



**permafrost**  
cci

**CCI+ PHASE 1 – NEW ECVS**

**PERMAFROST**

**D2.2 ALGORITHM THEORETICAL BASIS DOCUMENT  
(ATBD)**

**VERSION 1.0**

**28 FEBRUARY 2019**

**PREPARED BY**

**b·geos**

*J* **GAMMA REMOTE SENSING**



**UiO : University of Oslo**



**UNI  
FR**  
UNIVERSITÉ DE FRIBOURG  
UNIVERSITÄT FREIBURG



**UN** West University  
of Timisoara

**TERRASIGNA™**

## Document Status Sheet

Issue	Date	Details	Authors
0.1	04.12.2018	Initial version	T. Strozzi (GAMMA)
0.2	17.12.2018	Structure update	AB, SW
0.3	22.02.2019	Inclusion of material from existing documents	AB
0.4	25.02.2019	First draft	SW
0.5	26.02.2019	Draft update	SW, AB
1.0	28.02.2019	Final version	AB

## Author team

Sebastian Westerman, 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 CONTENTS

Executive summary .....	5
1 Introduction .....	6
2 Scientific background.....	11
3 Justification on the algorithm chosen.....	13
4 Processing line.....	15
5 Required input data .....	22
6 Output products .....	23
7 Error budget estimates and expected accuracy .....	24
8 Practical considerations for implementation .....	25
10 References .....	26

## EXECUTIVE SUMMARY

This document provides a detailed description of the algorithm chosen to generate the Permafrost ECV products to be delivered in the Permafrost\_cci project. The processing chain is described, including pre-processing of input data, the main processing steps using the ground thermal model CryoGrid CCI, and post-processing of model results. The CryoGrid CCI ground thermal model is based on existing permafrost models, but is optimized for efficient computation which is required to facilitate application for the entire permafrost domain. For each pixel, a model ensemble is computed in order to simulate the considerable subgrid variability (1km resolution) of the ground thermal regime observed in field-based studies. The model ensemble takes the spatial variability of landcover and ground stratigraphy, as well as snow depth into account, what allows for computation of a permafrost fraction. At the boundary of the permafrost domain, this makes it possible to quantitatively reproduce the well-established zonations in discontinuous and sporadic permafrost. In year 1, the processing chain makes use of remotely sensed land surface temperatures and landcover, which is supplemented by a range of ancillary information, such as ground stratigraphies and fields of the ERA-5 reanalysis. In years 2 and 3, further data sets from ongoing CCI projects such as Snow\_cci and Fire\_cci will be employed to improve Permafrost\_cci products. The spin-up of the ground thermal profile will be improved in years 2 and 3 of the project by starting simulations already in the years 1979/1980 (instead of 2003 as in year 1), so that a spin-up period of more than ten years can be achieved. The CryoGrid CCI model delivers the thermal profile from the surface downwards at sufficient temporal resolution to resolve the seasonal cycle. The output variables computed from the model state are mean annual ground temperature at different depths, annual active layer thickness and permafrost fraction within 1km pixels.

# 1 INTRODUCTION

## 1.1 Purpose of the document

The parameters required by the Global Climate Observing System (GCOS) for the Essential Climate Variable (ECV) Permafrost are Depth of Active Layer (the seasonal thaw layer) and Permafrost Temperature. ESA Permafrost\_cci will provide the ‘core’ ECV product Permafrost Temperature and the two ‘side’ ECV products Permafrost Active Layer and Permafrost Extent.

This document details the theoretical basis of the Permafrost\_cci algorithm designed to deliver the Permafrost ECV products.

## 1.2 Structure of the document

This document contains a glossary of permafrost specific terms in section 1.7. Section 2 describes the scientific background of the Permafrost\_cci algorithm. The remaining sections detail the theoretical and practical background of the selected algorithm, including input and output data as well as treatment of uncertainties.

## 1.3 Applicable Documents

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

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

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

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

#### 1.4 Reference Documents

[RD-1] Bartsch, A., Matthes, H., Westermann, S., Heim, B., Pellet, C., Onacu, A., Kroisleitner, C., Strozzi, T.(2019): ESA CCI+ Permafrost User Requirements Document, v1.0

[RD-2] Bartsch, A., Westermann, Strozzi, T., Wiesmann, A., Kroisleitner, C. (2019): ESA CCI+ Permafrost Product Specifications Document, v1.0

[RD-3] Bartsch, A., Westermann, S., Heim, B., Wieczorek, M., Pellet, C., Barboux, C., Kroisleitner, C., Strozzi, T. (2019): ESA CCI+ Permafrost Data Access Requirements Document, v1.0

[RD-4] Bartsch, A.; Grosse, G.; Kääh, 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-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](https://ipa.arcticportal.org/images/stories/AG_reports/IPA_AG_SucessorMap_Final_2016.pdf)

[RD-6] Westermann, S., Bartsch, A., Strozzi, T. (2019): ESA CCI+ Product Validation and Assessment Report, v1.0

#### 1.5 Bibliography

A complete bibliographic list that support arguments or statements made within the current document is provided in Section 10.1.

#### 1.6 Acronyms

A list of acronyms is provided in section 10.2.

## 1.7 Glossary

The list below provides a selection of term relevant for the parameters addressed in CCI+ Permafrost. A comprehensive glossary is available as part of the Product Specifications Document [RD-2].

### **active-layer thickness**

The thickness of the layer of the ground that is subject to annual thawing and freezing in areas underlain by permafrost.

The thickness of the active layer depends on such factors as the ambient air temperature, vegetation, drainage, soil or rock type and total water 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 ( $T < 0^{\circ}\text{C}$ ).

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 ( $T > 0^{\circ}\text{C}$ ) 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.

<u>Permafrost</u>	<u>English usage</u>	<u>Russian Usage</u>
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 metres below the ground surface as a result of past changes in surface and climate conditions. Below that depth, it will increase as a result of the geothermal heat flux from the interior of the earth. The mean annual ground temperature at the depth of zero annual amplitude is often used to assess the thermal regime of the ground at various locations.

## **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 SCIENTIFIC BACKGROUND

Permafrost and frozen ground are key elements of the terrestrial Cryosphere that will be strongly affected by a warming climate. With widespread permafrost degradation likely to occur in this century, remote sensing of permafrost is seeking to unveil the processes and causal connections governing this development, from the monitoring of variables related to the permafrost state to the mapping of the impacts of degradation and potential natural hazards on the ground.

Remote sensing of permafrost is challenging. The physical subsurface variables which characterize its thermal state – ground temperature, ice content and thaw depth – are not directly measurable through current remote sensing technologies. In order to gain spatially continuous information, numerical models are key. Permafrost\_cci will therefore produce Level 4 datasets, based on CDRs from other ECVs [RD-2].

### 2.1 Permafrost monitoring to date

At present, terrestrial systems are in place within which these physical variables are monitored at specific sites, either continuously (as typical in boreholes for ground temperature in the GTN-P) or sporadically during a season (as for active layer thickness at CALM sites), or even only every few years. While the terrestrial monitoring has been drastically expanded during and after the “International Polar Year” (IPY), the distribution of sites is strongly biased towards a few regions (typically where resource extraction and/or infrastructure projects have created easy access), leaving vast areas uncovered by monitoring. This in particular renders upscaling of trends in the permafrost ECV to global scale problematic. Furthermore, it complicates validation of Earth System Model output for to permafrost, which causes considerable uncertainty in future projections on the permafrost-carbon feedback.

The main requirement for EO-based algorithms for permafrost ECV generation is therefore to improve the spatial and temporal coverage compared to the existing in-situ networks, while at the same time providing consistent coverage of at least all relevant permafrost regions. Since EO-based algorithms necessarily operate at the spatial scale of individual pixels, the spatial resolution of the output must be put in context with the spatial variability of permafrost temperatures and active layer thickness.

### 2.2 Thermal subsurface modeling

In the ESA GlobPermafrost project, a simple equilibrium permafrost model has been employed to infer ground temperatures and permafrost extent based on thawing and freezing degree days derived from

remotely sensed data. This model mainly has the advantage of simplicity which makes global application possible without a high performance computing system. However, the transient response of the ground thermal regime to a changing climate forcing cannot be reproduced with an equilibrium scheme, and instead, a transient model must be used. Transient thermal subsurface models are now widely employed in permafrost science. While they use a wide variety of forcing data sets, all state-of-the-art models rely on the same model physics, numerically solving the heat conduction equation including the phase change of water/ice in soil pores. The feasibility and maturity of transient permafrost models in conjunction with EO-data over large spatial domains has recently been demonstrated in a pilot study in Northern Siberia (Westermann et al., 2017). EO data sets are employed to determine the upper boundary condition of the differential equation, while its coefficients (e.g. heat capacity and thermal conductivity) are selected according to landcover information (Westermann et al., 2017). With this, a spatial resolution of the final product of 1 km is possible, corresponding to “breakthrough” according to the WMO OSCAR database [RD-1]. Other transient permafrost models are well-established in the community, most notably GIPL 2 (Jafarov et al., 2012) at the University of the Arctic Fairbanks. Transient permafrost models are highly demanding computationally ( $>10^6$  more CPU hours compared to an equilibrium model), so that High Performance Computing frameworks must be used. On the other hand, validation with field data is possible in a much more strict way, as ground temperature and thaw depth measured at a specific depth and date can be compared.

### 3 JUSTIFICATION ON THE ALGORITHM CHOSEN

#### 3.1 User requirements

The Permafrost CCI project gathered existing user requirements from international consensus references, such as GCOS-200, the IGOS Cryosphere theme report (WMO, 2007) and the WMO RRR database (WMO, 2011a), as well as from the climate modelling community through the Modelling Working Group of the Permafrost Carbon network (PCN). The requirements review further covered user survey results from ESA DUE GlobPermafrost, a white paper in the framework of the WMO Polar Space Task Group (Bartsch et al. 2014), workshop reports (e.g. NRC 2014) and discussions with representatives of the International Permafrost Association (IPA) and the IPA Action Group ‘Specification of a Permafrost Reference Product in Succession of the IPA Map’. Views from climate researchers associated with the project (AWI, University of Fribourg, West University of Timisoara) were also taken into account.

Users demand a combination of extensive geographical coverage (global permafrost extent 20-30 Mio km<sup>2</sup>), high spatial resolution (target resolution 1km) including representation of subgrid variability, high temporal resolution (monthly data) and long temporal coverage (one to several decades back in time). These requirements go considerably beyond the state-of-the-art in remote permafrost ECV assessment, based on published studies and recently demonstrated progress. The main requirement for Permafrost CCI is to implement a Permafrost ECV processing system that can retain the capabilities (e.g. spatial resolution, assessment of subpixel variability, global coverage) of the GlobPermafrost products, but at the same time can evaluate the transient response of the ground thermal regime to changes in the climatic forcing.

In response to the user requirements and in agreement with GCOS Action T34, the Permafrost\_cci project will deliver novel products for ground temperature and active layer thickness, based on the transient permafrost model CryoGrid CCI, which is compiled for the project based on the existing Cryogrid 2 model (Westermann et al., 2013; 2017).

#### 3.2 Technical requirements

Apart from specific accuracy and resolution requirements, users have stressed the need for the development of a suitable benchmark dataset, as it does not exist yet to date. While the terrestrial monitoring has been drastically expanded during and after the “International Polar Year” (IPY), the

distribution of sites is strongly biased towards a few regions (typically where resource extraction and/or infrastructure projects have created easy access), leaving vast areas uncovered by monitoring.

Accuracy requirements stated by the users are strongly complicated by the fact that permafrost ECV physical variables (ground temperature and active layer thickness) often feature significant variations at spatial scales below the target requirement of 1km, which in the few documented cases exceed even the threshold requirement of an RMSE of 2.5K (e.g. Fig. 2 in Gislén et al., 2014). Therefore, even comparison of “perfect” 1km average temperatures to point temperature measurements in boreholes will feature a significant RMSE which in this case rather reflects the spread of temperatures in space than the accuracy of the method (assuming that boreholes are placed at random locations within a pixel). In real-world permafrost ECV assessment, a bias introduced by the method/model and the input data will overlap with this effect, which significantly complicates the evaluation of accuracies.

Since EO-based algorithms necessarily operate at the spatial scale of individual pixels, the spatial resolution of the output must be put in context with the spatial variability of permafrost temperatures and active layer thickness. In many permafrost regions, these can display a strong variability at spatial scales of meters, which is generally much finer than the footprint of EO sensors. For this reason, it is desirable to also derive permafrost extent (fraction) as permafrost ECV parameter (in line with the WMO RRR database), which is the aerial fraction within an area (pixel) at which the definition for the existence of permafrost (ground temperature  $< 0^{\circ}\text{C}$  for two consecutive years) is fulfilled. The characterization of the permafrost extent in terms of aerial coverage has been employed for decades in the permafrost community, e.g. in the classic IPA permafrost map displaying classes of continuous, discontinuous, sporadic and isolated permafrost. Furthermore, the ice content in the ground is an important factor for active layer thickness, especially when the active layer deepens in the course of a warming climate. Therefore, a spatially distributed product of ground stratigraphies is required as input to the permafrost model in order to achieve a satisfactory performance for the active layer thickness. At present, such a product does not exist globally, but Permafrost\_cci is in the process of creating an initial version combining various data sources to a global product.

## 4 PROCESSING LINE

### 4.1 The CryoGrid CCI ground thermal model

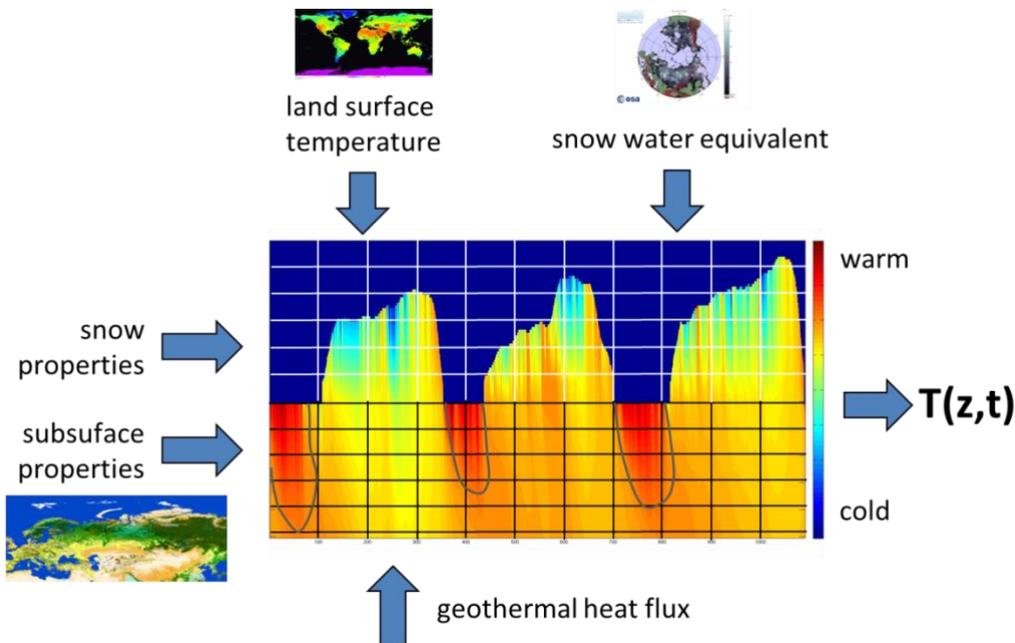
The CryoGrid family of ground thermal models has been employed for a wide variety of tasks, with the simple equilibrium model CryoGrid 1 (e.g. Westermann et al., 2015) used in ESA GlobPermafrost to compile the first EO-derived global map of permafrost extent and temperatures (Obu et al., in review). For Permafrost CCI, we mainly rely on the transient model CryoGrid 2 (Westermann et al. 2013), but improve it with elements of the more sophisticated land surface scheme CryoGrid 3 (Westermann et al., 2016), as well as the CROCUS snow model (Vionnet et al., 2012). Furthermore, the main prognostic state variable, which is integrated in time, is changed from temperature to enthalpy, in line with the GIPL2 model (Jafarov et al., 2012), which offers advantages with respect to energy conservation. The combined model scheme is referred to as CryoGrid CCI in the following.

CryoGrid 2 numerically solved the heat conduction equation in the enthalpy formulation:

$$\frac{dE}{dt} = \frac{d}{dz} \left( K(T(E)) \frac{dT(E)}{dz} \right),$$

with  $E$  denoting the enthalpy [ $\text{J}/\text{m}^3$ ],  $t$  time [sec],  $K$  the thermal conductivity [ $\text{W}/\text{mK}$ ],  $T$  temperature [ $\text{K}$ ] and  $z$  the vertical space coordinate [m]. The left-hand side of the equation corresponds to the derivative of enthalpy in  $t$  time, while the right-hand side corresponds to the divergence (i.e. gradient in one dimension) of the conductive heat flux. The temperature  $T$  can be computed from enthalpy using a unique function, which takes into account both the heat capacity and latent heat change effects due freezing/melting of water/ice. The thermal conductivity is again obtained as a function of temperature, as described in Westermann et al., (2013). The heat conduction equation is discretized in space using the method of lines, while we use the simplest possible time integration scheme, first-order forwards Euler (as in Westermann et al. 2013), to integrate the fields of enthalpy and temperature in time. The main advantage of the enthalpy formulation is that it is by definition satisfies conservation of energy for the given heat fluxes. Furthermore, the enthalpy variable does not feature a strong discontinuity at the freezing point of water,  $0^\circ\text{C}$ , as it is the case for temperature. The function temperature vs. enthalpy is parametrized analogous to many Earth system models (e.g. Gouttevin et al., 2012), with a freezing range between  $0^\circ\text{C}$  and a predefined freezing temperature  $T_{\text{freeze}}$ , between which the phase change of water/ice is assumed to occur. While this is only a simplification of a real freezing characteristic as employed in

CryoGrid 2 model, it features strong advantages in particular with respect to model runtime. Furthermore, the method allows for a consistent treatment of free water/ice, which allows full continuity between the treatment of the ground and the snow cover.



**Figure 1:** Schematic representation of the CryoGrid CCI model scheme, showing the input data land surface temperature and snow water equivalent flow which time series are required. Snow properties are computed by the model scheme, while subsurface properties are assigned based on the ESA CCI landcover classification. At the lower boundary of the model domain, a (constant) geothermal heat flux is prescribed. CryoGrid CCI delivers depth- and time-resolved fields of ground temperatures, from which annual averages at different depths, as well as annual active layer thickness (from thaw progression, grey line).

#### 4.1.1 Snow model

A major challenge for spatially distributed application of thermal models is the snow density, which is variable in space, but determines both snow depth (if snow water equivalent is provided) and the thermal conductivity of the snow cover. In regional studies, the snow density is mostly prescribed based on field measurements (e.g. Westermann et al., 2013; 2017), but this method already fails for larger regions with stronger environmental gradients (e.g. Nicolsky et al., 2017), making global application essentially impossible without factoring in the main drivers for snow depth. In Cryogrid CCI, we employ a snow compaction model oriented at the CROCUS scheme, taking into account a temperature- and wind-speed dependent initial snow density, snow compaction due to the overburden of overlying snow layers, as well as compaction by wind drift (see Vionnet et al., 2012). Note that application of a fully-fledged snow model is not possible in the context of Permafrost\_cci, as this would require explicitly solving the

surface energy balance, instead of relying on remotely sensed data sets. The snow model makes knowledge of wind speeds necessary, which control the initial density of fresh snow and the degree of compaction due to wind drift. In year 1, three snow layers are employed, while the number of layers is likely to increase in years 2 and 3. Phenomenologically, the goal of the snow scheme is to reproduce observed large-scale gradients of the snow density, i.e. snow density increases with snow water equivalent, temperature and windiness, which have already been accounted for in the ESA GlobPermafrost project (using the formulation of Onuchin and Burenina, 1996).

#### **4.1.2 Ensemble generation**

For each 1km pixel, time series of surface temperature, snowfall and snowmelt are provided. To estimate the spatial variability of ground temperature and active layer thickness within a pixel, we generate an ensemble corresponding to the best guess for equal-area subpixel fractions, taking a range of factors into account. For ensemble generation, we first use subpixel landcover fractions to define a stratigraphies for each ensemble member, according to the fractional coverage of landcover classes (see below). Note that landcover classes with low fractions are not accounted for if the number of ensemble members is not sufficient. For each landcover class present, snowfall is further scaled according to a lognormal distribution function, using coefficients of variations that are again dependent on landcover (according to Liston, 2004). This way, the natural variability of snow depths which gives rise to strong small-scale differences of ground temperatures especially in tundra areas (e.g. Gislén et al., 2014) is accounted for in a statistical way.

#### **4.1.3 Model operation and technical aspects**

The vertical model domain extends to a depth of 100m below the surface, with a vertical discretization of 0.1 m within the uppermost 1.5 m to resolve the active layer in detail. In deeper layers, discretization gradually increases up to a spacing of 10 m for the lowermost grid cells. A constant model time step of 1 hour (3600 sec) is employed in the time routine, which simplifies massively parallel application as required in Permafrost CCI. In year 1, the model is initialized to steady-state condition for the first two model years (2004 and 2004), following the computationally efficient step-wise spin-up procedure outlined in Westermann et al. (2013), which can estimate both the steady-state thermal gradient in deeper layers and the “transient” steady-state in the upper layers. Initialization with only two model years is clearly problematic, but inevitable due to the short time series of only 15 years (2003-2017). In year 2, this problem will be overcome by initializing model runs already in 1980 and running transiently until

the first epoch 1997, which so that a spin-up period of more than ten years is available. This is considered sufficient so that the near-surface temperature profile is independent of the applied initialization.

## 4.2 Pre-processing

Due to availability of the remote sensing data sets, the processing line is set for the period 2003 to 2017 in year 1 of Permafrost\_cci. The spatial resolution is 1km, in line with user requirements. The processing chain starts with pre-processing of forcing data sets for the CryoGrid CCI thermal model, which are compiled as eight day averages to drive the model. In the second step, the CryoGrid CCI thermal model is run in ensemble mode (i.e. simulating a number of different realizations for each pixel), using predefined distributions for the snow and higher (300m) resolution data for landcover as input. Finally, postprocessing of the model output leads to generation of the Permafrost\_cci data sets for the Permafrost ECV physical variables ground temperature and active layer thickness. In addition, ensemble mode (as demonstrated in ESA Globpermafrost) can deliver permafrost probabilities/ fractions demanded by the user community.

### 4.2.1 Land surface temperature

In year 1, land surface temperatures (LST) are employed from the daily level 2 version 6 MODIS LST collections MOD1A11/MYD1A11, using data sets from both the Terra and Aqua satellites. For cloud-free skies, these feature up to four values per day (one day and night-time value for each satellite). The overpass time is variable, as the overpass closest to nadir is selected for which the scene is free of clouds. From year 2 onwards, Permafrost\_cci aims to use newly compiled LST data sets from Temperature\_cci.

In high-latitude regions, prolonged periods of cloudiness constitute a major challenge for remotely sensed LST, and validation studies have demonstrated a considerable bias of several Kelvin in longer-term (weekly to seasonal) LST averages inferred from satellites (Langer et al., 2010; Westermann et al. 2011; 2012; Østby et al., 2014). For this reason, it is crucial to gap-fill the LST time series during cloudy periods. In Permafrost\_cci, we will follow the successful gap-filling approach taken in Westermann et al., 2015 and ESA Globpermafrost, which uses air temperatures from the ERA reanalysis for gap-filling. Other than in ESA GlobPermafrost, Permafrost\_cci can benefit from the much improved ERA-5 reanalysis which covers the entire time period and offers a much improved spatial and vertical resolution of the atmospheric model. The coarse-scale (~30km) reanalysis data are first downscaled to 1km resolution using a Digital Elevation Models (DEM, REF), relying on information from different pressure levels to account for atmospheric lapse rates in the target regions (see Westermann et al., 2015 for details). Finally, downscaled air temperatures are inserted to fill the gaps in the MODIS LST time series,

using the actual timestamp of the MODIS retrieval to define the gaps. From the combined time series, eight day (8d) LST averages are computed, which are used to force the CryoGrid CCI thermal model.

#### **4.2.2 Snow cover**

To take the insulating effect of the seasonal snow cover on the ground thermal regime into account, time series of snow depths are required in thermal modelling. Time series derived from space-borne passive microwave sensors have been employed to drive thermal models for a small region in Siberia (Westermann et al., 2017), but work in the ESA GlobPermafrost showed major deficits especially in mountain and coastal areas, as well as in areas with a high fraction of water bodies. These findings are in line with major validation studies (e.g. Takala et al., 2013; see also open discussion of Westermann et al., 2017). Furthermore, the coarse spatial resolution of ~25km would strongly decrease the final resolution of Permafrost\_cci products below threshold resolution [RD-1], and compromise acceptance in the permafrost community. Similar to ESA GlobPermafrost, we use a degree-day based snow model driven by downscaled reanalysis data in year 1 to overcome these problems, leading to a snow data at a spatial resolution matching the LST input data sets (1km). Snowfall data sets are generated by downscaling ERA-5 precipitation fields with altitude (5% increase per 1000m) and applying a threshold criterion of 0°C of the downscaled air temperatures (see above). For the melt rate, we again use downscaled air temperatures in conjunction of a melt threshold (0°C) and latitude- and day-of-year dependent degree-day factors, as in ESA GlobPermafrost. Both snowfall and melt are accumulated for the same eight day periods as the LST forcing.

#### **4.2.3 Land cover and subsurface information**

To apply spatially distributed data sets of surface and subsurface parameters, a spatial classification is required. As in ESA GlobPermafrost, Permafrost\_cci will rely on the Landcover\_cci classification at 300m resolution, aggregating classes to seven “super-classes”: bare, vegetated tundra, grassland, shrub, coniferous forest, deciduous forest and wetland. For application in ensemble mode of CryoGrid CCI, we calculate the areal fractions of each super-class for each 1km pixel.

For application of any thermal subsurface model, a stratigraphy of subsurface thermal parameters must be defined for the entire vertical model domain (see above). In Cryogrid CCI, we use the parametrizations of the Cryogrid 2 model (Langer et al., 2013; Westermann et al., 2013; 2017), which allows translating volumetric fractions of the mineral, organic, air and water/ice phases within the ground to the “native “ model parameters heat capacity and thermal conductivity (see above). The advantage of this method is that the abundance of field data sets on these parameters can be employed

in modelling, using data from soil cores and pits, as well as exposed sections along river banks and construction sites. In year 1, the process of assigning typical stratigraphies to each of the superclasses is still ongoing, so generic stratigraphies based on a previous modelling studies with CryoGrid models (Langer et al., 2013; Westermann et al., 2013; 2016; 2017) have been assigned. From year 2 onwards, we aim for using a novel stratigraphy product based on Landcover\_cci products in the Permafrost\_cci ECV compilation.

#### 4.2.4 Ancillary data

To drive the snow model of CryoGrid CCI (see above), an estimate for wind speeds at the model scale of 1km is required. As for precipitation/snowfall, we rely on downscaled ERA-5 reanalysis products for this. A major challenge is that surface wind speeds in the reanalysis strongly depend on the native roughness length, which is determined as a function of a coarse-scale (30km) landcover classification. This naturally does not match the Landcover\_cci classification employed in Permafrost\_cci (see above). To avoid inconsistencies, we therefore use pressure level information on wind speeds to interpolate both laterally and vertically to the target altitude, followed by application of a log-normal wind profile (i.e. assuming the simplest case of a neutral stratification) to obtain 2m wind speeds as required by the snow model. The roughness length is hence defined for each ensemble member in CryoGrid CCI, depending on the fractions of landcover superclasses within a 1km pixel. As a result, 2m wind speeds in e.g. forested areas are significantly lower than in tundra areas, which at least qualitatively matches reality and in the processing chain leads to a higher degree of snow insulation in forest compared to tundra areas. While not yet explored in detail, ERA-5 derived wind speeds appear strongly biased to low values in low-lying areas in high-topography regions, e.g. on Svalbard and the coast of Greenland.

### 4.3 Post processing

For each ensemble member, the resulting ground temperature field (in time and depth) is averaged over yearly periods for different depths (1m, 2m, and 5m in year 1), and the active layer thickness is determined as the depth of maximum thaw within each year. For each pixel, we then determine both mean and the standard deviation of temperatures. The standard deviation provides a robust first-order assessment of the uncertainty of the mean, corresponding to the expected spread in ground temperatures when measuring at a random location within a pixel. As Permafrost\_cci is strongly directed towards the climate modelling community, which in general uses only a single realization for a grid cell, the ensemble spread is exploited for the characterization of uncertainties for both ground temperatures and active layer thickness. Furthermore, the ensemble allows to calculate the fractional permafrost cover

which makes it possible to produce maps with the well-known permafrost zonation in continuous, discontinuous and sporadic permafrost. In year 1, we employ two different methodologies for inferring “permafrost presence” for an ensemble member. First, we use a simple threshold criterion at a depth of 2 m, which has the advantage of simplicity, but does not correspond to the definition of permafrost (below 0°C for two consecutive years) if the active layer is deeper than 2m. Therefore, we use a second method which explicitly checks if the first cell below the seasonal thaw layer (characterized by having both positive and negative ground temperatures) is constantly below the freezing point. While formally more correct, this method would classify short-term freezing of an upper ground layer due to an extended period of cold years as permafrost (potentially in accordance with the formal definition), although an unfrozen zone exists below the shallow frozen layer. The second method provides a formally correct way of deducing permafrost, so with a longer model time series in years 2 and 3, it is likely that this method will be finally adopted. However, data sets generated the simple first method might be useful for many applications of the science community, so that both methods will be employed at least in year 1. With permafrost presence inferred for each ensemble member, permafrost fractions can be calculated. Note that the low number of ensemble members in year 1 (5) already allows a first estimate for the permafrost zonations, with continuous permafrost corresponding to all realizations featuring permafrost, discontinuous to 3-4, sporadic to 1-2, while a pixel is flagged as permafrost-free if none of the ensemble members features permafrost.

## 5 REQUIRED INPUT DATA

Permafrost\_cci level 4 products are created from the analysis of lower level data, resulting in gridded, gap-free products [RD-2]. Permafrost\_cci will employ remotely sensed Land Surface Temperature, snow cover and landcover data sets, in conjunction with downscaled ERA-5 reanalysis data [RD-3]. Auxiliary data include ground stratigraphies from field investigations, as well as model parameterizations for crucial permafrost and snow processes.

In year 1, daily 1km MODIS LST data sets from 2003-2017 based on the level 3 version 006 collection (products MOD11A1 and MYD11A1) will be employed. These data are based on acquisitions of two satellites, Terra and Aqua, which in total provide four measurements per day for cloud-free conditions (two day and night retrievals each). Gap-filling as demonstrated in Westermann et al. (2015; 2017) and in ESA GlobPermafrost will be accomplished using ERA-5 reanalysis data which are downscaled to 1km resolution with the help of a digital elevation model (see below). Land cover information is required to define soil properties and snow redistribution (Westermann et al. 2015), and the ESA Landcover\_CCI product will be employed. For soil stratigraphies, generic estimates for the same landcover superclasses as employed in GlobPermafrost will be used. For downscaling of the ERA reanalysis data, a global digital elevation model is required which includes the Arctic and provides sufficient spatial resolution. The 1 km Global Multi-resolution Terrain Elevation Data 2010 (GMTDEM) complies to these requirements. From the ERA-5 reanalysis, surface fields for air temperature, precipitation, geopotential and wind speed are required, as well as air temperature, wind speed and geopotential at the pressure levels 800, 700, 600, 500, 400, and 200mbar.

In years 2 and 3, improved data sets from Temperature\_cci for land surface temperature will be employed instead of MODIS LST, if available in time for Permafrost ECV production. Furthermore, data of snow extent and fractional snow covered area from Snow\_cci will be ingested in the processing chain. External soil datasets developed for the Arctic (e.g. NCSCD, Hugelius et al. 2013, C-band SAR proxies as in Bartsch et al. 2016b and Reschke et al. 2012) will be combined with Landcover\_cci information to compile a spatially distributed ground stratigraphy product. To define stratigraphies for the classes, extensive field data sets from major permafrost regions (>7000 individual measurements) will be evaluated statistically, aiming for a more realistic representation of ground properties. In year 3, we will also attempt to employ data sets from the Fire\_cci project to explicitly account for areas, where fires have changed ground thermal properties, which results in major changes of e.g. the active layer thickness.

## 6 OUTPUT PRODUCTS

The primary variables requiring climate-standard continuity as defined by GCOS are ground temperature and active layer thickness. In addition, permafrost fraction and zone are secondary parameters, but of high interest to users. At product level 4, annual averages of ground temperature and annual maxima of thaw depth (active layer thickness) will be provided at 1km spatial resolution during year 1 of Permafrost\_cci. The output products address the main user requirements [RD-1], including those expressed by the Permafrost\_cci climate research group (CRG). All datasets will be provided in NetCDF format, with subsets for the two major permafrost domains in subsets in America and Eurasia.

## 7 ERROR BUDGET ESTIMATES AND EXPECTED ACCURACY

Here, we specify how product uncertainties are estimated reflecting the combination of sources of uncertainty. These include both the pre-processing of input data sets, the uncertainty propagation through the ground thermal model CryoGrid CCI, as well as additional model uncertainty due to neglected processes.

In principle, the uncertainty due to the first two points can be evaluated through ensemble methods, i.e. by drawing a randomized ensemble of model realizations perturbing input data sets and model parameters according to provided uncertainty characteristic of these data sets. This has been demonstrated in the study by Langer et al. (2013) for a site in Northern Siberia. The Permafrost\_cci algorithm based on CryoGrid CCI has been designed for application with such ensemble methods for uncertainty characterization. However, the main problem is that permafrost areas are mainly situated in high latitudes and in high mountain areas, for which e.g. characterizations of uncertainties of input data sets are often not applicable. This has been demonstrated for MODIS land surface temperature for which much higher uncertainties have been found in Arctic settings compared for benchmark sites typically located in mid-latitudes (Westermann et al. 2011; 2012; Østby et al., 2013). Therefore (in addition to the additional model uncertainty), it is questionable if automatic uncertainty characterization based on provided uncertainties of input data sets can provide meaningful characterization of uncertainties for the Permafrost ECV products. In a few cases, it can be possible to improve or establish additional uncertainty benchmarks of input data sets (e.g. land surface temperature or landcover) for permafrost sites, but a comprehensive evaluation cannot be accomplished in Permafrost\_cci, while it is also not possible due to the scarcity of validation sites for many parameters.

For this reason, it is important to synchronize any computed uncertainty through a direct comparison of the Permafrost ECV products with ground truth measurements. This is especially important since there are many more validation data sets available for the Permafrost ECV data sets ground temperature and active layer thickness, than there are for example for land surface temperature or snow extent from permafrost regions.

In Permafrost\_cci, we will therefore follow both ways of estimating the error budgets for the Permafrost ECV, with the goal to bring the two methods into agreement.

## 8 PRACTICAL CONSIDERATIONS FOR IMPLEMENTATION

Computational costs are one of the principal challenges of the transient Permafrost\_cci algorithm, with CryoGrid CCI application for the entire global permafrost domain in ensemble mode being considerably larger than previous applications of transient permafrost models (e.g. Jafarov et al., 2012; Westermann et al., 2013). For this reason, a scalable implementation on the Abel Supercomputing cluster at the University of Oslo the algorithm has been chosen, and the model code has been strongly improved with respect to runtime. Near-operational runs of the processing chain indicates that the entire permafrost domain (ca. 25 Mio km<sup>2</sup>) can be simulated for the period 2003-2017 with five ensemble members per grid cell for a computational cost of less than 100,000 CPU (which are confirmed for year 1). With the Permafrost\_cci algorithm implemented and tested, the Permafrost\_cci products can be produced at the specifications outlined in [RD-6].

A major challenge is that the considerable processing time (exceeding 2 months) leads to long turnover cycles, which considerably complicates model tuning to achieve an optimal performance. Furthermore, ongoing improvements of model input data sets, such as the employed ground stratigraphies, cannot be ingested continuously, but the improvements will only become apparent in the next scheduled model iteration. For this reason, model tuning cannot yet be finalized in year 1, but must be achieved in an ongoing process stretching over the entire Permafrost\_cci period.

## 10 REFERENCES

### 10.1 Bibliography

Bartsch, A., Höfler, A., Kroisleitner, C., Trofaier, A.M., 2016a. Land Cover Mapping in Northern High Latitude Permafrost Regions with Satellite Data: Achievements and Remaining Challenges. *Remote Sens.*, 8, pp.979.

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

Gisnås, K., Westermann, S., Schuler, T.V., Litherland, T., Isaksen, K., Boike, J. and Etzelmüller, B., 2014. A statistical approach to represent small-scale variability of permafrost temperatures due to snow cover. *The Cryosphere*, 8, pp.2063-2074.

Gouttevin, I., Krinner, G., Ciais, P., Polcher, J. and Legout, C., 2012. Multi-scale validation of a new soil freezing scheme for a land-surface model with physically-based hydrology. *The Cryosphere*, 6, pp.407-430.

Hugelius, G., Tarnocai, C., Broll, G., Canadell, J.G., Kuhry, P. and Swanson, D.K., 2013. The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions. *Earth System Science Data*, 5(1), pp.3-13.

Jafarov, E.E., Marchenko, S.S. and Romanovsky, V.E., 2012. Numerical modeling of permafrost dynamics in Alaska using a high spatial resolution dataset. *The Cryosphere*, 6(3), pp.613-624.

Langer, M., Westermann, S. and Boike, J., 2010. Spatial and temporal variations of summer surface temperatures of wet polygonal tundra in Siberia-implications for MODIS LST based permafrost monitoring. *Remote Sensing of Environment*, 114(9), pp.2059-2069.

Langer, M., Westermann, S., Heikenfeld, M., Dorn, W. and Boike, J., 2013. Satellite-based modeling of permafrost temperatures in a tundra lowland landscape. *Remote Sensing of Environment*, 135, pp.12-24.

Liston, G.E., 2004. Representing subgrid snow cover heterogeneities in regional and global models. *Journal of climate*, 17(6), pp.1381-1397.

Nicolosky, D.J., Romanovsky, V.E., Panda, S.K., Marchenko, S.S. and Muskett, R.R., 2017. Applicability of the ecosystem type approach to model permafrost dynamics across the Alaska North Slope. *Journal of Geophysical Research: Earth Surface*, 122(1), pp.50-75.

Onuchin, A.A. and Burenina, T.A., 1996. Climatic and geographic patterns in snow density dynamics, Northern Eurasia. *Arctic and Alpine Research*, 28(1), pp.99-103.

Reschke, J.; Bartsch, A.; Schlaffer, S.; Schepaschenko, D. Capability of C-Band SAR for Operational Wetland Monitoring at High Latitudes. *Remote Sens.* 2012, 4, 2923-2943.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E. and Willemet, J.M., 2012. The detailed snowpack scheme Crocus and its implementation in SURFEX v7. 2. *Geoscientific Model Development*, 5, pp.773-791.

Westermann, S., Langer, M. and Boike, J., 2011. Spatial and temporal variations of summer surface temperatures of high-arctic tundra on Svalbard—implications for MODIS LST based permafrost monitoring. *Remote Sensing of Environment*, 115(3), pp.908-922.

Westermann, S., Langer, M. and Boike, J., 2012. Systematic bias of average winter-time land surface temperatures inferred from MODIS at a site on Svalbard, Norway. *Remote Sensing of Environment*, 118, pp.162-167.

Westermann, S., Schuler, T., Gislås, K. and Etzelmüller, B., 2013. Transient thermal modeling of permafrost conditions in Southern Norway. *The Cryosphere*, 7(2), pp.719-739.

Westermann, S., Østby, T.I., Gislås, K., Schuler, T.V. and Etzelmüller, B., 2015. A ground temperature map of the North Atlantic permafrost region based on remote sensing and reanalysis data. *The Cryosphere*, 9(3), pp.1303-1319.

Westermann, S., Langer, M., Boike, J., Heikenfeld, M., Peter, M., Etzelmüller, B. and Krinner, G., 2016. Simulating the thermal regime and thaw processes of ice-rich permafrost ground with the land-surface model CryoGrid 3. *Geoscientific Model Development*, 9(2), pp.523-546.

Westermann, S., Peter, M., Langer, M., Schwamborn, G., Schirrmeister, L., Etzelmüller, B. and Boike, J., 2017. Transient modeling of the ground thermal conditions using satellite data in the Lena River delta, Siberia. *The Cryosphere*, 11(3), pp.1441-1463.

Østby, T.I., Schuler, T.V. and Westermann, S., 2014. Severe cloud contamination of MODIS land surface temperatures over an Arctic ice cap, Svalbard. *Remote Sensing of Environment*, 142, pp.95-102.

## 10.2 Acronyms

AD	Applicable Document
ALT	Active Layer Thickness
AWI	Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research
B.GEOS	b.geos GmbH

CCI	Climate Change Initiative
CDR	Climate Data Record
CRG	Climate Research Group
CRS	Coordinate Reference System
DARD	Data Access Requirements Document
ECV	Essential Climate Variable
EO	Earth Observation
ESA	European Space Agency
ESA DUE	ESA Data User Element
GAMMA	Gamma Remote Sensing AG
GCOS	Global Climate Observing System
GCMD	Global Change Master Directory
GIPL	Geophysical Institute Permafrost Laboratory
GTD	Ground Temperature at certain depth
GTN-P	Global Terrestrial Network for Permafrost
GUIO	Department of Geosciences University of Oslo
IPA	International Permafrost Association
IPCC	Intergovernmental Panel on Climate Change
LST	Land Surface Temperature
MAGT	Mean Annual Ground Temperature
MAGST	Mean Annual Ground Surface Temperature
NetCDF	Network Common Data Format
NSIDC	National Snow and Ice Data Center
PFR	Permafrost extent (Fraction)
PFF	Permafrost-Free Fraction
PFT	Permafrost underlain by Talik
PSD	Product Specifications Document
PSTG	Polar Space Task Group
PZO	Permafrost Zone
RD	Reference Document
RMSE	Root Mean Square Error
RS	Remote Sensing
SLF	Institut für Schnee- und Lawinenforschung, Davos
SU	Department of Physical Geography Stockholm University
TSP	Thermal State of Permafrost
UAF	University of Alaska, Fairbanks
UNIFR	Department of Geosciences University of Fribourg
URD	User Requirements Document
WGS 84	World Geodetic System 1984
WUT	West University of Timisoara