Derivation and analysis of a high-resolution estimate of global permafrost zonation

S. Gruber
Glaciology, Geomorphodynamics and Geochronology, Department of Geography, University of Zurich, Switzerland

Correspondence to: S. Gruber (stephan.gruber@geo.uzh.ch)

Received: 30 April 2011 – Published in The Cryosphere Discuss.: 23 May 2011
Revised: 22 November 2011 – Accepted: 10 December 2011 – Published: 17 February 2012

Abstract. Permafrost underlies much of Earth’s surface and interacts with climate, eco-systems and human systems. It is a complex phenomenon controlled by climate and (sub-) surface properties and reacts to change with variable delay. Heterogeneity and sparse data challenge the modeling of its spatial distribution. Currently, there is no data set to adequately inform global studies of permafrost. The available data set for the Northern Hemisphere is frequently used for model evaluation, but its quality and consistency are difficult to assess. Here, a global model of permafrost extent and dataset of permafrost zonation are presented and discussed, extending earlier studies by including the Southern Hemisphere, by consistent data and methods, by attention to uncertainty and scaling. Established relationships between air temperature and the occurrence of permafrost are re-formulated into a model that is parametrized using published estimates. It is run with a high-resolution (<1 km) global elevation data and air temperatures based on the NCAR-NCEP reanalysis and CRU TS 2.0. The resulting data provide more spatial detail and a consistent extrapolation to remote regions, while aggregated values resemble previous studies. The estimated uncertainties affect regional patterns and aggregate number, and provide interesting insight. The permafrost area, i.e. the actual surface area underlain by permafrost, north of 60° S is estimated to be 13–18 × 10^6 km^2 or 9–14 % of the exposed land surface. The global permafrost area including Antarctic and sub-sea permafrost is estimated to be 16–21 × 10^6 km^2. The global permafrost region, i.e. the exposed land surface below which some permafrost can be expected, is estimated to be 22 ± 3 × 10^6 km^2. A large proportion of this exhibits considerable topography and spatially-discontinuous permafrost, underscoring the importance of attention to scaling issues and heterogeneity in large-area models.

1 Introduction

Permafrost underlies much of Earth’s surface and interacts with climate (Walter et al., 2006), eco-systems (Jorgenson et al., 2001) and human systems (Nelson et al., 2002). The interaction between permafrost, or its degradation, and human activity is diverse and varies with environmental and societal conditions. Examples include ground subsidence (Nelson et al., 2001), vegetation changes on pastures (Wang et al., 2006), slope instability (Gruber and Haeberli, 2007; Lewkowicz and Harris, 2005), hydrological changes (Woo et al., 2008), damage to infrastructure (Larsen et al., 2008), and special requirements for construction (Peng et al., 2007; Bommer et al., 2010). This list is not exhaustive and it is likely that climate change will bring about unexpected permafrost phenomena and societal impact in the future.

Many aspects of permafrost research such as the generalization of monitoring results and the anticipation of climate change impacts require knowledge of its spatial distribution. This is especially true (and difficult) in the large transition zone between the area of near-continuous permafrost and areas that have virtually no permafrost. Because of the hidden nature of permafrost, the ability to estimate and visualize its spatial distribution is of special importance as input or validation for models, for multi-disciplinary research, and for the communication of research findings of societal relevance. Currently, the Circum-Arctic Map of Permafrost and Ground-Ice Conditions (Brown et al., 1997; Heginbottom and Dubreuil, 1993) published by the International Permafrost Association (short: IPA map) is the most widely used base reference for Northern-hemisphere permafrost distribution. Its wide acceptance has made it the standard, against which the output of many models is judged (e.g. Stendel and Christensen, 2002) and with which permafrost is often visualized in a geographic context (e.g. Romanovsky et al.,...
2 Background

2.1 Fundamental issues in modeling PE

Permafrost is defined as sub-surface material (excluding glaciers) having a temperature of less or equal to 0°C during at least two consecutive years (ACGR, 1988). It is interesting that already here, uncertainty in permafrost mapping begins because permafrost has been differently defined as Earth material containing ice (Wang and French, 1995) in the former Soviet Union and the People’s Republic of China. The processes of surface and subsurface energy transfer with respect to a one-dimensional column (soil, vegetation, snow) are described and understood rather well (e.g. Williams and Smith, 1989) and a number of models and parameterizations of the relevant processes exist (Riseborough et al., 2008). However, the clandestine character of permafrost and the influence of heterogeneous and variable (sub-) surface and climate conditions dictate fundamental limitations to the modeling of PE over large areas.

Commonly, permafrost is represented on maps using the concept of permafrost zones characterized by PE. This proportion is assigned qualitatively and sometimes adjacent PE classes are defined with a percentage value separating them. A confusing diversity of terminologies and PE-based definitions exist (Heginbottom, 2002, Table 1) and their tabular comparison is easily misinterpreted towards an unrealistic accuracy of the stated PE limits. Interestingly, some definitions of PE classes extend down to 0% and thus areas without permafrost would still be classified as e.g. Isolated Patches.

Mean annual ground temperature (MAGT) as well as permafrost occurrence and properties exhibit strong lateral variation over ranges of only few meters to few kilometers. This is mostly due to the effects of topography (Gruber and Haeberli, 2007), vegetation cover (Shur and Jorgenson, 2007), ground material (Gubler et al., 2011), water bodies and flow (Burn, 2005; Endrizzi et al., 2011) or snow distribution (Zhang, 2005; Sturm and Benson, 2004; Liston, 2004). Consequently, point (borehole) measurements in any but the most homogeneous conditions do not suffice to describe conditions over an area, making their use for model calibration and validation difficult. In this paper, the terms calibration (adjustment of model parameters to improve agreement with
a reference data set), validation (demonstration of satisfactory accuracy consistent with the intended application, usually by comparison with data obtained by measurement of the real system) and credibility (sufficient degree of belief in a model to justify its use) are used as defined by Rykiel et al. (1996).

The macroscopic characterization of an area requires deciding on a suitable descriptor, e.g. the average behavior within the area, the majority behavior within the area, or the approximation of an area by the behavior of a single point within it. PE is such a macroscopic descriptor and deriving it in a meaningful way requires multiple measurements and a form of up-scaling or aggregation based on the composition of this area (e.g. Wright et al., 2003; Nelson et al., 1997; Lewkowicz and Bonnaventure, 2008). Because no coherent study of this type that spans wide environmental gradient exist, the calibration and validation of large-area models of PE (or similar metrics) is challenged by missing suitable data. The original delineation of PE classes (cf. Heginbottom, 2002) is similarly based on a qualitative (and subjective) up-scaling of observations. In map generation, this is then transferred to larger areas by extrapolation, usually along MAAT isotherms. Here, observations can include the occurrence of landforms such as active ice-wedge polygons, rock glaciers, or palsas that are considered diagnostic for certain classes of PE but that do not directly allow to estimate PE itself.

Scaling between point and area is frequently omitted in the use of physics-based models by calibrating with point measurements and subsequent application on gridded data. The tacit assumption here is that the aggregated properties of a grid cell such as its mean elevation, vegetation type, or snow cover are suitable model input resulting in relevant output. While this practice is obviously problematic, its consequences are hard to quantify or even notice due to the difficulty of providing spatial ground truth of permafrost occurrence or properties. Giorgi and Avissar (1997) have stated this general problem for non-linear processes as \( F(x) \neq \bar{F} \), i.e. the aggregated effect of variable \( x \) on process \( F \) cannot be represented by using the average of \( x \) to evaluate process \( F \). Depending on the intended application, even the mean behavior such as aggregated MAGT of a grid cell may not be a suitable description of sub-grid phenomena. Both effects are illustrated by Fig. 1: while MAGT resulting from the yellow cell or averaged from finer resolution may be adequate to inform e.g. land-atmosphere coupling, it would not indicate the permafrost at high elevation. Similar examples can be constructed in lowland areas based on variable ground properties, vegetation and snow drift.

Even with careful attention to the aggregation of physics-based model results from fine to coarse scales, some important effects are not easily predicted because lateral interaction at the fine scale can produce emergent phenomena. Many prominent permafrost landforms such as ice-wedge polygons and rock glaciers are examples of this. Especially in the light of climate change impacts, the importance of emergent phenomena can hardly be overstated. While cold areas with continuous permafrost are often regarded as stable, field observations reveal e.g. drastic geomorphic changes (Jorgenson et al., 2006; Bowden et al., 2008; Gooseff et al., 2009) that emerge from altered feedback mechanisms in only a small fraction of the landscape but strongly affect its overall functioning.

2.2 The effects of past climate fluctuations and anthropogenic climate change

Permafrost temperatures and conditions usually reflect a long history of changes in climate and land-surface dynamics. Depending on its depth, ice content and surface conditions, the time of formation and decay of permafrost can range from few to tens of thousands of years. While this precludes the straightforward definition of equilibrium conditions between climate and permafrost, equilibrium models (e.g. Riseborough et al., 2008) are frequently used because large-area patterns dictated by atmospheric circulation averaged over few decades already give a valuable first order estimate of permafrost occurrence, irrespective of local deviations. For the level of detail possible at the global scale, a reference period of a few decades duration that is based on dense measurements is desirable. The interest in permafrost zonation is partly motivated by the strong and unexpected alterations that permafrost-affected systems may undergo as a reaction to climate change. The nature and rate of those changes however is mainly dependent on ground characteristics and thus controlled by phenomena that occur at a much finer scale than that of the model presented here. To keep results robust and traceable, PE is thus estimated for conditions with little anthropogenic climate change. This is useful because it conservatively indicates areas that may be subject to permafrost conditions and thus possibly exhibit pronounced or unexpected reactions to change.
3 Data and methods

3.1 Model of permafrost extent

PE at the global scale is modeled as a function of MAAT alone and the combined effects of phenomena with a fine-scale variability such as snow cover, exposition to solar radiation, vegetation, or subsurface properties are parametrized stochastically. At one point, the probability of finding MAGT ≤ 0 °C is a function of MAAT using the cumulative normal distribution

\[ F_{\text{MAGT}} \leq 0 = \frac{1}{2} \operatorname{erfc} \left( \frac{\text{MAAT} + \mu}{\sqrt{2}\sigma^2} \right), \tag{1} \]

the underlying probability density function is given by

\[ F_{\text{MAGT}} = 0 = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(\text{MAAT} + \mu)^2}{2\sigma^2}}, \tag{2} \]

where \( \mu \) describes the mean temperature difference MAGT-MAAT and \( \sigma^2 \) describes the spread of the distribution. This allows to scale from a point to an area by interpreting probabilities in terms of PE and can be regarded as a re-formulation and re-interpretation of established relationships between PE and long-term MAAT. Model parameters \( \mu_L \) and \( \sigma_L^2 \) (subscript \( L \) indicates these to be based on literature) can be found based on two points for which MAAT and PE are estimated, e.g. \( F_{\text{MAGT}} \leq 0 (-8) = 0.9 \) and \( F_{\text{MAGT}} \leq 0 (-2) = 0.1 \). Isolated occurrence of permafrost at high MAAT (e.g. Delaloye et al., 2003) as well as the existence of permafrost-free ground at very low MAAT (e.g. Burn, 2005) support the preference of the normal over e.g. a uniform distribution in the simple approximation of the relationship between MAAT and MAGT. The elevation data used is accurate in its mean and its influence on PE via sub-grid variance is modeled as \( \sigma_L^2 = (\sigma_E L)^2 \) where \( \sigma_E \) is the sub-grid standard deviation of elevation and \( L \) the atmospheric lapse rate. These variances finally combine to \( \sigma^2 = \sigma_L^2 + \sigma_S^2 \) being used in the model.

3.2 Model parameters

In the absence of possibilities for adequate calibration, the parameters \( \mu_L \) and \( \sigma_L^2 \) are estimated based on the interpretation of published values. The PE classes used in the IPA map (Brown et al., 1997) are also employed here: 90–100 % (Continuous Permafrost), 50–90 % (Extensive Discontinuous Permafrost), 10–50 % (Sporadic Discontinuous Permafrost), and smaller than 10 % (Isolated Patches). Despite sometimes differing or inexact terminology, MAAT limits can be approximated to be \(-7.5 \) to \(-8.5 \) °C for an extent of \( 90 \% \) and \(-1 \) to \(-2 \) °C for \( 10 \% \) using published values (e.g. Bockheim, 1995; King, 1986; Péwé, 1983; Brown and Peeve, 1973; Nikiforoff, 1928).

This derivation and especially the original estimates are of course subject to large uncertainty and subjectivity. For published MAAT and PE limits, the determination of PE but also of MAAT or the time period over which this MAAT is averaged or PE estimated are subject to large uncertainty. To allow the propagation of this uncertainty into model results, it is qualitatively described by a conservative (i.e. more permafrost, Cold) and an anti-conservative (i.e. less permafrost, Warm) variant of the Norm parameter set used (Table 1). The Cold and Warm variants are derived by using the upper and lower bounds of MAAT for PE = 0.1 and PE = 0.9 as derived from the literature and by shifting PE at the respective limit by \( \pm 5 \) percentage points. The resulting model relationships are illustrated by Fig. 2.

3.3 Elevation, surface type, borders

Excluding Antarctica, the SRTM30 data set (USGS, 2005) provides a global description of elevation and its 30 arc-second grid (i.e., < 1 km resolution, WGS84) is used for the data products described in this contribution. Between 56° S and 60° N it is based on the Shuttle Radar Topography Mission (Farr et al., 2007) and at higher latitudes or where no valid data is available from this it is based on GTOPO30 (USGS, 2004) data of mostly lower quality. For SRTM covered areas, the standard deviation of elevation within each cell is provided with SRTM30, for all other areas it is estimated from the maximum slope within a 3 × 3 pixel neighborhood based on a regression analysis within the SRTM data. Oceans were masked using the DIF product in the

![Fig. 2. Model relationships between MAAT and PE as defined in Table 1.](image-url)
S. Gruber: High-resolution global permafrost zonation

SRTM30 data set. Glaciers and ice caps are masked based on the Digital Chart of the World (Defense Mapping Agency, 1992) gridded to SRTM30 resolution. Obvious errors in the glacier mask such as the large squares present in the Pyrenees or Central Italy were deleted before processing. Circular glacier outlines like those present in Eastern Siberia were kept. While they distort local patterns, they are expected to contribute to statistics on country or global level in a meaningful way. The use of higher quality glacier inventories (cf. Raup et al., 2007) is complicated by their currently incomplete global coverage. Large inland lakes were gridded based on the Data & Maps World data set (ESRI, 2010) that was also used for the delineations of countries.

3.4 Air temperature
Gridded monthly mean air temperatures were averaged to obtain MAAT based on the CRU TS 2.0 (Mitchell et al., 2003) dataset covering the global land surface at 0.5° spatial resolution and the NCEP reanalysis (Kalnay et al., 1996) with a resolution of 2.5°. The reference period 1961–1990 is used for all MAAT values calculated throughout this study. Temperatures were reduced to anomalies at sea level, and then linearly interpolated and re-adjusted to the SRTM30 grid and elevation. Lapse rates were derived from the NCEP data by linear regression of MAAT against geopotential heights for the lowermost seven pressure levels (cf. Mokhov and Akperov, 2006). The resulting 30 arc-second grids of MAAT are here referred to as CRU30 and NCEP30 and have been sampled at 1056 locations for comparison with norm values of meteorological stations. Interpreting the grids as expected values of MAAT within a cell, data was extracted using nearest-neighbors and no correction for elevation differences between SRTM30 and stations was made. The comparison data set contains 202 stations of the Adjusted Historical Canadian Climate Data (AHCCD, Vincent et al., 2002), 957 stations of the Weltklima data compilation by the German National Meteorological Service and six norm values published by MeteoSwiss.

3.5 Terrain ruggedness
Topography is a major driver of heterogeneity in the characteristics of permafrost and for evaluating its importance, a measure is required that is robust with respect to the variable cell size caused by an unprojected grid. To achieve this, ruggedness is quantified using an area normalization (Melton, 1965)

$$M = \frac{\max(E) - \min(E)}{\sqrt{A}}.$$  (3)

$E$ being elevation and $A$ the total area of a $5 \times 5$ pixel neighborhood. Gaussian convolution with a standard deviation of 5 km is used to obtain comparable areas of support for this index over a broad range of latitudes, and after this, the result is scaled into the ruggedness index $I = \ln(M \times 100)$. This procedure is rather robust with respect to grid spacing and pixel areas changing with latitude, but does not account for the variable quality of SRTM30 data sources. Two-dimensional Gaussian convolution is achieved by successive one-dimensional convolutions with a sampled Gauss function

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}$$  (4)

one in north-south and one in east-west direction. Filter kernels are normalized, 50 cells wide, and given by

$$\Delta y = \frac{\pi R}{21600} \quad \text{and} \quad \Delta x = \cos(\phi) \Delta y$$  (5)

north-south ($\Delta y$) and west-east ($\Delta x$) for a 30 arc-second spherical grid. $R = 6371$ km is the mean radius of Earth and $\phi$ denotes latitude.

3.6 Simulations and further processing
In the absence of better knowledge, the average of CRU30 and NCEP30 is regarded as the best guess of MAAT and combined with the Norm variant (Table 1) of the PE model. To propagate the most important sources of uncertainty, additional simulations were performed by permutations of parameter sets and data sources (Table 2). A mask is applied to define the exposed land surface, excluding cells in oceans, large inland lakes as well as in ice sheets or glaciers. Because the accuracy of estimated PE cannot be demonstrated and many relevant fine-scale processes have to be neglected at the global scale, model results are interpreted as a permafrost zonation index (PZI) that serves to represent spatial patterns but that does not provide actual extent or probability of permafrost at a location. The PZI fringe is the zone of uncertainty over which PZI could extend under conservative estimates (cf. Table 1). Permafrost area (PA) is defined here as the area of a pixel with a PZI larger than a threshold (cf. Zhang et al., 2000). Unless otherwise stated, the threshold for PR is PZI $\geq 0.1$. For map products and visual display only ice sheets and large ice caps are excluded but not glaciers. This is because the glacier data used often has a very coarse resolution and would result in data gaps in mountain regions. When visualized in a GIS or virtual globe, glaciers can easily be discerned.

4 Model limitations and evaluation of products
4.1 General considerations
In the background section, I have argued that rigorous calibration and validation of a global PE model currently are
Table 2. Simulation parameters and input data.

<table>
<thead>
<tr>
<th>Code</th>
<th>Parameter set</th>
<th>MAAT source</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC</td>
<td>Warm</td>
<td>CRU30</td>
</tr>
<tr>
<td>WN</td>
<td>Warm</td>
<td>NCEP30</td>
</tr>
<tr>
<td>NM</td>
<td>Norm</td>
<td>(CRU30 + NCEP30)/2</td>
</tr>
<tr>
<td>CC</td>
<td>Cold</td>
<td>CRU30</td>
</tr>
<tr>
<td>CN</td>
<td>Cold</td>
<td>NCEP30</td>
</tr>
</tbody>
</table>

hardly possible. Model credibility however, can be supported by making the most important structural limitations understandable, by evaluating MAAT that is the most important input used, and by comparison with the IPA map. The model presented has a number of important limitations: Differences in the warming effect that winter snow exerts on ground temperatures and the topographic differentiation between shaded and sun-exposed slopes are not reproduced because neither precipitation nor solar radiation are taken into account. This implies that the relationship of PZI and actual PE or the occurrence of certain phenomena and landforms may have regional differences. In mountain areas, a sun-exposed slope will likely have less (or less likely) permafrost than one with the same PZI in a shaded location. Deep permafrost and the influence its presence or absence due to Pleistocene glacial and thermal conditions has on near-surface conditions today are not represented by the model. Cold air drainage and inversions, that often affect ground temperatures along valleys and depressions are also not included in the model used here. Similarly, the effects that ground thermal properties and vegetation have on permafrost are not included. Although one global $\sigma^2$ is used, the degree of sub-grid variability it represents may be larger in some regions (e.g. steep mountain areas or areas with diverse plant cover) than in others (e.g. areas with uniform ground cover). Transient effects are not represented.

4.2 Evaluation of MAAT grids

The differences between MAAT at 1056 meteorological stations and CRU30 have a mean of 0.09 °C and a standard deviation of 0.78 °C and for NCEP30 –0.16 °C and 1.28 °C. This provides quality control for the interpolation and regridding procedure but is not a validation because the stations used have partly been used for the original CRU and NCEP products. Figure 3 illustrates that the above figures by themselves could be greatly misleading: the spatial patterns of differences between NCEP30 and CRU30 (overall $\mu = -0.1$ °C, $\sigma = 1.53$ °C) reveal large absolute differences, especially in remote regions and mountain areas (cf. Anisimov et al., 2007; Ma et al., 2008). These patterns are visible also when directly plotting differences of NCEP resampled to the CRU grid without adjustment for elevation and thus are likely not artifacts of the processing presented here. The small averages of the differences between both data sets and between the station normals and the gridded data suggest that both CRU30 and NCEP30 describe plausible states of the global atmosphere but with strong differences in data-sparse areas. Given two widely accepted but differing global datasets of MAAT, it is beyond the scope of this study to try and improve them or to assign greater credibility to the one or the other. Instead, the effects of those differences are propagated into the final results of this study.

4.3 Ruggedness

The ruggedness index is a measure of how flat or mountainous a landscape is. It is intended to be robust with respect to global application in order to facilitate comparison and statistical evaluation. Figure 4 shows a global overview of ruggedness as well as three inset maps to facilitate comparison with familiar landscapes as a subjective reference. The five classes shown are flat (index of 0–1.5), undulating (1.5–2.5), hilly (2.5–3.5), mountainous (3.5–4.5) and rugged (> 4.5).

4.4 Evaluation of permafrost zonation index

In the absence of possibilities for proper validation, the new map of permafrost zonation is compared with the widely accepted IPA map. The spatial resolution of this comparison can only be that of the IPA map or coarser, and the ability of the new dataset to resolve higher resolution patterns has to be inferred based on its model structure. Removing glaciers and large inland lakes, the exposed land surface of the Northern Hemisphere is estimated to about $9.7 \times 10^6$ km². Using this mask, the IPA map has a permafrost region (PE ≥ 0.1) of $1.93 \times 10^6$ km² or $23.3 \times 10^6$ km² if polygons of isolated permafrost are included (cf. Zhang et al., 2000). By comparison, the new dataset here has a permafrost region (PZI ≥ 0.1) of $21.7 \times 10^6$ km² for the NM parameter set, with a low estimate of $18.7 \times 10^6$ km² (WC) and a high one of $24.3 \times 10^6$ km² (CN).

Figure 5 allows a spatial comparison of the IPA map with the new product in two regions. For this, the calculated PE has been contoured, simplified and smoothed, and short contours, especially in mountain areas, were manually removed. Perfect correspondence between the parameter combination NM and the IPA map would be indicated by the thick red, purple and blue lines coinciding with the boundaries between the four shades of grey that indicate IPA map PE classes. With some local deviations, this is almost the case for the North American region (Fig. 5b). The region between the conservative (CC, CN) and anti-conservative (WC, WN) estimates (cf. Table 1) has a similar latitudinal extent as the classes of Extensive Discontinuous Permafrost and Sporadic Discontinuous Permafrost in the IPA map. The situation in the region in Western Siberia (Fig. 5a) indicates larger differences. Here, the boundary between Isolated
Fig. 3. Differences between CRU30 and NCEP30 highlight how model-based (NCEP) and statistical (CRU) spatial interpolation causes differences, especially in remote and data-sparse regions.

Fig. 4. Ruggedness Index displayed in five classes on a global overview (A) and in three inset maps to aid subjective familiarization: (B) Europe, (C) Tibetan Plateau, and (D) Alaska.
Patches and Sporadic Discontinuous Permafrost of the IPA map corresponds with the anti-conservative (WC, WN) estimates whereas the boundary between Extensive Discontinuous Permafrost and Continuous Permafrost approximately corresponds with the conservative (CC, CN) parameter combinations. This implies that the IPA map classes here, and possibly also elsewhere in Russia and former countries of the Soviet Union, are spread over a narrower range of MAAT, at least with respect to the data used in this study.

5 Results

5.1 Permafrost zonation index map

Figure 6 provides a high-resolution display of the PZI map in an exemplary remote region. The area shown is situated in the Pamirs, in the Eastern part of the Wakhan Corridor, where the Islamic Republic of Afghanistan extends between the Republic of Tajikistan in the North and the Islamic Republic of Pakistan in the South. Major topographic features are represented by the new PZI map whereas the IPA map lacks sufficient detail and accuracy to allow interpretation at this scale. Figure 6b additionally visualizes the PZI fringe as the most conservative estimate of PR extension.

PZI or a map color indicate, to what degree permafrost exists only in the most favorable conditions (yellow in the legend of Fig. 6) or nearly everywhere (blue). These local conditions affecting permafrost occurrence will partly exhibit regional trends (e.g. mean snow cover characteristics or continentality), partly vary over typical distances on the order of several km (e.g. shaded or sun-exposed side of a mountain), and partly over tens to hundreds of meters (e.g. snow drift, vegetation, ground material). These conditions need to be assessed during interpretation, depending on the intended purpose of using the PZI map. This product is likely to be most valuable in remote regions where only sparse reliable information exists. In the future, areas with more available information may thereby serve as benchmarks for training the local interpretation of PZI, enabling later transfer to similar but data-sparse regions.

5.2 Summary statistics

Excluding Antarctica, the global PR is estimated to be \(21.7 \times 10^6 \text{ km}^2\) or 17% of the exposed surface area. PA is estimated to 12.9–17.8 \(\times 10^6 \text{ km}^2\) or 9–14%. Including Antarctic (Bockheim, 1995) and sub-sea (estimated based on Shakhova et al., 2010; Osterkamp, 2001; Péwé, 1983) permafrost, the global PA is approximately 16–21 \(\times 10^6 \text{ km}^2\). Table 3 provides detailed results and ranges of uncertainty and Table 4 shows estimated PR and PE for 25 countries. Many earlier estimates are well matched for e.g. the Northern Hemisphere (Zhang et al., 2000, and references therein) or Russia (Baranov and Kudryavtsev, 1963). However, this is not surprising because the general method for large-area delineation (MAAT) has remained the same and mainly the data basis, technical sophistication, reproducibility and communication of uncertainty have evolved with the present study. Especially outside the large areas of near-continuous permafrost in Russia, Canada, China and Alaska, the new figures provide an improvement of summary statistics and uncertainty estimates. Depending on topography and climate as well as on the uncertainty due to differences between NCEP30 and CRU30, the range of uncertainty for PR is large for some countries, including Finland (283%), Argentina (218%), Chile (246%) and North Korea (251%).

5.3 Selected characteristics of permafrost regions

Figure 7 displays PR estimates for the exposed land surface north of 60° S in relation to latitude, elevation, terrain ruggedness and PZI. The distribution with elevation and latitude reproduces the results of Zhang et al. (1999), adding an estimate of uncertainty and the Southern Hemisphere, that is shown cumulatively in part (a) at 20° N. Less than one third of the PR is situated north of the Arctic circle and about three quarters of the PR lies at elevations below...
Fig. 6. High-resolution example of the PZI map in a remote region in the Pamirs. Part (A) shows the best-guess PZI as derived from the NM parameter set. Part (B) additionally includes the PZI fringe that indicates the largest extension of the permafrost region within the uncertainty explored (cf. Table 1). A hill shade image provides an impression of the underlying topography. The new product provides markedly more detail than the IPA map that is subject to strong spatial (location and resolution of polygons) and categorical (assignment of classes) generalization.

Table 3. Summary of global permafrost distribution. Permafrost region is defined here as having a permafrost zonation index $\geq 0.1$ in the NM (cf. Table 2) simulation, numbers in brackets indicate range of results from the remaining parameter sets. Permafrost area is only indicated by its range over all parameter combinations due to the lack of calibration for permafrost extent. Sources other than this study are: A: Bockheim (1995), B: Shakhova et al. (2010); Osterkamp (2001); Péwé (1983), sub-glacial permafrost is not considered.

<table>
<thead>
<tr>
<th>Region or type</th>
<th>Exposed area $10^6$ km$^2$</th>
<th>Permafrost region $10^6$ km$^2$</th>
<th>% of exposed area</th>
<th>Permafrost area $10^6$ km$^2$</th>
<th>% of exposed area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global, north of 60° S</td>
<td>131.6</td>
<td>21.7 (18.8–24.4)</td>
<td>17 (14–19)</td>
<td>12.9–17.8</td>
<td>9–14</td>
</tr>
<tr>
<td>Northern Hemisphere</td>
<td>97.1</td>
<td>21.7 (18.7–24.3)</td>
<td>22 (19–25)</td>
<td>12.9–17.7</td>
<td>13–18</td>
</tr>
<tr>
<td>Southern Hemisphere, north of 60° S</td>
<td>34.5</td>
<td>0.05 (0.01–0.15)</td>
<td>0.14 (0.03–0.43)</td>
<td>0.018–0.019</td>
<td>0.052–0.055</td>
</tr>
<tr>
<td>Antarctica, ice-free (A)</td>
<td>0.28</td>
<td>0.28</td>
<td>100</td>
<td>0.28</td>
<td>100</td>
</tr>
<tr>
<td>Sub-sea (B)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>2.8</td>
<td>–</td>
</tr>
<tr>
<td>Global, exposed surface</td>
<td>131.9</td>
<td>22.0 (19.1–24.7)</td>
<td>17 (14–19)</td>
<td>13.2–18.1</td>
<td>9–14</td>
</tr>
<tr>
<td>Global, including sub-sea</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>16.0–20.9</td>
<td>–</td>
</tr>
</tbody>
</table>

1000 m.a.s.l. However, Fig. 7c demonstrates that the predominance of low elevations does not imply topography to be unimportant. On the contrary, a large proportion of the PR is situated in areas of considerable relief. Figure 7d shows that considerable areas of (near-) continuous permafrost exist, but more importantly, that most permafrost is spatially discontinuous and strongly controlled by local variability. The relative importance of pronounced
topography and spatially-discontinuous permafrost in this global analysis underscore the need to carefully consider their parameterization or representation in areal models and assessments of permafrost conditions and development.

6 Conclusions

The data described here represents half a century old knowledge applied to current data and enables relevant new insight. The conclusions of this study furthermore comprise lessons learned during model derivation and implications for further studies that arise from this.

– Limitations in modeling permafrost extent
The modeling of permafrost over large areas is challenged by sparse data and scaling issues in a fundamental way. For this reason, proper calibration and validation of corresponding models is difficult and often avoided. The large regional differences in MAAT (that should be one of the better-known climate variables) between two accepted gridded data sets underscores the importance of carefully investigating the effect of uncertain input in large-area permafrost analyses. While the dataset presented here can be used as a reference for model evaluation, it does then by no means represent reliable ground truth.

– Permafrost zonation
A global (north of 60° S) and high-resolution dataset of permafrost zonation has been described and evaluated. Based on its structure and coarse-scale comparison with the IPA map it is expected to have a similar credibility. While in well-researched areas, the IPA map may have greater quality, the new product likely overall provides a more reliable and consistent picture given the large amount of extrapolation involved. The increased spatial resolution is beneficial in mountain areas where the influence of topography on MAAT is important. A spatial representation of model uncertainty is included in the permafrost zonation map through a fringe area that spans the maximum plausible extension of the...
Table 4. Permafrost area and region for the top 25 countries, ranked by permafrost region for the NM (cf. Table 2) parameter set. Numbers in brackets indicate range of permafrost region from the remaining parameter sets. Permafrost area is only indicated by its range over all parameter combinations due to the lack of calibration for permafrost extent.

<table>
<thead>
<tr>
<th>Country</th>
<th>Permafrost region 10^3 km²</th>
<th>Permafrost area 10^3 km²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russia</td>
<td>10,968 (9,619–12,006)</td>
<td>6,966–9,541</td>
</tr>
<tr>
<td>Canada</td>
<td>6,031 (5,238–6,695)</td>
<td>3,637–4,978</td>
</tr>
<tr>
<td>China</td>
<td>20,561 (14,57–24,63)</td>
<td>6,73–16,76</td>
</tr>
<tr>
<td>United States</td>
<td>11,32 (902–1242)</td>
<td>530–877</td>
</tr>
<tr>
<td>Mongolia</td>
<td>530 (357–758)</td>
<td>165–394</td>
</tr>
<tr>
<td>Greenland</td>
<td>297 (282–303)</td>
<td>226–276</td>
</tr>
<tr>
<td>Norway</td>
<td>84 (35–127)</td>
<td>25–65</td>
</tr>
<tr>
<td>Kyrgyzistan</td>
<td>72 (47–88)</td>
<td>19,6–54</td>
</tr>
<tr>
<td>Sweden</td>
<td>59 (113–123)</td>
<td>4,7–43</td>
</tr>
<tr>
<td>India</td>
<td>58 (21–88)</td>
<td>7,8–65</td>
</tr>
<tr>
<td>Tajikistan</td>
<td>54 (30–62)</td>
<td>11,6–48</td>
</tr>
<tr>
<td>Pakistan</td>
<td>40 (12.2–60)</td>
<td>4,4–43</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>39 (27–49)</td>
<td>12,2–27</td>
</tr>
<tr>
<td>Finland</td>
<td>34 (0.9–97)</td>
<td>2,6–23</td>
</tr>
<tr>
<td>Argentina</td>
<td>30 (9.1–76)</td>
<td>3,5–33</td>
</tr>
<tr>
<td>Afghanistan</td>
<td>27 (12.4–40.8)</td>
<td>4,4–26</td>
</tr>
<tr>
<td>Nepal</td>
<td>15.7 (10.5–19.5)</td>
<td>4,5–12.4</td>
</tr>
<tr>
<td>Chile</td>
<td>14.0 (3.5–37)</td>
<td>1,4–15.2</td>
</tr>
<tr>
<td>Italy</td>
<td>5.1 (2.6–8.4)</td>
<td>0,3–3.2</td>
</tr>
<tr>
<td>North Korea</td>
<td>4.7 (0.7–12.5)</td>
<td>0.3–4.1</td>
</tr>
<tr>
<td>Georgia</td>
<td>4.7 (2.4–7.6)</td>
<td>0.8–3.2</td>
</tr>
<tr>
<td>Switzerland</td>
<td>4.5 (2.3–6.6)</td>
<td>0.7–2.5</td>
</tr>
<tr>
<td>Austria</td>
<td>3.6 (1.7–6.3)</td>
<td>0.5–2.3</td>
</tr>
<tr>
<td>Bhutan</td>
<td>3.1 (0.8–6.4)</td>
<td>0.3–3.2</td>
</tr>
<tr>
<td>Uzbekistan</td>
<td>2.8 (1.8–3.7)</td>
<td>0.6–1.8</td>
</tr>
</tbody>
</table>

permafrost region. Because the often multi-faceted response of permafrost to climate change moderated by local heterogeneity, a derivation of future permafrost zonation based on the map presented here is not useful.

- Global statistics and characteristics
  The permafrost area north of 60° S is approximately 13–18 × 10^6 km² or 9–14 % of the exposed land surface. The global permafrost area including Antarctic and sub-sea permafrost is estimated to be 16–21 × 10^6 km². The global permafrost region is estimated to be 22 ± 3 × 10^6 km². A large proportion of this area exhibits considerable topography and spatially-discontinuous permafrost, underscoring the importance that heterogeneity has for permafrost phenomena.

- Implications
  The importance of heterogeneity (in term of global permafrost area affected) implies that in large-area permafrost studies, the relevant output variables likely require parameterization or computation of sub-grid effects. The use of average or majority methods to aggregate sub-grid environmental conditions or model results may conceal the most important effects and lead to unrealistic conclusions. Given the lack of data for comparison at the grid-cell scale, this is difficult to detect. Attention to scaling (space, time, process or functional detail) and model evaluation is essential for reconciling field observations and model-based research and strongly determines the value of modeling studies.

The availability of consistent PZI data can motivate systematic comparison with local evidence or aggregated finescale models. Here, a global extension of consistent evidence inventories (e.g., Cremonese et al., 2011) would be instrumental. Finally, quantifying the population living in permafrost regions would be an important additional metric to inform future research strategies. First analyses based on gridded global population data (Balk et al., 2005) support the intuitive notion that the distribution of population in permafrost areas is shifted towards higher elevation, lower latitude and more rugged terrain in comparison to Fig. 7 shown for area, here. However, the inconsistent and often incomplete spatial disaggregation of population in remote areas currently precludes consistent global analyses.

7 Data availability
The Global Permafrost Zonation Index resulting from this study as well as important intermediate data are available at http://www.geo.uzh.ch/microsite/cryodata. The most important data can be viewed in Google Earth or in a GIS via a web-mapping service at 30 arc-second resolution and raw data can be downloaded as binary files.

Acknowledgements. NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. Stefanie Gubler has helped to improve the formulation of the model used. Ross Purves and Alexander Brenning have been great discussion partners concerning the model shown here and many more colleagues have helped by commenting earlier manuscripts.

Edited by: T. Zhang

References
ACGR: Glossary of permafrost and related ground ice terms, Associate Committee on Geotechnical Research (ACGR), National Research Council of Canada, Ottawa, 1988.


The Cryosphere, 6, 221–233, 2012

www.the-cryosphere.net/6/221/2012/


