Our response to the referee reports is structured in three parts. First, we give a point-by-point response to all points raised by Referees 1 and 2. We then summarize the main changes to the revised manuscript, followed by the revised version of the manuscript with changes marked in bold font.

On behalf of all authors,

Sebastian Westermann

Anonymous Referee #1

This study makes use of the satellite remote sensing and the weather reanalysis data to produce land surface temperatures (LST), which is an important component for modeling the permafrost thermal state, and correspondingly the permafrost extent. This work addresses how LST could be used by the semi-empirical simple permafrost model to calculate present permafrost extent and, more importantly, ground temperatures. I have several suggestions and comments that could be used to improve the current manuscript. I also have several questions. Overall, I enjoyed reading the manuscript and will suggest for the publication after revision.

In the Introduction section, I suggest to mention similar studies by Panda et al, (2014) and Luo et al, (2014). In previous studies, researchers used the GIPL1 model. It will be useful to mention how the GIPL1 is different from GryoGrid1.

We have inserted an additional paragraph in the Introduction, which refers to GIPL1 and the two mentioned studies.

Section 2.4. It is not clear why authors used ERA-Interim and not MERRA-Land or CRUNCEP or any other reanalysis dataset.

Our analysis was performed with ERA-interim, and we have not tested any other reanalysis data set, which we consider beyond the scope of the presented study. We have added an additional sentence on ERA-interim to Sect. 2.3, and paragraphs discussing the uncertainty of ERA-interim to Sect. 4.1.2.

Section 2.5. Authors mentioned the importance of snow and used snow fall (SF) parameter in Table 1 as a turning coefficient for the input parameters. It is not clear, what are the units for SF? Is it normalized or non-dimensional?

It is in m of water equivalent, we have added the information to the caption of the table.

It looks to me that SF is the major coefficient affecting the range in SD shown on Figure 2. If so, then some areas in the mountains could accumulate a lot of snow. Does that mean that we

should expect higher uncertainty in certain regions? Adding more background on where we should expect high/low uncertainties in SD will be useful.

In areas with high average snow depth (i.e. high snow fall), the spatial variability of winter ground surface temperatures can indeed be expected to be higher if one assumes that wind redistribution leads to a pattern of bare-blown spots and snow drifts. At the same time it also plays a role how cold the winters are, i.e. how many FDD are accumulated. We have inserted a new paragraph in Sect. 2.7, which explains the motivation for the different ranges of nf and rk.

Authors mentioned that input parameters in Table 1 are drawn from previous studies. I suggest expanding this by adding more background information on how and why the ranges for nt and rk have been selected. How are the ranges from Table 1 applicable to other geographic regions?

A new paragraphhas been inserted in Sect. 2.7, which explains the motivation for the different ranges of nf and rk. From the study area, there are only few published values for nf and rk, but more background information is provided in Sect. 2.7 of the revised version. It must remain unclear in how far the ranges are applicable to other regions and this issue should be investigated further. While the provided ranges give a satisfactory match with borehole temperatures in the study region, it is not unlikely that other ranges must be assumed e.g. for more continental climate conditions. In Sect. 4.3, we state that "it may become necessary to refine the parameterizations for the ranges of nf and rk, possibly by introducing functional dependences on other environmental variables."

P758. L24. More background is needed on how the cloudy scenes were identified (manually or by using some sort of algorithm)?

Cloudy regions are automatically masked out by the MODIS cloud detection, so that such data are not contained in the employed MODIS LST products. An explanatory sentence has been inserted: "Cloudy regions are automatically detected and removed by the MODIS cloud mask (Frey et al., 2008)".

P 762. L 5. Why is assumed to be equal to 1?

MODIS LST provides the radiative skin temperature, which can be considered equal to the "ground surface temperature" if a dense canopy is lacking. In densely forested areas, one should investigate if there is a "surface offset" between radiative skin and ground surface temperature in long-term FDD and TDD. However, such conditions are not typical for the permafrost domain in the study area. We have inserted: "In case of a dense canopy, where MODIS LST may rather represent top-of-canopy instead of surface temperatures, a different nt may be required, but such conditions rarely occur in permafrost areas in the study region."

Figure 2. Typically TSP temperatures are measured at depth of 20 m. The TTOP calculates temperatures at the top of the permafrost table. It is not clear how did authors compared

those two temperatures? If temperatures calculated by GryoGrid1 get extrapolated to the deeper depth then that has to be mentioned in the manuscript.

This is a very valid point, and we are grateful to the reviewer for pointing this out. The borehole temperatures compiled in International Permafrost Association (2010) are extremely inhomogeneous and provided different depths, as well as for different time periods or points in time. We have used the single one value for the ground temperature provided in International Permafrost Association (2010) for comparison, without any interpolation to a specific depth. However, we emphasize that CryoGrid 1 delivers "Mean Annual Ground Temperatures", which is a hypothetical model temperature that is attributed to some depth at the top of the permafrost (Appendix A).

The mismatch/"error" of the in-situ measurements is estimated as follows: assuming steady-state conditions as in the derivation of the TTOP equation, the vertical temperature gradient $\Delta T/dz$ is equal to $-F/K$, where F is the geothermal heat flux and K the thermal conductivity of the ground. For typical values of $F=-50mW/m^2$ and K=2W/mK, the gradient is 0.25K/10m, so a temperature provided from a depth of e.g. 20m may be around 0.5K warmer than at the top of the permafrost at e.g. 0.5m depth. However, ground temperatures in many boreholes have also been warming in the past decade (Romanovsky et al., 2010), and a borehole temperature provided for e.g. 20 m depth may be strongly influenced by climate conditions before 2002, i.e. the time before satellite measurements of LST are available.

With the presented equilibrium approach, this problem cannot be solved. However, we argue that the direction of the error is not necessarily systematic, and generally smaller than the accuracy of 2.5K. We have A) clarified in Sect. 3.1 what was done in the comparison, and B) inserted a new subsection 4.1.1 discussing the above issue. A part of 4.2 has been moved to this new subsection.

In the conclusion section, I suggest to add a paragraph describing how this regional study can be carried over globally? What researchers need to know when use GryoGrid1 in Mongolia or Russia?

An additional statement has been inserted in the Conclusion section: "…, but uncertainties related to satellite-based land cover maps and associated ranges of model parameters must be investigated and resolved prior to global application."

Adding one more figure with calculated uncertainties (SD corresponding to Figure 4) will be extremely useful as well as describing the uncertainty map.

A new Figure 5 visualizing the standard deviation of all model realizations has been inserted, and it is explained in detail in a new paragraph under 3.1.

References

Panda SK, Marchenko SS, Romanovsky VE. 2014. High-resolution permafrost modeling in Denali National Park and Preserve. :1-46.Natural Resource Technical Report

NPS/CAKN/NRTR-2014/858. National Park Service, Fort Collins, Colorado. [https://irma.nps.gov/App/Reference/Profile/2208990]

Luo D L, Jin H J, Marchenko S, Romanovsky V. 2014. Distribution and changes of active layer thickness (ALT) and soil temperature (TTOP) in the source area of the Yellow River using the GIPL model Science China Earth Science, 09/2014; 57(8):1834–1845. DOI: 10.1007/s11430-014-4852-1

Anonymous Referee #2

Westermann et al. present a modelling scheme based on MODIS acquired LST and ERA-Interim reanalysis products in order to derive MAGST at 1 km at a continental scale. Subgrid variability is addressed by computing a distribution of MAGST's using a simple equilibrium model (CryoGrid 1) for each grid-cell. The approach is applied to approximately 5 million km2 in the North-Atlantic region. The approach is evaluated against a network of 143 boreholes (IPA 2010) and suggests a model accuracy > 2.5degC. The probabilistic approach allows a classification of each grid cell as continuous, discontinuous and sporadic permafrost.

In general I found the manuscript to be an interesting approach (and enjoyable read) that could be an important contribution to large area permafrost modelling and would recommend publishing after consideration of the following suggestions and comments.

MAIN COMMENTS

1. TOPOGRAPHY AND SURFACE COVER

The scheme accounts for subgrid variability by computing a range of physically plausible nf (roughly the offset between GST and surface T generally caused by snowpack) and rk (roughly the offset between GST and temperature at depth) parameter values. You mention rk to be primarily dependent upon water content of the soil but also of significance is the thermal properties of the surface cover itself – blocky, vegetation, fine grained material etc. Is this somehow considered in computing variability?

We agree, this would be important. At this stage, however, it can in practice not be included in such modeling approaches, since reliable global maps/ classifications of such phenomena are lacking. Therefore, we restricted ourselves to only two simple classes.

We have inserted a paragraph under Sect. 2.7, which clarifies such restrictions: "The rk values are chosen close to 1 for both classes, which implies that large changes in the thermal conductivities do not occur. Areas with a potentially large thermal offsets (i.e. low rk values), such as wetlands or extensive block field areas (Gruber and Haeberli, 2007), are clearly not represented by this choice. However, with the currently available land cover products, it is not possible to reliably detect such areas (see Sect. 2.7), so that they cannot be accounted for by the employed scheme.

Another very important source of subgrid variability, even at the 1 km scale is topography, particularly elevation and aspect (although reduced influence in northern climates). It would be nice if this could somehow be considered in some kind of ruggedness index as obviously this effect on uncertainty is far greater in a steeper region of coastal Greenland compared to flat regions of the Russian Arctic.

We also agree with this comment, but there are is a conceptual problem with including topography and exposition: The applied FDD and TDD is a composite of MODIS LST, i.e. LST values reflecting the true thermal surface temperature conditions of a certain pixel, and downscaled ERA reanalysis. For the latter, including topographical variations would be straight-forward, as the employed DEM provides a standard deviation of altitudes for each pixel. For the LST measurements, however, there is no accepted procedure to obtain the variation within a pixel. The signal received at the sensor is an average one, which is a result of the distributions of exposition, altitude, shaded areas, surface albedo, etc. within the footprint area. Simply applying an atmospheric lapse rate to estimate the spread would not do justice to these many factors, especially considering that the measurements are taken during clear-sky conditions. Moreover, in really steep topography, the satellite footprints can vary strongly from scene to scene depending on the respective satellite orientation and view angle, so that MODIS LST is of limited use in high mountain areas, such as the European Alps.

Accounting for the topographic variations only for the ERA reanalysis-derived values is conceptually not satisfactory, since their fraction is highly variable throughout the study area (Fig. 1), and the "weight" of the topographic variations would basically depend on the abundance of clear-sky conditions.

We have therefore chosen not to include topographic variations as a factor in the model approach and rather clearly state this as a limitation. We believe that the presented modeling cannot be more than a first-order approximation of the permafrost distribution in high mountain areas with strong topographic variations and that its main potential lies in areas with rather gentle topography. However, this is the case in the large part of the study area. Of all grid cells with modeled negative mean ground temperatures, only about 2-3% feature a standard deviation of the altitudes larger than 75m (according to the employed DEM), and even in countries like Norway, the majority of the permafrost occurs in high-lying plains rather than in jagged mountain areas.

We have included the following statement within Sect. 2.7: "Furthermore, variations of altitude and exposition within grid cells are an additional source of spatial variability of ground temperatures which is not accounted for in the presented model scheme. In mountain areas with strong topographic variations, spatial variability of ground temperatures within a grid cell is most likely underestimated."

Furthermore, we added a statement to the discussion in Sect. 4.3: "Furthermore, in mountain areas with strong topographic variations, satellite-based LST measurements at 1 km scale cannot sufficiently capture variations of altitude and exposition, so that

the modeling scheme at best can be expected to deliver a first-order approximation of the permafrost distribution."

2. UNCERTAINTY OF ERA-INTERIM AIR TEMPERATURE FIELD

You address some of the downscaling issues with using a coarse scale product such as ERA-Interim and I agree with the simple approach for your application. However, you don't discuss the spatial variability of bias in a reanalysis product such as ERA-Interim which rely on stations / upper air measurements to constrain the weather model. Such variability could lead to strong differences in regional patterns of bias. Some references (if available) on the performance of ERA-Interim in the North could be useful. You mention the finding that model levels below grid level do not yield as good results and in general this was found to be related to poor representation of the surface boundary layer. One significant effect of this that is worth mentioning is that valley inversions are not captured in the temperature field.

In the revised version, we refer to two studies, which have independently validated the performance of the ERA-interim analysis for areas adjacent to our study region. However, these are only unsystematic snapshots, and we are not aware of a conclusive study on regional biases of near-surface air temperatures in the study region. Furthermore, we have added a statement on the limitations of the downscaling procedure, and possible ways to overcome such differences: "We emphasize that the procedure cannot account for many regional and local climate and weather conditions, such as temperature inversions in valley systems. Downscaling of reanalysis data using e.g. the Weather Research and Forecasting (WRF) Model, though computationally expensive, may be a way to overcome such difficulties (Aas et al., 2015)."

Ref: Schanke Aas, K., Berntsen, T., Boike, J., Etzelmüller, B., Kristjánsson, J., Maturilli, M., Schuler, T., Stordal, F., Westermann, S.: A comparison between simulated and observed surface energy balance at the Svalbard archipelago, Journal of Applied Meteorology and Climatology, in print, doi: http://dx.doi.org/10.1175/JAMC-D-14-0080.1, 2015.

3. UNCERTAINTY OF ERA-INTERIM SNOW DEPTH

There are well documented biases in reanalysis precipitation fields (eg Schmidli et al 2006) and such coarse resolution of ERA-Interim will never capture the variability of snowdepth in complex topography which in turn has a large effect on spatial patterns of the ground thermal regime. I think these possible sources of bias should be mentioned in the scheme evaluation. Also, how are snowdepths at each grid cell used to determine the range of nf factors? I don't see this described anywhere.

We have included the reference in 2.5, and mentioned documented biases. As detailed in 2.5, the aim of using ERA snowfall is to capture large-scale precipitation patterns, and adjust the ranges of nf-factors, so that the true range is contained within. This means that in an area with high snowfall, such as the W coast of Norway, the minimum nf should be significantly lower compared to an area with low snowfall, such as N

Greenland. In the revised version, this is explained in more detail in Sect. 2.7, where we describe how nf is determined from ERA snowfall.

4. UNCERTAINTY OF MODIS LST

You mention in Section 2.2 the seasonal average cold biases of up to 3K and therefore question the reliability of MODIS. To overcome this problem you say you have a composite product of reanalysis and MODIS. However, the reanalysis is used to gap fill cloudy days. Does this mean the 3K bias is due to cloudiness or some other effects? It would be good to be clear how the composite directly addresses the bias in MODIS and is not just a gap filling strategy.

Yes, the bias is mainly due to cloudiness, as shown in Westermann et al., (2012) and Østby et al. (2014). We have modified the respective passage to better explain this: "This is in particular attributed to prolonged cloudy periods with systematically different average surface temperatures that are not captured by the satellite measurements. Furthermore, detection of clouds is imperfect in particular during polar night conditions (Liu et al., 2004), so that cloud top temperatures are contained in the MODIS LST time series. Validation studies have shown that these effects can lead to a systematic cold-bias of up to 3 K in seasonal averages (Westermann et al., 2012; Østby et al., 2014).

6. CLOUDINESS

How is cloudiness detected and subsequently MODIS scenes rejected? Some details on methods, thresholds used would be useful.

Cloudy regions are automatically masked out by the MODIS cloud detection, so that such data are not contained in the employed MODIS LST products. An explanatory sentence has been inserted: "Cloudy regions are automatically detected and removed by the MODIS cloud mask (Frey et al., 2008)".

7. BIAS

There appears to be a cold bias in "North America" (Figure 3) compared to "Nordic" and "Russia". Do you have any suggestions why?

It is already stated in the manuscript that there is a cold-bias for the boreholes in N America. Possible reasons are: 1) the comparatively small number of boreholes located only at a few different sites, which makes it an unrepresentative selection; this would at least leave the possibility open that the bias in reality does not exist; 2) all boreholes are located in a high-Arctic setting, where the performance of the employed reanalysis may be less good – a similar bias may well occur e.g. in N Greenland, where there are no boreholes; 3) the employed parameterizations in particular for nf may be biased for the high-Arctic setting.

Since it is not entirely clear if there really is such a regional bias, we do not wish to speculate about the potential reasons. We have inserted a sentence in Sect. 3.1: "…,

while there is a slight cold-bias for the available boreholes in N America. However, the latter cannot be fully secured due to the comparatively weak data basis of only 12 boreholes, all of which are assigned large model standard deviations of 1.9 to 2.7 degree C (see Supplement) which indicate a potentially large spatial variability of ground temperatures."

8. VISUALIZATION OF UNCERTAINTY

Linking back to comment 1, among other factors – key sources of uncertainty would be related to ruggedness of topography in a grid cell. Would it be possible to devise an uncertainty map based around a topographic index?

As outlined above, this cannot be done in a satisfactory way since the footprint of a MODIS satellite pixel is to some extent variable, which causes problems exactly in areas with high roughness. The uncertainty will depend on the exact nature of this footprint area.

Another figure I would like to see would be a spatial representation of the variability of modelled values for each grid-cell. This would again give some insight into regional patterns of uncertainty at least with respect to the perturbed model variables. You could also overlay bias point values on this map from the borehole-modelled results.

A new Figure 5 visualizing the standard deviation of all model realizations has been inserted.

TECHNICAL COMMENTS

1. *P754 l6: on continental scale > at continental scales.*

Done

2. *P755 l16:18: reads a bit like this sentence has just been thrown in there without too much context.*

Changed to "In order to improve such shortcomings, Gruber (2012) recently derived a global high-resolution data set…"

3. *P755 l25:26: which SWE product do you refer too and why do you not use instead of ERA-Interimsnow-depth.*

In the Langer et al.-study, GlobSnow SWE at a spatial scale of 25km is used, which has a significantly higher resolution compared to ERA-interim snow depth. We have inserted more information in the text.

4. *P756 l5 Schmid et al. 2012 could be a good reference here too.*

Reference inserted

5. *P756 l18: on continental scale > at continental scales.*

Done

6. *P756 l18: factor > variable?*

We agree, done

7. *Section 2.3: why was ERA-Interim chosen?*

Our analysis was performed with ERA-interim, and we have not tested any other reanalysis data set, which we consider beyond the scope of the presented study. We have added an additional sentence on ERA-interim to Sect. 2.3, and paragraphs discussing the uncertainty of ERA-interim to Sect. 4.1.2.

8. *You mention SST and upper-air soundings but surface stations are a significant source of data.*

Inserted

9. *P762 l6 round > ground*

Done, thanks!

10. *P762 l19 remove 'a'*

Done, thanks!

11. *P765 l10 "regime on <a> continental scale"*

done

12. *P769 l8 a proxy of temperature is measured and converted to LST via an algorithm*

13. *P769 l7:8 You mention that the strength of LST is to provide an actual measurement – perhaps its worth moderating that statement as there is a lot of algorithm going into deriving the final LST value. In that sense you could argue MODIS LST is closer to a modelled value. This separation is of course not black and white - but I think you would not classify MODIS LST similarly to a sensor directly at the ground surface.*

We agree, this separation is not black and white. However, we argue that satellite LST measurements are significantly closer to surface observations than to a coarse-scale atmospheric model (even when downscaled). The main difference is that reanalyses use the implemented model physics to compute e.g. skin or air temperatures from sparse observational data sets that are generally acquired FAR from the point of interest. However, the MODIS LST algorithm relies on an observational data set (i.e. received radiances in a whole range of spectral bands) DIRECTLY from the point of interest. This implies that only information taken at or at least close to the actual spatial and temporal resolution is employed to derive the LST value. So, if there is e.g. a forest fire or a volcanic eruption, the satellite measurement will see the increased temperature, but

the model will not, since neither the fire nor the volcano are part of its model physics. If a large Arctic river adverts heat and warms a narrow area adjacent to its banks, the satellite measurements will detect it, but the model will not. If there is a temperature inversion in a clear-sky night in a 2km wide valley system, it is clearly visible in satellite product, but not in a coarse-scale model. We agree that the underlying set of algorithms is complicated and prone to error, but many state-of-the-art measurement systems and devices (eddy covariance systems, for instance) share this characteristics.

We have reformulated the statement to be more neutral: "This approach combines the strengths of both products - the capacity of satellite sensors to provide actual measurements of the footprint area at a high spatial resolution, and the dense, gap-free time series of the reanalysis product."

REFERENCES

Schmidli, Jürg, Christoph Frei, and Pier Luigi Vidale. "Downscaling from GCM Precipitation: A Benchmark for Dynamical and Statistical Downscaling Methods." International Journal of Climatology 26, no. 5 (April 1, 2006): 679–89. doi:10.1002/joc.1287.

Schmid, M.-O., S. Gubler, J. Fiddes, and S. Gruber. "Inferring Snowpack Ripening and Melt-out from Distributed Measurements of near-Surface Ground Temperatures." The Cryosphere 6, no. 5 (October 2012): 1127–39. doi:10.5194/tc-6-1127-2012.

Changes to the manuscript:

1. Sect. 2.2: A new paragraph has been inserted that provies information on the reasons for the cold-bias in MODIS LST averages.
2. Sect. 2.7: More information on the choice for the ranges of nf and rk provided in the Table are provided .
3. Sect. 3.2: A new Figure 5 displaying the standard deviation of all model realizations has been inserted, along with a new paragraph describing it.
4. Sect 4.1: A new Sect. 4.1.1 on the equilibrium model approach has been inserted. This section in particular addresses the problem of comparing the results of such an equilibrium approach to real-world measurements, which are taken at a certain depth and time.
5. Sect. 4.1.2: A paragraph emphasizing on the uncertainties of ERA interim model results is provided, as well as more details on how such shortcomings might be resoled in the future.
6. Appendix: The derivation of Eq. 1 for MAGT has been corrected. The way the variables were defined in the previous version, it is incorrect to set MAGT=T_{TOP}, since the latter refers to the winter temperature at the bottom of the active layer. MAGT is the weighted average of winter and summer temperatures at the bottom of the active layer. This has been corrected by modifying Eq. A1, and adding Eq. A2.

Manuscript prepared for The Cryosphere Discuss.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 7 May 2015

A ground temperature map of the North Atlantic permafrost region based on remote sensing and reanalysis data

S. Westermann[1,2], T. Østby[1], K. Gisnås[1], T. V. Schuler[1], and B. Etzelmüller[1]

[1]Department of Geosciences, University of Oslo, P.O. Box 1047, Blindern, 0316 Oslo, Norway
[2]Center for Permafrost (CENPERM), Department of Geosciences and Natural Resource Management, University of Copenhagen, Øster Voldgade 10, 1350 Copenhagen K., Denmark

Correspondence to: S. Westermann (sebastian.westermann@geo.uio.no)

Abstract

Permafrost is a key element of the terrestrial cryosphere which makes mapping and monitoring of its state variables an imperative task. We present a modeling scheme based on remotely sensed land surface temperatures and reanalysis products from which mean annual ground temperatures (MAGT) can be derived at a spatial resolution of 1 km at continental scales. The approach explicitly accounts for the uncertainty due to unknown input parameters and their spatial variability at subgrid scale by delivering a range of MAGTs for each grid cell. This is achieved by a simple equilibrium model with only few input parameters which for each grid cell allows scanning the range of possible results by running many realizations with different parameters. The approach is applied to the unglacierized land areas in the North Atlantic region, an area of more than 5 million km^2 ranging from the Ural mountains in the East to the Canadian Archipelago in the West. A comparison to in-situ temperature measurements in 143 boreholes suggests a model accuracy better than 2.5 °C, with 139 considered boreholes within this margin. The statistical approach with a large number of realizations facilitates estimating the probability of permafrost occurrence within a grid cell so that each grid cell can be classified as continuous, discontinuous and sporadic permafrost. At its southern margin in Scandinavia and Russia, the transition zone between permafrost and permafrost-free areas extends over several hundred km width with gradually decreasing permafrost probabilities. The study exemplifies the unexploited potential of remotely sensed data sets in permafrost mapping if they are employed in multi-sensor multi-source data fusion approaches.

1 Introduction

Permafrost shapes approximately a quarter of the landmass of the Northern Hemisphere (Brown et al., 1997) and is thus one of the largest elements of the terrestrial cryosphere. Permafrost occurs mainly in arctic regions which will be strongly impacted in the future warmer climate. The recent and predicted future warming of the global climate may lead

to widespread thawing which can trigger climatic feedbacks from local to global scale and severely impact ecosystems, infrastructure and communities in the Arctic. Thawing organic-rich permafrost is projected to spark substantial emissions of the greenhouse gases CO_2, CH_4 and N_2O (Walter et al., 2006; Schuur et al., 2008; Elberling et al., 2010, 2013) which may be relevant for future climate projections and hereof derived mitigation strategies (e.g. Schaefer et al., 2011; Schneider von Deimling et al., 2012). As a consequence, considerable efforts and resources are currently dedicated to improving the representation of permafrost processes in the General Circulation Models (GCMs) for the Intergovernmental Panel on Climate Change (IPCC) assessment. Due to the large thermal inertia of the frozen ground and associated long time periods required for thawing, future projections critically depend on the knowledge of current thermal ground conditions. However, the more than 15 years old, but still widely used global permafrost map from the International Permafrost Association (Brown et al., 1997) features a very coarse scale and lacks quantitative information on permafrost state variables, in particular of ground temperatures. In order to improve such shortcomings, Gruber (2012) recently derived a global high-resolution data set of permafrost probabilities based on downscaled air temperatures from reanalysis data.

Permafrost is a thermally defined subsurface phenomenon. In contrast to e.g. glaciers, permafrost cannot be "seen" at the surface by its spectral signature and hence not directly monitored by satellite sensors. Therefore, monitoring and mapping of permafrost is restricted to either direct point observations or coarse-scale modeling using atmospheric circulation models. However, there exist a variety of satellite data sets which can be employed in numerical permafrost models to compute the thermal state of the ground, in particular Land Surface Temperature (LST) and Snow Water Equivalent (SWE). In a first proof-of-concept study, Langer et al. (2013) demonstrated the potential of such concepts by forcing a transient ground thermal model at a spatial scale of 1 km by time series of MODIS LST and downscaled GlobSnow SWE (Takala et al., 2011). While the agreement with measured in-situ data of ground temperatures and thaw depth is striking, the study also highlights the considerable sensitivity of physically-based approaches to model parameters, which are generally not known for very large spatial domains, e.g. an entire continent. On the

other hand, field studies suggest a considerable variability of ground temperatures at spatial scales much smaller than 1 km (e.g. Schmid et al., 2012; Gisnås et al., 2014), which questions the validity of a single model scenario per grid cell. Multi-scenario-runs with e.g. different snow depths and ground thermal properties (similar to tiling approaches employed in GCMs) would be an elegant way to represent the small-scale variability in a statistical way, but this implies a significant increase in the computational demands if all combinations of model parameters should be explored. **In such a context, simple schemes with few parameters, such as the modified Kudryavtsev (Sazonova and Romanovsky, 2003) or the TTOP approach (Smith and Riseborough, 1996) can offer advantages over more sophisticated models. They have been developed to a state of operational maturity and are employed for large-scale mapping of permafrost properties with a wide variety of input data sets (e.g. Luo et al., 2014; Panda et al., 2014).**

In this study, we employ a simple semi-empirical equilibrium model (CryoGrid 1, based on the TTOP approach, Gisnås et al., 2013) with a small number of model parameters in a statistical framework which is computationally efficient enough to be applied for large spatial domains. We demonstrate high-resolution statistical mapping of ground temperatures based on a combination of remotely sensed land surface temperatures and reanalysis data. The model results are benchmarked against measurements of ground temperatures in boreholes from the "Thermal State of Permafrost" (TSP) data base (Romanovsky et al., 2010a; International Permafrost Association, 2010). The results suggest that mapping of ground temperatures at continental scales is feasible, a variable highly demanded by the scientific community, as well as policymakers and industries in affected countries.

2 Modeling tools and satellite data sets

2.1 Study region and in-situ data

The study focuses on the permafrost areas bordering the North Atlantic Ocean which comprise a large gradient of ground temperatures, from permafrost-free areas in Scandinavia

to some of the coldest permafrost on Earth in N Greenland and Ellesmere Island, Canada. In detail, the region ranges from the Ural mountains, Novaja Zemlja and Franz Josef Land in the East over Scandinavia, Svalbard, Iceland and Greenland to the Eastern parts of the Canadian Archipelago. The area contains a wide range of permafrost environments, from high-relief mountain ranges to wetlands and subarctic mires underlain by permafrost, which makes it a well-suited test region for a modeling scheme. In-situ monitoring programs of ground temperatures in boreholes (Christiansen et al., 2010; Romanovsky et al., 2010b; Smith et al., 2010) have created a excellent record to which the model results can be compared. In total, we employ 143 borehole temperatures in N America, the Nordic region and W Russia (see Fig. 1) from the "Thermal State of Permafrost (TSP) Snapshot Borehole Inventory" (International Permafrost Association, 2010). In addition, regional-scale modeling studies have compiled maps of ground temperatures and permafrost distribution at spatial scales of 1 km (e.g. Gisnås et al., 2013; Westermann et al., 2013, 2015). The study is based on data sets from the ten-year period 2003 to 2012, in which most of the validation data sets have been compiled.

2.2 MODIS LST

The "Moderate Resolution Imaging Spectroradiometer" (MODIS) operationally delivers remotely sensed land surface temperatures on global scale at a spatial resolution of 1 km. The MODIS sensor is carried on board of two satellites of NASA's "Earth Observing System", Terra (launched in 2000) and Aqua (launched in 2002). As such, a time series of more than ten years is available for the combined Terra/Aqua data set. For this study, we make use of the level 3 products MOD11A1 (Terra) and MYD11A1 (Aqua) in the version 5. They contain a day and a night value for LST each, so that four values per day are available, thus in principle capturing the diurnal temperature cycle. **Cloudy regions are automatically detected and removed by the MODIS cloud mask (Frey et al., 2008).** In practice, frequent cloudiness leads to a drastic reduction of the data density with prolonged periods without measurements. To cover the study region, in total 16 MODISL LST tiles with an area of $1200\,\mathrm{km} \times 1200\,\mathrm{km}$ each were employed (h15v01, h15v02, h15v03,

h16v00, h16v01, h16v02, h17v00, h17v01, h17v02, h18v00, h18v01, h18v02, h18v03, h19v01, h19v01, h19v02, h20v02). For each of the tiles, a ten year period from 2003 to 2012 was evaluated, so that more than 116 000 individual files were processed.

While the target accuracy of individual LST measurements is 1 K (Wan et al., 2002), several studies from arctic region suggest a significantly reduced accuracy both for individual measurements and hereof derived long-term LST averages (Westermann et al., 2011, 2012; Østby et al., 2014). **This is in particular attributed to prolonged cloudy periods with systematically different average surface temperatures that are not captured by the satellite measurements. Furthermore, detection of clouds is imperfect in particular during polar night conditions (Liu et al., 2004), so that cloud top temperatures are contained in the MODIS LST time series. Valdiation studies have shown that these effects can lead to a systematic cold-bias of up to 3 K in seasonal averages (Westermann et al., 2012; Østby et al., 2014).** It is therefore questionable if MODSIS LST alone can be used to model ground temperatures. To overcome this problem, we synthesize a composite of MODIS LST and ERA–interim reanalysis data as input for the ground thermal model CryoGrid 1 (see Sect. 2.4).

As exemplified by Østby et al. (2014) for the MODIS LST tile h18v00, the land mask employed in the MODIS LST production algorithm is shifted by about 5 km for the tiles N of 80° N, so that land surface temperatures are produced for sea areas, while they are missing for some land areas. In the study region, this mainly affects the N parts of Svalbard, Greenland and Ellesmere Island.

2.3 Downscaled near-surface air temperatures from ERA–interim reanalysis

Reanalysis products provide global datasets of meteorological variables based on atmospheric models in which a range of observations, such as meteorological surface observations, remotely sensed sea surface temperatures, or vertical temperature and humidity profiles from radio soundings, are assimilated. In this study, we make use of air temperatures of the ERA–interim reanalysis (Simmons et al., 2007; Dee et al., 2011), which is available from 1979 until now at a spatial resolution of $0.75° \times 0.75°$. **ERA–interim is the state-of-the-art**

reanalysis product from the European Centre for Medium-Range Weather Forecast (ECMWF), and it has been successfully employed in previous permafrost modeling studies (Fiddes et al., 2015). While the atmospheric model can provide a gap-free time series with four values per day (at 00:00, 06:00, 12:00, and 18:00 UTC), the coarse spatial resolution requires a down-scaling scheme to correct for the difference between the true altitude of a point and the coarse-scale ERA orography, which in the study region can be as much as 1500 m in mountainous terrain. To obtain a downscaled air temperature for a point in space and time, Fiddes and Gruber (2014) propose a four-dimensional interpolation between different reanalysis grid cells, pressure levels and time steps. Since only long-term freezing and thawing degree days are required as input for CryoGrid 1 (Sect. 2.7), a simplified version of this scheme (similar to Gruber, 2012) is employed:

1. An average atmospheric lapse rate is derived for each day of year (DOY) and MODIS grid cell, based on ERA–interim surface fields and the 700 mbar pressure level, which corresponds to an altitude of approximately 3000 m, slightly higher than the highest non-glacierized areas within the study region. For each DOY, all available ERA–interim data from 2003 to 2012 are employed.

2. The downscaled air temperature is computed from the lapse rate and the difference between the ERA orography (interpolated to the center position of a MODIS grid cell) and the "true" altitude.

The "true" altitude of the 1 km MODIS LST grid cell is derived by interpolation of the Global Multi–resolution Terrain Elevation Data 2010 (GMTED2010) at 30 arcsec resolution (Danielson and Gesch, 2011) to the center position of each MODIS 1 km grid cell. This data set is specifically compiled for continental-scale applications and includes the best available global elevation data. This procedure ensures that the spatial and seasonal variability of the lapse rate is accounted for so that e.g. inversions of the air temperature during winter in continental regions can be reproduced.

2.4 Data fusion of MODIS LST and ERA-interim

Due to frequent cloudiness, the MODIS LST time series contains a large number of gaps and only clear-sky LST averages can be derived from remote sensing products alone. To eliminate or at least moderate the effect of clouds, the MODIS LST time series is merged with the gap-free time series of downscaled air temperatures from the ERA reanalysis (see Sect. 2.3). For this purpose, the acquisition time of each MODIS day and night LST measurement (which varies by several hours from day to day) is converted to UTC, and the ERA value closest in time to each LST measurement subsequently removed. From the resulting time series, daily averages are computed from which freezing and thawing degrees are derived. Depending on the cloudiness, the fraction of MODIS LST measurements in the composite MODIS/ERA data set is variable (Fig. 1). In areas with frequent cloudiness, such as Iceland, less than a quarter of the data points are derived from MODIS LST, while the fraction of MODIS LST is more than 50 % in other areas.

2.5 ERA snowfall

The average yearly snowfall from 2003 to 2012 was determined for each grid cell by interpolation of the snowfall surface field coarse-scale ERA–interim reanalysis data. These snowfall data are employed to determine the range of the winter n_f factors for the statistical modeling (in conjunction with a remotely sensed land cover data, see below). **While there are well documented biases in the precipitation fields of the reanalysis (e.g. Schmidli et al., 2006)**, it captures large-scale precipitation patterns, such as the high snowfall in the S Norwegian mountains close to the Atlantic Ocean. On the other hand, it cannot account for small-scale variations of winter precipitation. However, since only the range of possible n_f factors is assigned in a statistical modeling framework with many different model realizations (Sect. 2.7), utilizing large-scale precipitation patterns may indeed be adequate.

2.6 MODIS land cover

The land cover has strong implications for the small-scale distribution of the snow cover: in areas with low vegetation or bare ground, redistribution of snow due to wind drift can lead to a strong spatial variability of snow depths and ground temperatures (e.g. Gisnås et al., 2014). In contrast, a more uniform snow depth is expected in forested areas. To represent such effects in the modeling (Sect. 2.7), the MODIS land cover product MCD12Q1 for the year 2006 at a spatial resolution of 500 m is employed. Due to the poor performance of the landcover classification in some parts of the study region (see Sect. 4.1), the 500 m grid cells are aggregated to the 1000 m MODIS LST grid and the land cover classes merged to two units: high vegetation/forest ("International Geosphere–Biosphere Programme" – IGBP classes 1 to 9, 11) and bare ground/low vegetation (IGBP classes 10, 15, 16). The classes 12 to 14 (cropland/urban) occur only sparsely in permafrost areas and are treated like the high vegetation class.

2.7 Statistical modeling with CryoGrid 1

CryoGrid 1 is an equilibrium model for ground temperatures (Gisnås et al., 2013) based on the TTOP approach (e.g. Smith and Riseborough, 1996). It computes mean annual ground temperature (MAGT) based on thawing and freezing degree days ($\mathrm{TDD_s}$ and $\mathrm{FDD_s}$) of surface temperatures, according to

$$\mathrm{MAGT} = \begin{cases} \frac{1}{\tau}\left(n_\mathrm{f}\,\mathrm{FDD_s} + r_k\, n_t\,\mathrm{TDD_s}\right) & \text{for} \quad \mathrm{MAGT} \leq 0 \\ \frac{1}{\tau}\left(n_\mathrm{f}\,\mathrm{FDD_s} + \frac{1}{r_k}\, n_t\,\mathrm{TDD_s}\right) & \text{for} \quad \mathrm{MAGT} > 0 \end{cases} \tag{1}$$

where τ is the period for which TDD and FDD are evaluated, while r_k, n_f and n_t are semi-empirical parameters which aim to capture a variety of key processes in a single variable. The winter n_f factor relates the freezing degree days at the surface (here derived from

MODIS/ERA) to the freezing degree days at the ground surface as

$$\mathrm{FDD_{gs}} = n_\mathrm{f}\,\mathrm{FDD_s}\,. \tag{2}$$

It thus captures the insulating effect of the snow cover which is a result of snow depth, the thermal properties of the snow cover, and the ground thermal regime itself. For bare surfaces, this variable is unity, while n_f factors as low as 0.2-0.3 are reported from in-situ measurements (e.g. Gisnås et al., 2013) for high snow cover in Scandinavia. The summer n_t factor is defined in a similar way as

$$\mathrm{TDD_{gs}} = n_t\,\mathrm{TDD_s}\,. \tag{3}$$

If air temperatures are employed as surface forcing $\mathrm{TDD_s}$, variable n_t factors are considered to account for differences in air and ground surface temperatures (e.g. Gisnås et al., 2013). **In this study, we assume remotely sensed LST in conjunction with ERA–interim to be a satisfactory representation of ground surface temperature and set n_t to unity. In case of a dense canopy, where MODIS LST may rather represent top-of-canopy instead of surface temperatures, a different n_t may be required, but such conditions rarely occur in permafrost areas in the study region.** The active layer damping factor r_k gives rise to the thermal offset between average ground surface and ground temperatures (Osterkamp and Romanovsky, 1999). It is related to the average thermal conductivities in the active layer in fully thawed and frozen states (k_t, k_f) as $r_k = k_\mathrm{t}\,/\,k_\mathrm{f}$, so that it is closely related to the water/ice content of the soil. The thermal conductivities of water and ice ($k_\mathrm{w} = 0.45\,\mathrm{W\,m^{-1}\,K^{-1}}$, $k_\mathrm{i} \approx 2.2\,\mathrm{W\,m^{-1}\,K^{-1}}$) confine the range of physically possible values of r_k between one for dry soil or rock and approximately 0.2 for pure water.

The statistical modeling is based on the assumption that $\mathrm{FDD_s}$ and $\mathrm{TDD_s}$ can be exactly determined from the MODIS/ERA-composite, while the two remaining variables in Eq. (1), n_f and r_k are unknown and only confined by physically plausible limits (note that $n_t = 1$ is assumed). For each of the land cover units “high vegetation/forest” and “bare ground/low vegetation” (Sect. 2.6), upper and lower bounds for n_f and r_k are defined (Table 1), oriented at physical constraints and published values from previous studies. However, the upper and

lower bounds of n_f and r_k are also tuning parameters which can be adjusted to achieve the best possible match between model and observations (Sect. 3.1) for each land cover unit. We emphasize that remotely sensed data sets on precipitation or snow depth from which the n_f factors could be estimated (Gisnås et al., 2013) are not available at 1 km scale and the interval of possible values for n_f was chosen according to large-scale ERA–interim snowfall data (Sect. 2.5). **For the "bare ground/low vegetation" class, the upper bound for n_f is set to 1, since bare-blown spots without insulating snow cover can occur. The lower bound decreases with increasing snow fall, i.e. for areas with high snow fall, the spread of possible snow depths and thus n_f-values is larger than for areas with low snowfall. For the mountain areas in Norway, Gisnås et al. (2013) assumed a minimum n_f of 0.3, while the minimum lower bound of n_f is around 0.1 for Norway in this study. However, Gisnås et al. (2013) employed grid-cell averages of snow depths to calculate n_f, while in this study the lower bound of n_f aims to represent the areas with maximum snow depths within each grid cell (e.g. snow drifts). The parameterization of n_f vs. snow depth applied by Gisnås et al. (2014) to represent such small-scale variability of snow depths yields n_f values between 0.1 and 0.2 for snow depths of several meters, which regularly occur in the Norwegian mountain areas, in agreement with the choice in this study. For the "high vegetation/forest" class, both upper and lower bound for the winter n_f-factor decrease linearly with snow fall, while the difference between minimum and maximum n_f is held constant. The r_k values are chosen close to 1 for both classes, which implies that large changes in the thermal conductivities do not occur. Areas with potentially large thermal offsets (i.e. low r_k values), such as wetlands or extensive block field areas (Gruber and Haeberli, 2007), are clearly not represented by this choice. However, with the currently available land cover products, it is not possible to reliably detect such areas (see Sect. 4.1.3), so that they cannot be accounted for by the employed scheme. Furthermore, variations of altitude and exposition within grid cells are an additional source of spatial variability of ground temperatures which is not accounted for in the presented model**

scheme. In mountain areas with strong topographic variations, the spatial variability of ground temperatures within a grid cell is most likely underestimated.

For each grid cell, CryoGrid 1 is run for 30 values of n_f and r_k (in equally spaced intervals from the lower to the upper limit), yielding a total 900 different values for the mean annual ground temperature (MAGT). The statistics of these values is a representation of both the small-scale (i.e. $< 1\,\mathrm{km}$) variability of model parameters and the model uncertainty due to the generally unknown "true" parameter distribution of a grid cell. As final result, we calculate the average of all 900 realizations of MAGT, as well as the standard deviation, and the maximum and minimum MAGT.

3 Results

3.1 Comparison to in-situ data

We compare the mean of all model realizations to in-situ measurements of ground temperatures from the "IPA-IPY Thermal State of Permafrost (TSP) Snapshot Borehole Inventory" (International Permafrost Association, 2010) documented in more detail for N America (Smith et al., 2010), Russia (Romanovsky et al., 2010b), and the Nordic areas (Christiansen et al., 2010). **While this data basis provides only a single value for ground temperatures measured at different depths and different points in time, it is the most extensive compilation of permafrost temperatures available for model validation (see Sect. 4.1.1). For comparison,** we generally select the grid cell closest to the borehole location. For boreholes located close to a larger water body, a nearby grid cell located at least 1 km from the shore is employed to avoid contamination of the MODIS LST record with surface temperatures from the water body. For the boreholes north of 80° N (five boreholes near Alert, Ellesmere Island, Canada), no MODIS LST measurements are available due to the erroneous land mask (Sect. 2.2). We therefore use the closest grid cell featuring MODIS LST measurements E of Alert which is located on land. The results of the comparison are displayed in Fig. 2, with 12 boreholes located in N America (Canadian

Archipelago), 69 in the Nordic area (Scandinavia, Greenland, Svalbard, Iceland) and 62 in Russia. For each borehole, the error bars represent one SD of the results of all model realizations for a grid cell which depends on the ranges of n_f and r_k. For the large majority of the boreholes, the agreement between measured and modeled ground temperatures is better than 2 °C, with 67 (47 %) contained within one, 108 (75 %) within two and 128 (90 %) within three SDs of all model realizations from the mean (see Supplement).

A histogram of deviations between measured and the mean modeled ground temperatures is displayed in Fig. 3. It roughly follows a Gaussian distribution with width of 1.2 °C. For 60 % of the boreholes, the mean of all model realizations is within 1 °C of the measurement, while the agreement is better than 2 °C for 93 % and better than 2.5 °C for 97.5 % of the boreholes. From the comparison with in-situ measurements, we conclude that the accuracy of the modeled ground temperatures is on the order of 2 to 2.5 °C for the study region. There is no significant regional bias for Russia and for the Nordic areas, while there is a slight cold-bias for the available boreholes in N America. **However, the latter cannot be fully secured due to the comparatively weak data basis of only 12 boreholes, all of which are assigned large model standard deviations of 1.9 to 2.7°C (see Supplement) which indicate a potentially large spatial variability of ground temperatures.**

While the borehole data of the TSP network are well suited to validate the large-scale performance of the modeling, only very few quantitative studies on the spatial variability of ground temperatures within areas of a model grid cell exist to which the ensemble of all model realizations could be compared. Recently, Gisnås et al. (2014) presented a one-year data set of ground surface temperature measurements based on 40 to 100 temperature sensors distributed in three areas of approximately 0.5 km^2, located in Svalbard and S Norway. Since the thermal offset between average ground surface and ground temperatures is considered small for the sites, we compare the measured distribution of average ground surface temperatures (AGST) to the ensemble of modeled ground temperatures. For Finse (60°34′ N, 7°32′ E), 30 % of the temperature sensors displayed AGSTs below 0 °C, compared to 51 % of the model relization for the MODIS grid cell comprising the study site (for ground temperatures). The measured range of AGST was -2 to $+2.5$ °C, while minimum

and maximum modeled ground temperatures were −2.2 and +2.1 °C. For the Juvvasshøe site (61°41′ N, 8°23′ E), 76 % of the measured AGSTs were below 0 °C, compared to 77 % of the model realizations. The modeled temperature range (−3.8 to 1.2 °C) was slightly larger than the measured range of AGST (−2 to +1.5 °C). For Ny-Ålesund (78°55′ N, 11°50′ E), the modeled temperatures are slightly cold-biased, with measured borehole temperatures of −2.3 °C (International Permafrost Association, 2010) compared to a mean of all model realizations of −3.8 °C. Nevertheless, the modeled temperature range of 4.9 °C between minimum and maximum corresponds well to the range of measured AGST (−4.5 to +0.5 °C). Finally, the spatial variability of ground temperatures has been modeled for the high-Arctic Zackenberg site in NE Greenland, taking into account the spatial variability of the snow cover, ground surface and ground properties (Westermann et al., 2015). For the period 2002–2012, modeled annual average ground temperatures at 1 m depth range from −9.0 to −3.5 °C, while the model realizations of CryoGrid 1 yield a temperature range from −11.0 to −2.9 °C. While not a validation in a strict sense, the satisfactory to good agreement for these few examples suggests that the ensemble of all model realizations can represent the small-scale spatial variability of ground temperatures, at least for the land cover class "bare ground/low vegetation". Similar field data sets or modeling studies do not exist for the "high vegetation/forest" class, so that is not possible to benchmark the modeled temperature range at this point.

3.2 Modeled ground temperatures in the N Atlantic permafrost region

The model approach facilitates large-scale mapping of the ground thermal regime on a continental scale. As shown in the previous section, the comparison to in-situ measurements suggests that the mean of all model realizations reproduces the ground thermal regime within an accuracy of 2 to 2.5 °C. Figure 4 displays the resulting ground temperature map for the permafrost regions bordering the N Atlantic Ocean. The modeled ground temperatures span a wide range, from −15 °C in N Greenland and Ellesmere Island to more than +5 °C in S Scandinavia. Ground temperatures below 0 °C are modeled in the Canadian Archipelago, Greenland, Iceland, Svalbard, the Scandinavian mountains and the coastal

regions of the Pechora Sea E of the Kanin peninsula. In the high elevations of the Ural mountains, subzero temperatures occur as far south as 60° N. In the lowlands E of the Ural mountains, negative modeled ground temperatures reach significantly farther south compared to the W side. In N Scandinavia and Russia, large areas with modeled ground temperatures just above 0 °C exist, for which it is not possible to establish permafrost presence or absence considering the accuracy of 2 to 2.5 °C accuracy limit (see Sect. 3.3). The modeled permafrost distribution in the N Atlantic region is in good qualitative agreement with the "classic" permafrost map of the IPA based on field evidence (Brown et al., 1997), including prominent features like the asymmetry of the permafrost extent E and W of the Urals.

The standard deviation of all model realizations is displayed in Fig. 5. The highest values of around 3°C are reached in N and E Greenland, as well as on Baffin Island, areas with a large number of freezing degree days that are classified as "bare ground/low vegetation" and hence are assigned a wide range of n_f-values. The large spread therefore stems from the strong variability of winter ground surface temperatures reflected in the term $n_f FDD$ in Eq. 1. Permafrost areas further south, e.g. in Scandinavia, feature lower numbers of FDD and hence a smaller standard deviation, generally between 1 and 2°C. Areas classified as "high vegetation/forest" have standard deviations of 1°C or less, mainly due to the smaller range of winter n_f-factors, which reflect the more homogeneous snow cover in such areas.

At 1 km spatial resolution, it is meaningful to "zoom" into subregions for a detailed assessment of the ground thermal regime. Figure 6 displays a ground temperature map for Greenland and Iceland, with a significantly finer resolution compared to the 25 km scale assessment of Daanen et al. (2011). In the ice-free parts of N Greenland, modeled ground temperatures range from −10 and −15 °C, while they gradually increase southwards in the ice-free regions of NE Greenland to around −3 °C in the Scoresbysound region. Further South, only few ice-free land areas exist on the E coast, but subzero ground temperatures are modeled in the Tasilaq region at 65.5° N. On the W coast, the permafrost extends from the outer coast line to the ice sheet in the extensive ice-free land regions S of Disko Bay,

with modeled ground temperatures warmer than $-5\,^{\circ}\mathrm{C}$. Starting just N of Nuuk, the coastal areas become permafrost-free, but subzero ground temperatures are still modeled at higher elevations in S Greenland.

In Iceland, subzero ground temperatures are restricted to the interior, ranging from the central Sprengisandur between Hofsjökull and Vatnajökull north to the mountain ranges of the Tröllaskagi peninsula, where active rock glaciers suggest a periglacial environment (Lilleøren et al., 2013). The modeled permafrost extent is in good agreement with previous regional estimates based on mean annual air temperatures (Etzelmüller et al., 2007; Farbrot et al., 2007).

On Svalbard, modeled ground temperatures in the ice-free areas range from -2 to $-3\,^{\circ}\mathrm{C}$ at the W coast to -5 to $-6\,^{\circ}\mathrm{C}$ in the ice-free parts of Nordaustlandet. In Scandinavia, the southernmost areas with subzero modeled ground temperatures are located at high elevations in the S Norwegian mountains. Within the accuracy, the modeled ground temperatures agree well with the spatially distributed modeling of Westermann et al. (2013), showing rather warm permafrost with ground temperatures generally warmer than $-3\,^{\circ}\mathrm{C}$. In general, more areas are mapped with subzero temperatures compared to the previous study (especially in the Hardangervidda area), but the differences in the absolute modeled temperatures are mostly less than $1.5\,^{\circ}\mathrm{C}$ (see also Sect. 3.3), and patchy permafrost is known to exist in these regions (Etzelmüller et al., 2003). In N Scandinavia, permafrost is modeled in higher elevations of the Swedish and Norwegian mountains. In the latter, the pattern is again in agreement with the regional assessments of Farbrot et al. (2013) and Gisnås et al. (2013) who employed CryoGrid 1 with gridded air temperature and snow depth products.

In the rolling plains of the Finnmarksvidda, the model shows temperatures just above $0\,^{\circ}\mathrm{C}$, while sporadic permafrost is found in high-lying fell areas and palsa mires (Farbrot et al., 2013). In this area, the model approach clearly fails to reproduce the permafrost patterns. Most likely, a main reason is the employed MODIS landcover product, in which almost the entire area is uniformly classified as “open shrubland” (and thus as “high vegetation/forest” in our approach), although it is characterized by a complex pattern of fell areas, tree-less mires and areas covered by mountain birch forest. If the model is

applied to the area with the parameter set for the “bare ground/low vegetation class” (Table 1), subzero ground temperatures are modeled for a large part of the area. Therefore, the model approach would indeed yield a pattern of permafrost and permafrost-free areas if a landcover product capable of distinguishing mountain birch forest from bare fell and tundra areas is employed. The same issue applies to modeled ground temperatures on the Kola peninsula in Russia, where ground temperatures around 0 °C are modeled for the N and E parts.

3.3 Permafrost probability maps from statistical modeling

With an estimated accuracy of 2 to 2.5 °C, it is impossible to delineate exact boundaries for the permafrost extent of the study region based on the mean of all model realizations. Around the thaw threshold, the statistical model approach features model realizations below and above 0 °C, depending on the input parameters n_f and r_k. However, permafrost and permafrost-free areas coexist in this transition zone also in nature. This long-known fact (e.g. Baranov, 1959) gave rise to the classification system “continuous” ($>$ 90 % of an area underlain by permafrost), “discontinuous” (50–90 %), and sporadic ($<$ 50 %) permafrost, as defined by the IPA long before the advent of sophisticated modeling techniques (Brown et al., 1997).

In contrast to model approaches with only a single model realization for each grid cell, the ensemble of realizations of CryoGrid 1 facilitates deriving a similar permafrost zonation from the modeling. Figure 8 displays the fraction of model realizations with subzero ground temperatures in the same classes as the IPA permafrost map for Scandinavia and W Russia. The potential permafrost extent is larger than in Figs. 4 and 7, with the most significant differences in lowland areas in Scandinavia and Russia. The interior of N Scandinavia (for which the mean of all model realizations is above 0 °C, Sect. 3.2) is classified as sporadic permafrost in this approach, which is in good agreement with observed field evidence for this area. In Russia, sporadic to discontinuous permafrost is modeled from the Kola and Kanin peninsulas to the mouth of the Pechora River, east of which the permafrost is clas-

sified as continuous. Also in this probability map, the individual permafrost zones extend significantly further south on the east side of the Ural Mountains than on the west side.

This asymmetric permafrost distribution around the Ural mountains is a prominent feature in the IPA permafrost map (Brown et al., 1997), which the presented satellite-based modeling approach is capable of reproducing. In contrast, the global map of Gruber (2012), which also delivers comparable probabilities of permafrost occurrence, shows a more or less symmetric permafrost extent around the Urals. Further pronounced differences occur on the eastern tip of the Kola peninsula where the map by Gruber (2012) shows no permafrost, other than the IPA permafrost map and the satellite-based modeling approach. Compared to the IPA map, the transition zone from permafrost-free areas to the continuous permafrost zone in NW Russia is broader in the satellite-based map, which is an indication that some model scenarios (i.e. combinations of model parameters) do not occur in nature. A similar broadening of the transition zones is visible in the probability map of Gruber (2012).

4 Discussion

4.1 Uncertainties and prospects for improvements

4.1.1 The equilibrium model

The TTOP-approach is one of the older modeling schemes conceived to obtain ground temperatures and thus permafrost occurrence from near-surface meteorological variables (e.g. Smith and Riseborough, 1996, 2002). It delivers a "mean annual ground temperature" at the depth corresponding to the top of the permanently frozen ground under the strong assumption that the ground thermal regime is in thermal equilibrium with the applied surface forcing data (Appendix A). Since this assumption is violated to a certain degree for real–world conditions, a comparison to in-situ measurements, as conducted in Sect. 3.1, is not a straight-forward task. For a typi-

cal geothermal gradient of 0.025°C m^{-1}, measured $MAGT$s would be less than 1°C warmer than modeled $MAGT$s for borehole depths of up to 40 m. On the other hand, ground temperatures have been warming by a similar magnitude even at depths of 10 to 20 m in the model period (Romanovsky et al., 2010a), so that recorded borehole temperatures may represent colder climate conditions compared to the model period. With the equilibrium approach, it is thus principally impossible to exactly reproduce ground temperatures that can be measured in the field. However, we note that the combined uncertainty due to measurement depth and transient effects is not necessarily systematic, and presumably well below the estimated accuracy of the modeling scheme of 2 K for most boreholes.

4.1.2 Land surface temperature

In this study, land surface temperatures are derived from satellite measurements during clear-sky conditions, while air temperatures from reanalysis products are employed for cloud-covered periods when satellite measurements are unavailable. This approach combines the strengths of both products – the capacity of satellite sensors to provide actual measurements of the footprint area at a high spatial resolution, and the dense, gap-free time series of the reanalysis product. On the other hand, it strongly reduce the uncertainty associated with both input data sets when used separately: even when downscaled with altitude, reanalysis products only provide a large-scale temperature field which does not account for the heterogeneity of the surface temperature caused by spatially variable surface conditions. Employing only MODIS LST and thus clear-sky LST to calculate long-term averages can result in a serious bias since cloudy periods with a significantly different surface temperature regime are not contained (e.g Westermann et al., 2012; Østby et al., 2014). However, these validation studies also concluded that a significant number of highly erroneous measurements due to undetected clouds is contained in the MODIS LST time series. These measurements are not removed in the synthesized MODIS/ERA time series, which could in principle introduce a bias in the freezing and thawing degree days. Since including ERA reanalysis data strongly increases the total number of values in the time series, these

biased values will attain a smaller weight in the FDD_s and TDD_s calculation compared to employing MODIS LST only, so that their influence is moderated. Nevertheless, it should be investigated if highly erroneous MODIS LST can be detected by a consistency check with downscaled ERA air temperatures, e.g. by setting a threshold value for the difference between the two temperatures.

In this study, the ERA–interim reanalysis is employed to complement satellite-based LST measurements. Since a large number of operational meteorological observations are assimilated in the reanalysis (Dee et al., 2011), independent validation data sets are rare especially in Arctic regions. Over the central Arctic ocean during summer, Jakobson et al. (2012) found near-surface temperatures to be warm-biased by up to 2°C, but the performance improved at higher altitudes. For Ireland, the performance for near-surface temperatures was excellent with a root mean square error of generally less than 0.5°C, but a slight warm bias during the winter months (Mooney et al., 2011). While these examples give a hint on the expected accuracy, the performance should be investigated further, especially in regions with pronounced topography. The coarse-scale temperature field of the ERA–interim reanalysis is downscaled using high-resolution topography and a seasonal lapse-rate calculated from the surface and the 700 mbar levels. The similar downscaling scheme presented by Fiddes and Gruber (2014), which is based on a 4-D interpolation between all available pressure levels and time steps, did not provide satisfactory results for grid cells with elevations below the ERA orography. This case is common e.g. at the coasts of Svalbard and Greenland, where the elevation difference can be more than 1000 m. In these cases, the interpolation relies on the "near-surface lapse rate", i.e. the modeled temperature difference between the surface and the first pressure level, which is subsequently extrapolated to altitudes far below the surface level. In particular if an ERA grid cell is snow-covered, this near-surface lapse-rate can feature strong temperature inversions, so that extrapolation to lower altitudes leads to a strong cold-bias of the downscaled temperature. To moderate the effect of near-surface stratifications, we employ an "average lapse rate" for the entire ERA air column between the surface and the 700 mbar pressure level located at approximately

the elevation of the highest non-glacierized areas in the study region. **We emphasize that the procedure cannot account for many regional and local climate and weather conditions, such as temperature inversions in valley systems. Downscaling of reanalysis data using e.g. the Weather Research and Forecasting (WRF) Model, though computationally expensive, may be a way to overcome such difficulties (Aas et al., 2015).**

4.1.3 Land cover

For each 1 km grid cell, the range of possible values for n_{f} and r_k is assigned based on the MODIS landcover product and large-scale snowfall data sets. Hereby, the main goal is to distinguish areas with bare ground or low vegetation, where strong redistribution of snow due to wind drift can occur, from areas with trees and high vegetation where the snow cover is more uniform. At least in some areas, the performance of the MODIS landcover product in distinguishing between the two classes seems to be poor: in the rolling plains of Finnmarksvidda, N Norway, mountain birch forest alternates with bare fell areas, while the entire area is classified as “open shrubland” in the MODIS landcover and thus as “high vegetation/forest” in the modeling. In the studied region, this problem is restricted to N Scandinavia and Russia, where permafrost occurs close to and within forested areas. In regions with more extensive forest cover in permafrost areas, this issue should be investigated in more detail. The new global land cover classification delivered by the “Climate Change Initiative” of the “European Space Agency” (ESA CCI, Hollmann et al., 2013) may constitute an improved product for permafrost modeling, but it’s performance in distinguishing areas with low trees and shrubs from bare tundra should be investigated in more detail. However, with this product, it may be feasible to distinguish additional classes which differ in their parameter ranges employed in the modeling (Table 1), at least in regional applications. Ideally, a pan-arctic land cover classification should be compiled which is specifically tailored for permafrost applications. Such a product should not only deliver information on the vegetation and surface state, but also on subsurface properties, which could e.g. be employed to better estimate the r_k factors in the presented model scheme.

4.1.4 Precipitation and snow depth

The presented ground temperature map is compiled without employing fine-scale data sets on precipitation or snow depth. Instead, only plausible ranges of values for n_f are assigned, which account for both total winter precipitation (i.e. spatially averaged snow depth) and wind redistribution of snow. In a recent study for Svalbard and Norway, Gisnås et al. (2014) demonstrated that annual average ground surface temperatures within areas of approximately one modeling grid cell (1 km) vary by up to 5 °C in areas with pronounced wind drift, while a borehole can only provide one point value sampled from the distribution of ground temperatures. The modeling approach explicitly accounts for this spatial variability by computing ground temperatures for a whole range of n_f value. We find that the agreement with borehole temperatures, despite using large-scale data sets of winter precipitation, is around ± 2 °C and thus similar to natural spatial variability of ground temperatures.

4.2 Deterministic vs. statistical modeling of ground temperatures

In the version employed for this study (Eq. 1), the TTOP approach accounts for the insulating effect of the snow cover, as well as the thermal offset between mean annual ground surface and ground temperatures. The complex physical processes which give rise to these effects, are lumped into one variable each (n_f and r_k) which must be adjusted to achieve adequate results. While both n_f and r_k are in principle empirical constants, they are subject to physical constraints which confines them to a certain range. Other than deterministic modeling schemes which seek to employ a single set of model parameters for each grid cell, we use the cumulative information content of many realizations, i.e. combinations of different parameters selected from the physically possible range (Sect. 4.1.1). Due to its simplicity and the small number of input parameters, the TTOP-approach is computationally efficient enough to compute a large number of realizations. The scheme can be expected to deliver adequate results if the uncertainty due the simplified model (Appendix A) is smaller than the uncertainty due to the unknown model parameters. We note that the model approach can not separate the spatial variability of the input parameters within a model grid cell from

the uncertainty due to generally unknown values of these parameters. Ideally, it delivers a range of possible MAGT's in which the true range of MAGTs (as documented by measurements, e.g. Gisnås et al., 2014) is contained. While the agreement between modeled and measured ranges is excellent for the few in-situ data sets on small-scale spatial variability of ground temperatures (Sect. 3.1), more work and in particular improved validation data sets are needed for a robust assessment of modeled and true ranges of MAGTs.

More sophisticated modeling schemes (e.g. Jafarov et al., 2012; Zhang et al., 2012; Fiddes et al., 2015; Westermann et al., 2013) feature significantly more model parameters which are required to achieve a realistic description of physical processes. While these parameters can generally be estimated to a certain level of confidence in regional-scale applications (e.g. Fiddes et al., 2015), their variation and spatial distribution is hard to access on continental to global scale. It is therefore questionable whether the performance of such computationally expensive schemes is better than the simple CryoGrid 1 model. Furthermore, the significant number of input parameters and the computational costs make sensitivity studies with many realizations unpractical, at least for many grid cells. In the CryoGrid 1 based scheme, however, a sensitivity is performed for each grid cell so that confidence limits for the results can be given. On the point scale, Langer et al. (2013) performed a sensitivity analysis for a transient permafrost model. They found modeled ground temperatures to vary by up to 4 °C if the input parameters related to snow and soil properties were allowed to vary within physically plausible limits. This suggests that on a continental scale the performance of more sophisticated schemes could be on the order of the much simpler TTOP approach.

4.3 Towards a global high-resolution ground temperature map

The "classic" permafrost map compiled by the IPA more than 15 years ago (Brown et al., 1997) is based on available field evidence for permafrost occurrence and properties. Since then, new global data sets have become available which so far have only rarely been used for permafrost mapping on large spatial scales. An exception is the global approach by Gruber (2012) who derived the probability of permafrost occurrence at 1 km scale from

downscaled air temperatures (obtained from atmospheric model output) using empirical probability functions.

In this study, we have demonstrated the continental-scale application of a statistical model approach with global satellite data sets and a reanalysis product as input data. In addition to probabilities of permafrost occurrence (as in Gruber, 2012), the approach facilitates mapping of ground temperatures and thus quantifying the thermal state of the permafrost. The model was successfully applied to an area of more than 5 million km^2 with strong spatial differences in ground temperatures and a wide variety of permafrost landscapes which suggests that the approach is scalable to include the entire contiguous permafrost extent on the Northern Hemisphere, an area of approximately 23 million km^2 (Brown et al., 1997). Hereby, it may become necessary to refine the parameterizations for the ranges of n_f and r_k (Table 1), possibly by introducing functional dependences on other environmental variables. A main uncertainty is the performance in forested regions which could not be sufficiently investigated in the N Atlantic study region. At least in case of dense forests or strong off-nadir angles remotely sensed LST may represent top-of-canopy temperatures instead of skin temperatures at the ground or snow surface, as required by CryoGrid 1. Thus, it should be investigated whether this effect leads to a reduced accuracy or even a systematic bias of modeled ground temperatures.

In addition to a global mapping, the presented method is also feasible for regional-scale applications, especially if improved data sets on e.g. surface cover, air temperature or snow depth are available. An example is the ground temperature map of Norway compiled by Gisnås et al. (2013) through application of CryoGrid 1 with gridded air temperature data sets, which could be improved both by integrating remotely sensed LST and by accounting for the subgrid variability of the snow cover in a statistical framework. However, we emphasize the semi-empirical nature of the model approach so that it should only be applied in areas with sufficient in-situ data sets for validation. **Furthermore, in mountain areas with strong topographic variations, satellite-based LST measurements at 1 km scale cannot sufficiently capture variations of altitude and exposition, so that the mod-**

eling scheme at best can be expected to deliver a first-order approximation of the permafrost distribution.

5 Conclusions

In this study, we present a modeling approach for ground temperatures using remote sensing data and reanalysis products as input data that could be operationally applied on continental to global scale at a scale of 1 km. Based on the equilibrium permafrost model CryoGrid 1, remotely sensed land surface temperatures and land cover are employed in conjunction with air temperatures and snowfall from the ERA–interim reanalysis. The coarse spatial resolution of reanalysis data is compensated by a statistical modeling framework which scans a range of possible model parameters and thus facilitates estimating the uncertainty. The approach is applied to the permafrost areas around the North Atlantic, and area with a wide variety of permafrost environments an strong differences in ground temperatures.

- The mean of all model realizations is within 2.5 °C of in-situ measurements for 139 of 143 boreholes contained in the "Thermal State of Permafrost" network. The differences between modeled and measured ground temperatures can be approximated by a Gaussian distribution with width 1.2 °C centered around zero.

- For tundra areas, the modeled spread of ground temperatures within the ensemble of model realizations is roughly in agreement with measurements of the spatial variability of the ground thermal regime caused by variable snow depths due to wind drift of snow.

- The ensemble approach with many model realizations facilitates assigning fractions of permafrost or permafrost probabilities for each grid cell. It is therefore possible to revamp the "classic" permafrost zonation system with sporadic, discontinuous and continuous permafrost at a scale of 1 km.

– Within the accuracy established by comparison to in-situ measurements, the modeled permafrost extent agrees well with field observations and previous studies in all investigated regions.

The study demonstrates the potential of remotely sensed data sets for permafrost modeling if they are employed in multi-source data fusion approaches. It constitutes an important step towards satellite-based mapping of the ground thermal regime at unprecedented spatial resolutions, **but uncertainties related to satellite-based land cover maps and associated ranges of model parameters must be investigated and resolved prior to global application.** Such novel ground temperature products bear significant potential for permafrost science, e.g. by improving the validation of climate model results. Furthermore, they can be of high value for societies and industries in permafrost-affected countries, e.g. for the planning of infrastructure.

Appendix A: Alternative derivation of the TTOP equation

A derivation of the TTOP equation (Eq. 1) is found in Romanovsky and Osterkamp (1995). Here, we present an alternative derivation which allows insight in the simplifications that are applied in the TTOP model.

We consider the permafrost case, $MAGT \leq 0°\mathrm{C}$. During the summer, the active layer thaws and features a thermal conductivity k_t. At the bottom of the active layer (i.e. at the top of the permafrost) at depth d, we assume a constant temperature of 0°C. Furthermore, d and k_t are assumed constant during the entire thaw season, which is a gross simplification of the natural processes. At each time t, the heat flux in the permafrost is then computed as a linear interpolation between the ground surface temperature $T_{\mathsf{gs}}(t)$ as

$$F_{\mathsf{summer}}(t) = -k_t \, \frac{T_{\mathsf{gs}}(t) - 0°\mathsf{C}}{d} \, .$$

During winter, the active layer is frozen with thermal conductivity k_f and a temperature of T_{TOP} at depth d, resulting in a heat flux of

$$F_{\mathsf{winter}}(t) = -k_f \, \frac{T_{\mathsf{gs}}(t) - T_{\mathsf{TOP}}}{d} \, .$$

In thermal equilibrium between summer and winter (ignoring the geothermal heat flux), the cumulated summer and winter heat fluxes must balance each other:

$$\int\limits_{\mathsf{summer}} F_{\mathsf{summer}}(t)\, dt + \int\limits_{\mathsf{winter}} F_{\mathsf{winter}}(t)\, dt = 0 \, .$$

Rearranging the equations yields the winter equilibrium temperature T_{TOP} as

$$\tau_w T_{\mathsf{TOP}} = \left(\frac{k_t}{k_f} \int\limits_{\mathsf{summer}} T_{\mathsf{gs}}(t)\, dt + \int\limits_{\mathsf{winter}} T_{\mathsf{gs}}(t)\, dt \right) \tag{A1}$$

with τ_w being the duration of the winter period. With τ the total period for which forcing data are available and τ_s the duration of the summer season, one can define the mean annual ground temperature $MAGT$ as the average of winter and summer temperatures at depth d as

$$MAGT := \frac{\tau_w T_{\mathsf{TOP}} + \tau_s \, 0^\circ\mathsf{C}}{\tau} \, . \tag{A2}$$

Inserting Eq. A1 in Eq. A2, one obtains Eq. 1 with the additional simplifications

$$\int\limits_{\mathsf{summer}} T_{\mathsf{gs}}(t) dt = n_t \int\limits_{T_{\mathsf{s}} > 0^\circ\mathsf{C}} T_{\mathsf{s}}(t) dt := n_t TDD_{\mathsf{s}}$$

and

$$\int\limits_{\mathsf{winter}} T_{\mathsf{gs}}(t)\, dt = n_f \int\limits_{T_{\mathsf{s}} \leq 0^\circ\mathsf{C}} T_{\mathsf{s}}(t)\, dt := n_f FDD_{\mathsf{s}} \, ,$$

where T_s is the surface temperature. For the non-permafrost case, $MAGT > 0°C$, the derivation is identical, but summer and winter change roles since the winter temperature at the bottom of the frozen surface layer is confined to 0°C in this case.

Acknowledgements. We gratefully acknowledge the help of the following individuals and institutions: Andreas Kääb provided advice on remote sensing-derived Digital Elevation Models. The MODIS LST level 3 and MODIS landcover data products are courtesy of the online Data Pool at the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota. GMTED2010 elevation data courtesy of the U.S. Geological Survey. ECMWF ERA–interim data used in this study have been obtained from the ECMWF data server. This project was funded by CryoMet (project no. 214465; Research Council of Norway), COUP (project no. 244903/E10; JPI Climate; Research Council of Norway), and the Department of Geosciences, University of Oslo.

References

Aas, K. S., Berntsen, T. K., Boike, J., Etzelmüller, B., Kristjánsson, J. E., Maturilli, M., Schuler, T. V., Stordal, F., and Westermann, S.: A comparison between simulated and observed surface energy balance at the Svalbard archipelago, Journal of Applied Meteorology and Climatology, in print, doi:10.1175/JAMC-D-14-0080.1, 2015.

Baranov, I. Y.: Geographical distribution of seasonally frozen ground and permafrost, in: General Geocryology, pp. 193–219, V. A. Obruchev Institute of Permafrost Studies, USSR Academy of Sciences, Part I, Ch. 7, Moscow, 1959.

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

Christiansen, H., Etzelmüller, B., Isaksen, K., Juliussen, H., Farbrot, H., Humlum, O., Johansson, M., Ingeman-Nielsen, T., Kristensen, L., Hjort, J., et al.: The thermal state of permafrost in the nordic area during the international polar year 2007–2009, Permafrost and Periglacial Processes, 21, 156–181, 2010.

Daanen, R. P., Ingeman-Nielsen, T., Marchenko, S. S., Romanovsky, V. E., Foged, N., Stendel, M., Christensen, J. H., and Hornbech Svendsen, K.: Permafrost degradation risk zone assessment using simulation models, The Cryosphere, 5, 1043–1056, doi:10.5194/tc-5-1043-2011, http://www.the-cryosphere.net/5/1043/2011/, 2011.

Danielson, J. and Gesch, D.: Global multi-resolution terrain elevation data 2010 (GMTED2010), U.S. Geological Survey Open-File Report 2011–1073, p. 26, 2011.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, 2011.

Elberling, B., Christiansen, H. H., and Hansen, B. U.: High nitrous oxide production from thawing permafrost, Nature Geoscience, 3, 332–335, 2010.

Elberling, B., Michelsen, A., Schädel, C., Schuur, E. A., Christiansen, H. H., Berg, L., Tamstorf, M. P., and Sigsgaard, C.: Long-term CO2 production following permafrost thaw, Nature Climate Change, 3, 890–894, 2013.

Etzelmüller, B., Berthling, I., and Sollid, J.: Aspects and concepts on the geomorphological significance of Holocene permafrost in southern Norway, Geomorphology, 52, 87–104, 2003.

Etzelmüller, B., Farbrot, H., Guðmundsson, Á., Humlum, O., Tveito, O. E., and Björnsson, H.: The regional distribution of mountain permafrost in Iceland, Permafrost and Periglacial Processes, 18, 185–199, 2007.

Farbrot, H., Etzelmüller, B., Schuler, T., Guðmundsson, Á., Eiken, T., Humlum, O., and Björnsson, H.: Thermal characteristics and impact of climate change on mountain permafrost in Iceland, Journal of Geophysical Research, 112, F03S90, doi:10.1029/2006JF000541, 2007.

Farbrot, H., Isaksen, K., Etzelmüller, B., and Gisnås, K.: Ground thermal regime and permafrost distribution under a changing climate in northern Norway, Permafrost and Periglacial Processes, 24, 20–38, 2013.

Fiddes, J. and Gruber, S.: TopoSCALE v.1.0: downscaling gridded climate data in complex terrain, Geoscientific Model Development, 7, 387–405, doi:10.5194/gmd-7-387-2014, http://www.geosci-model-dev.net/7/387/2014/, 2014.

Fiddes, J., Endrizzi, S., and Gruber, S.: Large-area land surface simulations in heterogeneous terrain driven by global data sets: application to mountain permafrost, The Cryosphere, 9, 411–426, 2015.

Frey, R. A., Ackerman, S. A., Liu, Y., Strabala, K. I., Zhang, H., Key, J. R., and Wang, X.: Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5, Journal of Atmospheric and Oceanic Technology, 25, 1057–1072, 2008.

Gisnås, K., Etzelmüller, B., Farbrot, H., Schuler, T., and Westermann, S.: CryoGRID 1.0: Permafrost distribution in Norway estimated by a spatial numerical model, Permafrost and Periglacial Processes, 24, 2–19, 2013.

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

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

Gruber, S. and Haeberli, W.: Permafrost in steep bedrock slopes and its temperature-related destabilization following climate change, Journal of Geophysical Research: Earth Surface (2003–2012), 112, doi:10.1029/2006JF000547, 2007.

Hollmann, R., Merchant, C., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., De Leeuw, G., Forsberg, R., et al.: The ESA climate change initiative: Satellite data records for essential climate variables, Bulletin of the American Meteorological Society, 94, 1541–1552, 2013.

International Permafrost Association: IPA-IPY Thermal State of Permafrost (TSP) Snapshot Borehole Inventory, Boulder, Colorado USA: National Snow and Ice Data Center, doi:10.7265/N57D2S25, 2010.

Jafarov, E. E., Marchenko, S. S., and Romanovsky, V. E.: Numerical modeling of permafrost dynamics in Alaska using a high spatial resolution dataset, The Cryosphere, 6, 613–624, 2012.

Jakobson, E., Vihma, T., Palo, T., Jakobson, L., Keernik, H., and Jaagus, J.: Validation of atmospheric reanalyses over the central Arctic Ocean, Geophysical Research Letters, 39, doi:10.1029/2012GL051591, 2012.

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

Lilleøren, K. S., Etzelmüller, B., Gärtner-Roer, I., Kääb, A., Westermann, S., and Guðmundsson, Á.: The Distribution, Thermal Characteristics and Dynamics of Permafrost in Tröllaskagi, Northern Iceland, as Inferred from the Distribution of Rock Glaciers and Ice-Cored Moraines, Permafrost and Periglacial Processes, 24, 322–335, 2013.

Liu, Y., Key, J., Frey, R., Ackerman, S., and Menzel, W.: Nighttime polar cloud detection with MODIS, Remote sensing of environment, 92, 181–194, 2004.

Luo, D., HuiJun, J., Marchenko, S., and Romanovsky, V.: Distribution and changes of active layer thickness (ALT) and soil temperature (TTOP) in the source area of the Yellow River using the GIPL model, Science China Earth Sciences, 57, 1834–1845, 2014.

Mooney, P., Mulligan, F., and Fealy, R.: Comparison of ERA-40, ERA-Interim and NCEP/NCAR reanalysis data with observed surface air temperatures over Ireland, International Journal of Climatology, 31, 545–557, 2011.

Østby, T. I., Schuler, T. V., and Westermann, S.: Severe cloud contamination of MODIS Land Surface Temperatures over an Arctic ice cap, Svalbard, Remote Sensing of Environment, 142, 95–102, 2014.

Osterkamp, T. and Romanovsky, V.: Evidence for warming and thawing of discontinuous permafrost in Alaska, Permafrost and Periglacial Processes, 10, 17–37, 1999.

Panda, S. K., Marchenko, S. S., and Romanovsky, V. E.: High-Resolution Permafrost Modeling in Denali National Park and Preserve, Natural Resource Technical Report, National Park Service, Fort Collins, Colorado, pp. 1–44, https://irma.nps.gov/App/Reference/Profile/2208990, 2014.

Romanovsky, V. and Osterkamp, T.: Interannual variations of the thermal regime of the active layer and near-surface permafrost in northern Alaska, Permafrost and Periglacial Processes, 6, 313–335, 1995.

Romanovsky, V., Smith, S., and Christiansen, H.: Permafrost thermal state in the polar Northern Hemisphere during the international polar year 2007–2009: A synthesis, Permafrost and Periglacial processes, 21, 106–116, 2010a.

Romanovsky, V. E., Drozdov, D. S., Oberman, N. G., Malkova, G. V., Kholodov, A. L., Marchenko, S. S., Moskalenko, N. G., Sergeev, D. O., Ukraintseva, N. G., Abramov, A. A., Gilichinsky, D. A., and Vasiliev, A. A.: Thermal state of permafrost in Russia, Permafrost and Periglacial Processes, 21, 136–155, doi:10.1002/ppp.683, 2010b.

Sazonova, T. and Romanovsky, V.: A model for regional-scale estimation of temporal and spatial variability of active layer thickness and mean annual ground temperatures, Permafrost and Periglacial Processes, 14, 125–139, 2003.

Schaefer, K., Zhang, T., Bruhwiler, L., and Barrett, A.: Amount and timing of permafrost carbon release in response to climate warming, Tellus B, 63, 165–180, doi:10.1111/j.1600-0889.2011.00527.x, 2011.

Schmid, M.-O., Gubler, S., Fiddes, J., and Gruber, S.: Inferring snowpack ripening and melt-out from distributed measurements of near-surface ground temperatures, The Cryosphere, 6, 1127–1139, 2012.

Schmidli, J., Frei, C., and Vidale, P. L.: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods, International journal of climatology, 26, 679–689, 2006.

Schneider von Deimling, T., Meinshausen, M., Levermann, A., Huber, V., Frieler, K., Lawrence, D., and Brovkin, V.: Estimating the near-surface permafrost-carbon feedback on global warming, Biogeosciences, 9, 649–665, 2012.

Schuur, E., Bockheim, J., Canadell, J., Euskirchen, E., Field, C., Goryachkin, S., Hagemann, S., Kuhry, P., Lafleur, P., Lee, H., et al.: Vulnerability of permafrost carbon to climate change: Implications for the global carbon cycle, BioScience, 58, 701–714, 2008.

Simmons, A., Uppala, S., Dee, D., and Kobayashi, S.: ERA-Interim: New ECMWF reanalysis products from 1989 onwards, ECMWF Newsletter, 110, 25–35, 2007.

Smith, M. and Riseborough, D.: Permafrost monitoring and detection of climate change, Permafrost and Periglacial Processes, 7, 301–309, 1996.

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

Smith, S., Romanovsky, V., Lewkowicz, A., Burn, C., Allard, M., Clow, G., Yoshikawa, K., and Throop, J.: Thermal state of permafrost in North America: A contribution to the International Polar Year, Permafrost and Periglacial Processes, 21, 117–135, 2010.

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Karna, J.-P., Koskinen, J., and Bojkov, B.: Implementing hemispherical snow water equivalent product assimilating weather station observations and spaceborne microwave data, in: Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International, pp. 3768–3771, IEEE, 2011.

Walter, K., Zimov, S., Chanton, J., Verbyla, D., and Chapin III, F.: Methane bubbling from Siberian thaw lakes as a positive feedback to climate warming, Nature, 443, 71–75, 2006.

Wan, Z., Zhang, Y., Zhang, Q., and Li, Z.: Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data, Remote Sensing of Environment, 83, 163–180, 2002.

Westermann, S., Langer, M., and Boike, J.: 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, 908–922, 2011.

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

Westermann, S., Schuler, T., Gisnås, K., and Etzelmüller, B.: Transient thermal modeling of permafrost conditions in Southern Norway, The Cryosphere, 7, 719–739, 2013.

Westermann, S., Elberling, B., Højlund Pedersen, S., Stendel, M., Hansen, B., and Liston, G.: Future permafrost conditions along environmental gradients in Zackenberg, Greenland, The Cryosphere, 9, 719–735, doi:10.5194/tc-9-719-2015, 2015.

Zhang, Y., Li, J., Wang, X., Chen, W., Sladen, W., Dyke, L., Dredge, L., Poitevin, J., McLennan, D., Stewart, H., Kowalchuk, S., Wu, W., Kershaw, P., and Brook, R. K.: Modelling and mapping permafrost at high spatial resolution in Wapusk National Park, Hudson Bay Lowlands, Canadian Journal of Earth Sciences, 49, 925–937, 2012.

Table 1. Intervals of input parameters assumed in the TTOP modeling. SF_{ERA}: average yearly snowfall from ERA–interim reanalysis in m of water equivalent.

class	$n_{\mathrm{f,min}}$	$n_{\mathrm{f,max}}$	n_t	$r_{k,\mathrm{min}}$	$r_{k,\mathrm{max}}$
bare ground/low vegetation	0.725–0.625$\times$$SF_{ERA}$	1	1	0.8	1
high vegetation/forest	0.625–0.625$\times$$SF_{ERA}$	0.925–0.625$\times$$SF_{ERA}$	1	0.7	0.9

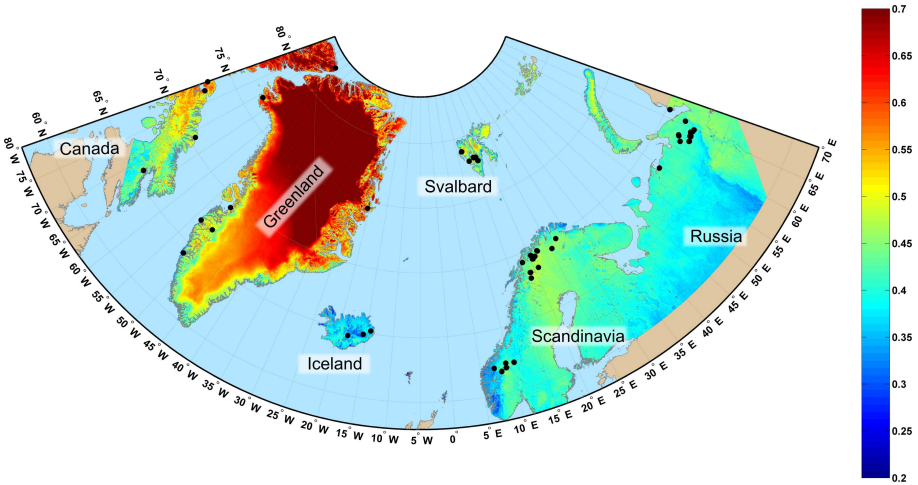

Figure 1. Fraction of MODIS LST measurements in the combined MODIS/ERA surface temperature product. Black dots: borehole sites from International Permafrost Association (2010) used for validation of the modeled ground temperatures. Note that a dot can represent several boreholes. Further boreholes outside the displayed regions are employed.

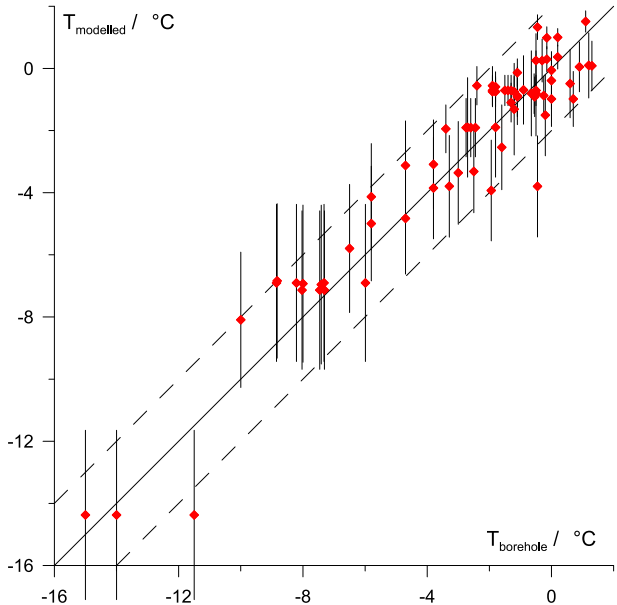

Figure 2. Measured vs. modeled ground temperatures for 143 borehole sites (International Permafrost Association, 2010). The error bars represent one SD of all model realizations, the dashed lines are the $\pm 2\,^{\circ}\mathrm{C}$ intervals around the $1:1$ line (in solid).

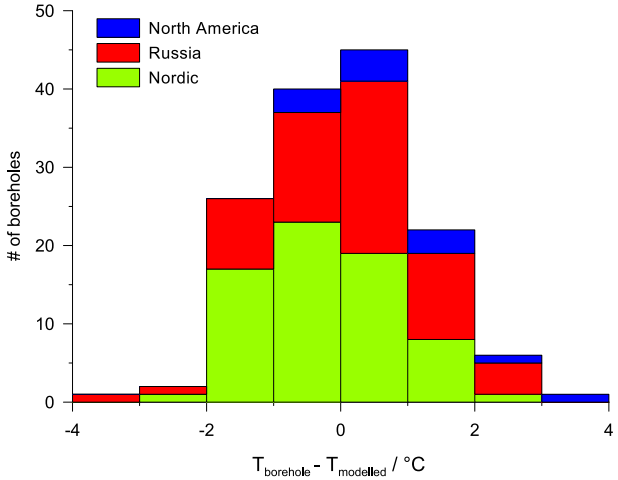

Figure 3. Histogram of the difference between modeled and measured ground temperatures for 143 borehole sites (International Permafrost Association, 2010).

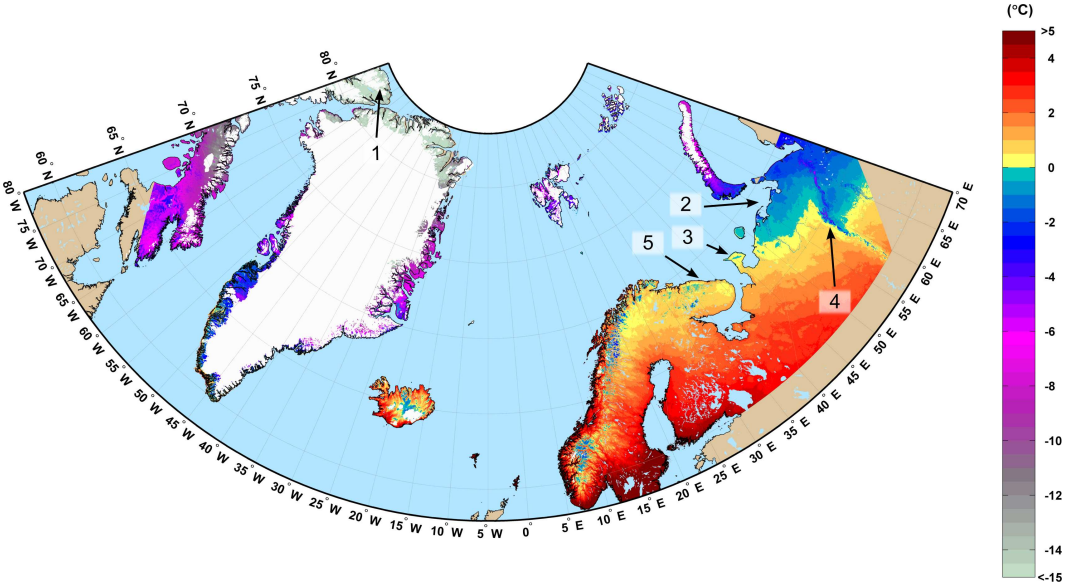

Figure 4. Mean of all model realizations for the N Atlantic study region. Glacierized areas are shown in white, water surfaces (ocean and larger lakes) in light blue. 1: Ellesemere Island; 2: Pechora Sea; 3: Kanin peninsula; 4: Ural mountains; 5: Kola peninsula.

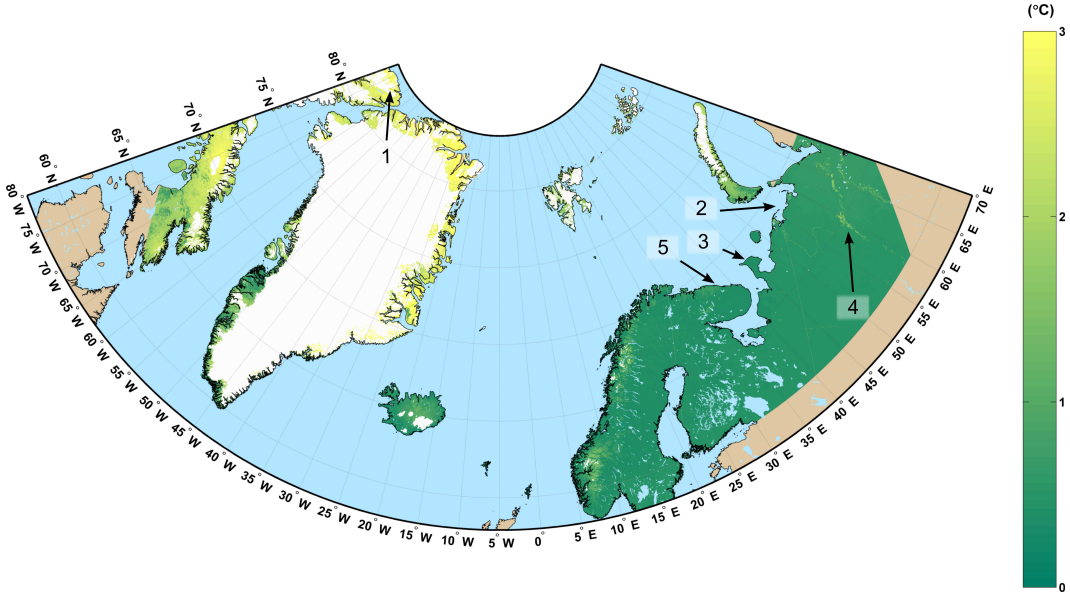

Figure 5. Standard deviation of all model realizations for the N Atlantic study region. Glacierized areas are shown in white, water surfaces (ocean and larger lakes) in light blue. 1: Ellesemere Island; 2: Pechora Sea; 3: Kanin peninsula; 4: Ural mountains; 5: Kola peninsula.

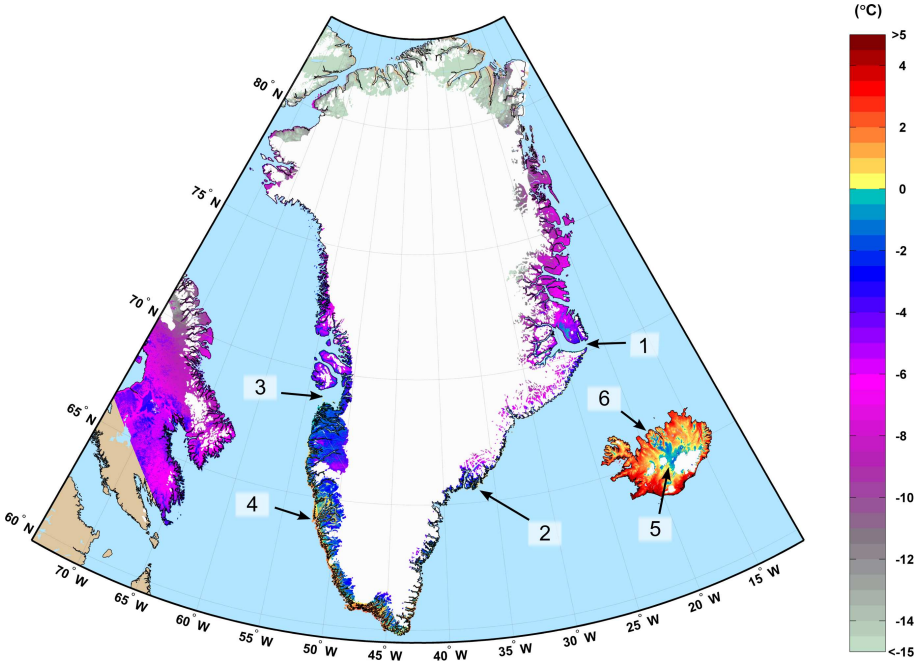

Figure 6. Mean of all model realizations for in Greenland and Iceland. Glacierized areas are shown in white, water surfaces (ocean and larger lakes) in light blue. 1: Scoresby Sound; 2: Tasilaq region; 3: Disko Bay; 4: Nuuk region; 5: Sprengisandur; 6: Tröllaskagi peninsula.

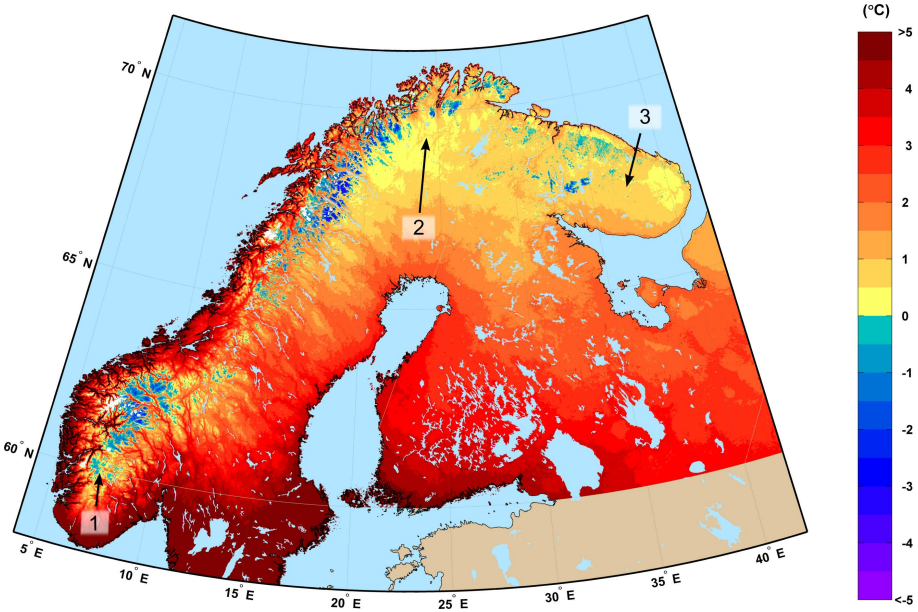

Figure 7. Mean of all model realizations for Scandinavia. Glacierized areas are shown in white, water surfaces (ocean and larger lakes) in light blue. 1: Hardangervidda; 2: Finnmarksvidda; 3: Kola peninsula.

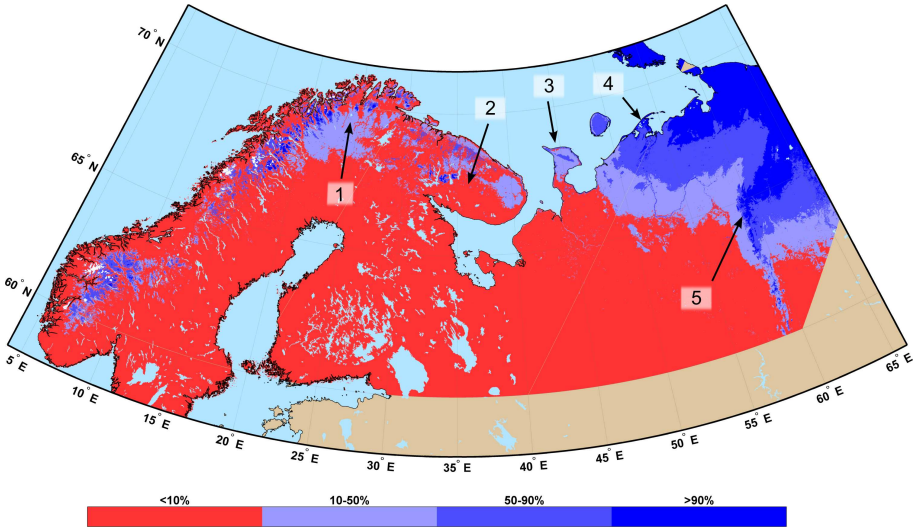

Figure 8. Permafrost zonation inferred from the statistical modeling in classes corresponding to continuous ($> 90\,\%$ of realizations MAGT $< 0\,^{\circ}$C), discontinuous (50–90 %) and sporadic (10–50 %) permafrost. Glacierized areas are shown in white, water surfaces (ocean and larger lakes) in light blue. 1: Finnmarksvidda; 2: Kola peninsula; 3: Kanin peninsula; 4: mouth of Pechora River; 5: Ural mountains.