doi: 10.3189/172756400781820291

Kroisleitner, C., Bartsch, A. and Bergstedt, H., 2018, Circumpolar patterns of potential mean annual ground temperature based on surface state obtained from microwave satellite data. The Cryosphere, 12, 2349-2370. https://doi.org/10.5194/tc-12-2349-2018

Kudryavtsev V.A., Kondrat'eva K.A. and Gavrilov A.G., 1978, Geocryological Map of the USSR, General Permafrost Studies, Materials for the Third International Conference on Permafrost. Nauka: Novosibirsk; scale 1:2,500,000 (in Russian).

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.

Nicolsky, D.J., Romanovsky, V.E., Kholodov, A.L., Dolgikh K. and Hasson, N., 2020, ABoVE: Soil Temperature Profiles, USArray Seismic Stations, AK and Canada, 2016-2019. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1767

Nicolsky, D.J., Romanovsky, V.E., Kholodov, A.L., Dolgikh K. and Hasson, N., 2019, ABoVE: Soil Temperature Profiles at USArray Seismic Stations, Alaska, 2016-2018. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1680

Nötzli J., Arenson, L.U., Bast, A., Beutel, J., Delaloye, R., Farinotti, D., Gruber, S., Gubler, H., Haeberli, W., Hasler, A., Hauck, C., Hiller, M., Hoelzle, M., Lambiel, C., Pellet, C., Springman, S.M., Vonder Muehll, D. and Phillips, M., 2021, Best Practice for Measuring Permafrost Temperature in Boreholes Based on the Experience in the Swiss Alps. Frontiers in Earth Science, 9. https://doi.org/10.3389/feart.2021.607875

Paulik, C., Melzer, T., Hahn, S., Bartsch, A., Heim, B., Elger, K. and Wagner, W., 2014, 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). PANGAEA, https://doi.org/10.1594/PANGAEA.832153

Rey, D.M., Ebel, B.A., James, S.R., Minsley, B.J. and Walvoord, M.A., 2020, Geophysical and related field data from the West Fork of Dall Creek, AK 2017-2019. U.S. Geological Survey data release, https://doi.org/10.5066/P977NU4B

Romanovsky, V.E., Smith, S.L. and Christiansen, H.H., 2010, Permafrost thermal state in the polar northern hemisphere during the international polar year 2007-2009: a synthesis. Permafrost and Periglacial Processes, 21, 106–116. https://doi: 10.1002/ppp.689

Smith, S., Brown, J., Nelson, F.E., Romanovsky, V.E., Zhang, T., Christiansen, H.H., et al., 2009, Permafrost and Seasonally Frozen Ground. Rome: Global Terrestrial Observing System.

Streletskiy, D.A., Biskaborn, B.K., Smith, S., Noetzli, J., Vieira, G. and Schoeneich, P., 2017, Strategy and Implementation Plan 2016–2020 for the Global Terrestrial Network for Permafrost GTN-P. Washington DC: The George Washington University, 42p.

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

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

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. and Hill, K., 2018, A Synthesis Dataset of Permafrost-Affected Soil Thermal Conditions For Alaska, USA. Earth System Science 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.

WMO, 2021. Chapter 2 - Measurement of Temperature (2.1.4.2.2 Measuring soil temperatures), in: Guide to Instruments and Methods of Observation Volume I – Measurement of Meteorological Variables, WMO-No. 8. World Meteorological Organization (WMO), Geneva, Switzerland, 90–124.

### 6.2 Acronyms

| | |
| --- | --- |
| ABoVE | Arctic-Boreal Vulnerability Experiment |
| AD | Applicable Document |
| ALD | Active Layer Depth |
| ALT | Active Layer Thickness |
| 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 |
| CH | Switzerland |
| CRDP | Climate Research Data Package |
| CRG | Climate Research Group |
| DUE | Data User Element |
| ECV | Essential Climate Variable |
| EO | Earth Observation |
| EU | European Union |
| ESA | European Space Agency |
| FRM | Fiducial Reference Measurement |
| FT2T | Freeze-Thaw to Temperature |
| GAMMA | Gamma Remote Sensing |
| GCOS | Global Climate Observing System |
| GT | Ground Temperature |
| GTD | Ground Temperature per Depth |
| 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 |
| NASA | National Aeronautics and Space Administration |

| | |
|---|---|
| NSF | National Science Foundation |
| NSIDC | National Snow and Ice Data Center |
| OSCAR | Observing Systems Capability Analysis and Review Tool |
| PANGAEA | Data Publisher for Earth & Environmental Science |
| PERMOS | Swiss Permafrost Monitoring Network |
| PFR | Permafrost Fraction |
| PVP | Product Validation Plan |
| RD | Reference Document |
| RHM | Roshydromet |
| RMSE | Root Mean Square Error |
| RR | Round Robin |
| RRR | Rolling Requirements Review |
| SWE | Snow Water Equivalent |
| std dev. | Standard Deviation |
| TR | Technical Requirement |
| TSP | Thermal State of Permafrost |
| 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 |
| WGMS | World Glacier Monitoring Service |
| WMO | World Meteorological Organization |
| ZAA | Zero Annual Amplitude |