The Cryosphere Discuss., 7, 5853–5887, 2013
www.the-cryosphere-discuss.net/7/5853/2013/
doi:10.5194/tcd-7-5853-2013

The Cryosphere Discussions
Open Access

Large area land surface simulations in heterogeneous terrain driven by global datasets: application to mountain permafrost

J. Fiddes[1], **S. Endrizzi**[1], **and S. Gruber**[2]

[1]Department of Geography, University of Zurich, Switzerland
[2]Department of Geography and Environmental Studies, Carleton University, Ottawa, Canada

Received: 6 November 2013 – Accepted: 22 November 2013 – Published: 11 December 2013

Correspondence to: J. Fiddes (joel.fiddes@geo.uzh.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Numerical simulations of land-surface processes are important in order to perform landscape-scale assessments of earth-systems. This task is problematic in complex terrain due to: (i) high resolution grids required to capture strong lateral variability, (ii) lack of meteorological forcing data where it is required. In this study we test a topography and climate processor, which is designed for use with large area land surface simulation, in complex and remote terrain. The scheme is driven entirely by globally available datasets. We simulate air temperature, ground surface temperature, snow depth and test the model with a large network of measurements in the Swiss Alps. We obtain RMSE values of 0.64 °C for air temperature, 0.67–1.34 °C for non-bedrock ground surface temperature, and 44.5 mm for snow depth, which is likely affected by poor input precipitation field. Due to this we trial a simple winter precipitation correction method based on melt-dates of the snow-pack. We present a test application of the scheme in the context of simulating mountain permafrost. The scheme produces a permafrost estimate of 2000 km^2 which compares well to published estimates. We suggest that this scheme represents a good first effort in application of numerical models over large areas in heterogeneous terrain.

1 Introduction

Numerical simulation is an increasingly important tool for assessment of the energy and mass balance at the earth's surface for many fields of research and application (e.g. Wood et al., 2011; Barnett et al., 2005; Gruber, 2012). In addition, numerical methods allow for transient assessment of past and future states, an essential step for change detection of (-near) surface conditions (Etzelmüller, 2013). Numerical approaches may also provide the means to simulate land surface variables where there is insufficient data for statistical methods e.g. remote areas or future periods.

TCD
7, 5853–5887, 2013

Landscapes that are heterogeneous in terms of e.g., topography, vegetation or redistribution of snow (e.g. Smith and Riseborough, 2002; Liston and Haehnel, 2007), provide a great challenge in this respect, as surface and subsurface conditions may vary on various, and often short, length scales, creating highly spatially differentiated surface-atmosphere interactions. This poses, in particular, a challenge to large-area simulations, which can be summarised as: (1) high resolution grids are required to capture surface heterogeneity, which is often numerically prohibitive over large areas, efficient methods are therefore required to make this task scalable; (2) there is often a lack of a representative forcing at the site or scale that it is required, particularly in remote regions.

Recent efforts in this respect include spatially explicit or distributed simulations i.e. Jafarov et al. (2012) produced a transient run of the ground thermal state in Alaska to assess permafrost dynamics under IPCC change scenarios. Another meso-scale modelling effort Gisnå s et al. (2013) provides an equilibrium model of permafrost distribution in Norway at a spatial resolution of 1 km^2. While representing major steps in application of numerical models over large areas, the grid resolution of 1–2 km is too coarse to simulate relevant spatial differentiation on fine scales, particularly under heterogeneous terrain.

In global climate models, a spatially detailed representation of the sub-grid land surface still remains some way behind the implementation of the atmosphere, yet is recognised to be key in accurately simulating feedbacks to the atmosphere (Pitman, 2003) e.g., surface albedo-atmosphere exchanges in the energy balance (Betts, 2009). For example, land surface heterogeneity is often represented in tiled approaches (Koster and Suarez, 1992), where surface types are represented by a limited number (or even a single) surface types, energy and mass balance is then computed independently, and finally aggregated at grid level. Here too, methods capable of representing fine-scale land surface heterogeneity efficiently could be useful. Finally, methods exist (e.g., SAFRAN-Crocus scheme; Durand et al., 1993, 1999) which classify topography ac-

TCD
7, 5853–5887, 2013

cording to fixed classes based on terrain parameters and enable application of numerical models over large areas in a semi-distributed fashion.

In previous papers (Fiddes and Gruber, 2012, 2013), methods have been developed and tested which enable (i) physically-based land surface models (LSM) to be applied over large areas using a sub-grid scheme that samples land surface heterogeneity and (ii) a method that scales gridded climate data necessary to drive an LSM, to the sub-grid using atmospheric profiles. The philosophy behind these approaches is to develop methods that depend only on globally available datasets to derive high-quality local results in heterogeneous and/or remote regions.

We perform this study in the context of the ground thermal regime in the European Alps as a test case, however, the globally consistent nature of methods and datasets employed suggest it may be applied in other geographical areas and fields of research.

The main aim of this study is to evaluate the performance of this combined scheme by:

1. conducting a test application of the combined schemes to derive land surface variables air temperature (TAIR), ground surface temperature (GST) and snow depth (SD) over a large area of the European Alps at a resolution of 30 m, and additionally a derived permafrost estimate.

2. Evaluating the performance of the combined schemes against a large network of TAIR, GST and SD measurements in the Swiss Alps,

3. interpreting results together with uncertainties in the model chain.

2 Methods

The model chain used in this study uses two previously described methods, (i) TopoSUB (Fiddes and Gruber, 2012, hereafter FG2012), (ii) TopoSCALE (Fiddes and Gruber, 2013, hereafter FG2013) together with a numerical LSM, GEOtop (Endrizzi et al.,

TCD
7, 5853–5887, 2013

2013). A brief synopsis of the tools used are given here to enable full understanding of the current study, but for further details and results of testing of these tools please see the respective publications.

2.1 TopoSUB

TopoSUB is a scheme which samples land surface heterogeneity at high resolution (here, 30 m). Input predictors describing relevant dimensions of variability are clustered with a K-means algorithm to reduce computational units in a given simulation domain (here, $0.75^{\circ} \times 0.75^{\circ}$). A 1-D LSM is then applied to each sample. For example, in FG2012 we show that reduction of a domain from 10^6 pixels to 258 samples yields comparable results to a full distributed 2-D baseline simulation. The main outcome is that the computational load is effectively reduced by a factor of 10^4, with an acceptable reduction in the quality of results. The scheme transfers model results to high-resolution pixels by membership functions (crisp or fuzzy) for a spatialised mapping of simulation results or statistical descriptions of the sub-grid domain. Additionally, we have an optional informed-scaling training routine, which regresses model results against input predictors after a training run in order to adjust the weighting of each input according to its significance, and in doing so improve the quality of final result. Limitations to this fundamentally 1-D approach include the fact that lateral mass and energy transfers can only be parameterised and not modelled explicitly. This approach allows for: (1) modelling of processes at fine grid resolutions, (2) efficient statistical descriptions of sub-grid behaviour, (3) a "sub-grid aware" aggregation of simulated variables to coarse grids and (4) enables validation of results with fine-scale ground truth.

2.2 TopoSCALE

TopoSCALE is a scheme which provides forcing to the LSM at fine-scale using gridded climate datasets. It works on the assumption that vertical gradients are often more important than horizontal in complex topography. Climate datasets are employed as they

TCD
7, 5853–5887, 2013

provide consistent fields required for LSM simulation in 3-D, therefore providing a detailed description of the atmospheric profile. In addition, they provide data with global coverage and so enable simulation in remote data-poor regions. Finally, they provide for the possibility of simulating future conditions. The basic principles of the scheme are: (1) interpolate data available on pressure levels: air temperature (Tair), relative humidity (RH), wind speed (Ws), wind direction (Wd) vertically above and below target site to provide a scaling according to atmospheric conditions at each model timestep; (2) downwelling longwave radiation (LWin) is scaled according to TAIR, RH and sky emissivity; (3) topographic correction to downwelling radiation fields (SWin/LWin); (4) lapse-rate with elevation applied to precipitation, P (optional disaggregation scheme based on climatology for site simulation only as this is spatially explicit). The final output is the full set of scaled fluxes required to drive a numerical model at 3 h timestep. It is a flexible scheme that can be used to supply inputs to models in 1-D/2-D or lumped configurations. The scheme has been shown in FG2013 to improve the scaling of driving daily fields compared to reference methods e.g. commonly used lapse rates or parameterisations.

2.3 Land surface model

The LSM used in this study, GEOtop, is a physically-based model originally developed for hydrological research. It should be noted that this model is not an LSM in the conventional sense (e.g. Mosaic, CLM, NOAH, Koster and Suarez, 1992; Dai et al., 2003), as it has not been designed to feedback to the atmosphere. However, it couples the ground heat and water budgets, represents the energy exchange with the atmosphere, has a multilayer snow pack and represents the water and energy budget of the snow cover. GEOtop simulates the temporal evolution of the snow depth and its effect on ground temperature. It solves the heat conduction equation in one dimension and the Richard's equation for water transport in one or three dimensions, describing water infiltration in the ground as well as freezing and thawing processes. GEOtop is therefore a suitable tool to model permafrost relevant variables such as snow and ground

TCD
7, 5853–5887, 2013

temperatures. It can be applied in high mountain regions and allows accounting for topographic and other environmental variability. Further information is given by Bertoldi et al. (2006); Rigon et al. (2006); Endrizzi (2007); Dall'Amico (2010). Further details specific to experiments are given in Sect. 2.5. A full description of the model is given in Endrizzi et al. (2013) and a description of model uncertainty and sensitivity is given by Gubler et al. (2013).

2.4 Model chain and modes

The model chain can be employed in two main configurations: point mode (MP) and spatial mode (MS) (Fig. 1). MS requires TopoSUB and TopoSCALE while MP requires only TopoSCALE. In terms of output, MP simulates points-scale results, whilst MS simulates a spatially-explicit mapped result from samples. The basic model chain employs TopoSCALE to derive a forcing at simulation points or samples, depending upon mode employed. The LSM simulates target variables at the computed points or sample centroids. TopoSUB is used in MS to pre-process topography and post-process results.

2.5 Snow correction method

Precipitation is highly variable in time and space, and fields computed by climate models often do not capture the frequency and/or intensity distribution of events correctly (Piani et al., 2009; Manders et al., 2012; Dai, 2006). Additionally, sub-grid topographic features may place large controls on the distribution of precipitation events (Leung and Ghan, 1998). Therefore, a method is required to correct magnitudes of precipitation inputs due to the important influence of this field on land surface processes. The method we test in this study relies on detection of the melt-date (MD) of the snow-pack, a parameter which summarizes both energy and mass inputs to the snow-pack. We vary a parameter in the model which applies a multiplicative correction on precipitation inputs called snow correction factor (SCF). We vary this parameter over the range 0.5–3 in steps of 0.25 and run a simulation for each correction factor. MD's are computed ac-

TCD
7, 5853–5887, 2013

cording Schmid et al. (2012) for each simulation and observation site using GST (which avoids circularity). We fit the simulation MD's to observed MD to obtain a correction factor for precipitation input. This method is based on cumulative winter precipitation and assumes summer and winter distributions of precipitation biases are similar, which is likely not the case. However, our primary aim is to address the thermal influence of the winter snow-pack. Please note, the SCF method is applied in the SCF experiment alone, all other results use only TopoSCALE precipitation.

3 Data

3.1 Input data

All input data used in this experiment are available free of charge, globally. This does not assume, however, that data quality is consistent globally.

3.1.1 Driving climate

Driving climate data are obtained from the ERA-Interim (ERA-I) dataset, which is an atmospheric reanalysis produced by the ECMWF (Dee et al., 2011). ERA-I provides meteorological data from 1 January 1979 and continues to be extended in near-real time. Gridded products include a large variety of 3 hourly (00:00, 03:00, 06:00, 09:00, 12:00, 15:00, 18:00 and 21:00 UTC) grid-surface fields (GRID) and 6 hourly (00:00, 06:00, 12:00, 18:00 UTC) upper-atmosphere products available on 60 pressure levels (PL) with top of the atmosphere located at 1 mb. ERA-I relies on a 4-D-VAR system which uses observations within the windows of 15:00 UTC to 03:00 UTC and 03:00 UTC to 15:00 UTC (in the next day) to initialize forecast simulations starting at 00:00 UTC and 12:00 UTC, respectively. In order to allow sufficient spin-up, the first nine hours of the forecast simulations are not used. All fields used in this study were extracted on the ECMWF reduced Gaussian N128 grid ($0.75^{\circ} \times 0.75^{\circ}$). Six PLs are used in this study

TCD
7, 5853–5887, 2013

covering the range of 1000–500 mb (1000, 925, 850, 775, 650, 500), corresponding to approximately an elevation range of 150–5500 m a.s.l.

3.1.2 Surface data

The DEM used in this study is the ASTER GDEM 2 (Tachikawa, 2011) available at approximately 30 m. Landcover was derived by a combined bedrock/debris classification which relies primarily on slope angle and a vegetation mask from a soil adjusted vegetation index (SAVI) derived from Landsat TM/ETM+ sensors. Further detail on the construction of this landcover dataset are available in Boeckli et al. (2012a). This resulted in three landcover classes, which also define sub-surface properties according to Gubler et al. (2013): (i) bedrock, (ii) coarse-blocks and (iii) vegetation. Table 1 gives a description of sub-surface properties associated with each class.

3.2 Validation datasets

The validation dataset covers a broad elevation range of 1560–3750 m a.s.l., full range of slopes from flat to vertical rock walls, full range of aspects and main alpine surface cover types, of alpine meadows, coarse-debris and bedrock (Fig. 3). The entire Alpine-space of Switzerland is well sampled by the datasets (Fig. 2). Table 2 gives an overview of each dataset.

3.2.1 IMIS

The SLF IMIS stations are used to evaluate TAIR, GST and SD. This network is biased towards high alpine locations (there are few valley stations) but represents strong topographical heterogeneity, in terms of elevation, slope and aspect. Network elevation range is 1562–3341 m a.s.l. This dataset is quite well behaved in that generally sites represent mainly elevation gradients. The dataset used covers years 1996–2011. It does not contain MAGST below zero degrees.

TCD
7, 5853–5887, 2013

3.2.2 Data loggers

The data logger dataset comprises two individual datasets and is used to evaluate GST only. The PERMOS dataset (www.permos.ch) contains data loggers of various types covering years 1995–2012 (ongoing) distributed throughout the Swiss Alps and managed by a number of institutions in Switzerland. The dataset is not homogeneous but has been compiled using consistent methods. This dataset covers a great diversity of locations but biased towards permafrost monitoring sites and therefore clustered around mean annual ground surface temperatures (MAGST) of 0 °C. In the analysis, PERMOS data is split into two groups: (a) PERMOS1, predominantly coarse debris, (b) PERMOS2: bedrock (mainly steep rock-walls). The second logger dataset, iBUTTONS, originates from a single study (Gubler et al., 2011; Schmid et al., 2012). It covers years 2010–2011 in a single region in the Engadin. While broader climatic heterogeneity is not represented by this dataset, it does cover strong topographic variability. The dataset is arranged as 10 m × 10 m "footprints" each containing 10 data loggers. In this study footprint mean values are used.

3.2.3 Data quality control

Observations outside acceptable limits were removed automatically by applying physically plausible thresholds to all datasets. Non-changing values beyond prescribed time limits were also screened out. These checks follow the methods of Meek and Hatfield (1994). Thresholds of maximum 10 % missing data in any given year qualified that year as a valid MAGST value. As datasets and sites within datasets differ in number of valid MAGST years (as defined above), validation values are computed as a mean of all available MAGST years and compared to the mean of the same modelled years.

TCD
7, 5853–5887, 2013

4 Simulation experiments

The simulation domain covers an area of approximately 500 km × 250 km, centred over the Swiss Alps (Fig. 2). The domain contains of 18 coarse-grid ERA-I boxes which supply the driving climate data. We simulate results for both MP and MS modes. TopoSUB is run at 200 sample resolution on each coarse-grid unit. The simulation period is 1984–2011. Spin-up is performed over 50 yr (10 times, 1979–1983 period). The LSM runs on an hourly timestep. LSM model parameters are fixed in all simulations as a mean value of prior distributions defined in Gubler et al. (2013). We compute mean annual air temperature (MAAT), mean annual ground surface temperature (MAGST) and mean annual snow depth (MASD). Focus is placed on mean annual values as we are primarily interested in analysing the performance of the spatial prediction of the scheme. In computing a permafrost estimate, we define a permafrost pixel as one in which the maximum daily ground temperature at 10 m depth (GT_{10}) over the entire observation period time is $\leq 0\,^{\circ}C$. Results are analysed statistically using the root mean squared error (RMSE), correlation coefficient (CORR) and mean bias (BIAS), defined as:

$$\mathrm{BIAS} = \overline{\mathrm{mod} - \mathrm{obs}}. \quad (1)$$

5 Results

5.1 Evaluation: simulated variables

Figure 4 gives MP and MS simulated results validated against measurement sites for MAGST, MAAT and MASD. MAGST results are validated against IMIS, PERMOS and iBUTTON datasets. The scheme most successfully simulates IMIS sites with low error and bias, however, there is, respectively, cold/warm bias at cold/warm sites. The iBUTTON data cover the largest range of MAGST and demonstrate good performance of the scheme in cold (i.e. MAGST < 0) locations. Both iBUTTON and PERMOS site validation shows the ability of the scheme to capture results influenced by the fine scale

TCD
7, 5853–5887, 2013

variability of the topography (Fig. 3). PERMOS1 sites (non-bedrock) are reproduced with greater success than PERMOS2 sites (steep bedrock). MP gives improved results for IMIS and PERMOS2 whereas the converse is true for IBUTTON and PERMOS1. Over all datasets an RMSE of 1.29 is obtained for MS and 1.21 for MP. However, these figures should be interpreted with caution as there is an implicit weighting based on available datapoints, which are unlikely to be representative of the distribution of underlying surfaces in the simulation domain. MAAT is well modelled at IMIS stations with low error and a bias of only 0.18 °C, although a counteracting slight warm/cold bias at cold/warm sites is visible in the data. MS and MP give statistically identical results. MASD is not captured well due to large biases in driving precipitation fields. The bias becomes more pronounced at higher values of SD. A slight improvement is seen in MP over MS. Overall, MP generally shows an improvement over MS in reproducing observations which would be expected as the spatial uncertainty introduced by TopoSUB is removed. However the difference is generally quite small (most significant difference is MAGST IMIS) and this is encouraging in that it seems MS does not introduce significant uncertainty over MP simulation. Figure 5 gives a visual impression of MS simulated MAGST results as a transect through the experiment domain.

5.2 Snow bias correction method

Figure 6 shows the effect of the snow bias correction factor on MASD. Figure 6a shows the large bias in precipitation inputs, particularly at sites with large snow accumulations (i.e MASD >50 cm). The effect of the snow correction method in successfully reproducing the spatial differentiation of precipitation quantities is shown in Fig. 6b by greatly reducing error and bias. Figure 6c gives an example which demonstrates how the method functions in obtaining a correction factor which best fits actual melt-out date of the snow-pack. A common problem is seen in Fig. 6c in that even by fitting the melt-date, snow depths can be underestimated. This could be due to the fact that SWE is reproduced accurately but parameters governing density of the snow-pack or

TCD
7, 5853–5887, 2013

wind erosion are not correct. Without SWE evaluation data however, this is difficult to confirm.

5.3 Test application: permafrost estimate

In this study we produced an estimate of 1974.9 km^2 permafrost within Switzerland based on the stated definition. Figure 7 gives a visual comparison of permafrost extent computed by this study with a state-of-the-art statistical model (Boeckli et al., 2012b) derived from Alpine specific datasets. The current method compares well in terms of spatial patterns with results of Boeckli et al. (2012b). Comparison of methods is only intended for reference as uncertainties exist based on methods or definitions of permafrost area as well as different observation periods. Boeckli et al. (2012b) is based on climate normals 1961–1990 whereas the current estimate covers the period 1984–2011. The method we have shown benefits from the simplicity of definition in actually simulating permafrost (i.e. ground $< 0\,^{\circ}C$ for more than two years), although depth of simulation remains an arguable point. In addition, Table 3 shows that the current study produces an estimate that fits a range of key estimates from the literature, well.

5.4 Macroclimatic distribution of error

Figure 8 shows the distribution of bias for all IMIS stations for TAIR, GST and SD at the macro-climatic scale. The purpose was to investigate whether there are any significant biases or sign changes of bias at the mountain range scale. Such biases would largely be a result of how well the driving climate is simulated in different topo-climatic settings e.g north or south slope of the Alps, inner Alpine-regions or west to east. TAIR bias is well distributed in sign and generally small in magnitude. There is no clear pattern in error distribution, although the north slope seems most well modelled. GST bias is well distributed in magnitude but positively biased (as shown in Fig. 8). Again north slope seems to be modelled most successfully. SD bias is very negative and error magnitude seems to fit magnitudes of precipitation i.e. greater in north-west, less in

TCD
7, 5853–5887, 2013

Engadin. However, the first stations on the north-slope appear to be modelled well. Overall, there is no clear evidence of topo-climatic gradients in error patterns at the mountain range scale.

6 Discussion

6.1 Model chain uncertainty

In order to place these results in context we provide a semi-quantitative analysis of uncertainty through the model chain. The main sources of uncertainty we identify are: (1) bias in driving fields, (2) error in scaling approach, (3) uncertainty in the lumped scheme, (3) LSM uncertainties (parameters and processes) and, (4) surface data based uncertainties (scale discussed separately).

1. Uncertainty in the driving fields can be due to bias, spatial/temporal issues or model physics and parametrisations. This issue was explored in FG2013 and reasonable to good results were reported for the variables tested. The exception being precipitation. Additionally, reanalysis datasets are expected to vary spatially and temporally with density of observations assimilated. Bias in driving precipitation is a commonly reported problem in climate models (e.g. Dai, 2006; Boberg et al., 2008) and we have attempted to address this issue with the correction method detailed in this study. Two notes of caution are worth mentioning with respect to this method, (a) this method is only valid currently at site scale, and (b) it relies on GST measurements. Other uncertainties in modelling snow precipitation lie in the definition of the snow/rain threshold, the fact that errors are cumulative over a season, and also that significant inputs are relatively infrequent, discrete events, which means that missing an event can have a large impact on season totals.

TCD
7, 5853–5887, 2013

2. In discussing uncertainty in the scaling approach (TopoSCALE) we focus on TAIR as this is the only driving variable evaluated in this study due to its importance in driving the ground thermal regime. Other driving fields were previously evaluated in FG2013. Frei (2013) reports a TAIR RMSE of 1.5 °C in the Alps using a sophisticated interpolation technique of station data. While these are daily values and cannot be compared directly to an RMSE of 0.64 obtained in this study, in FG2013 we show that TopoSCALE is able to achieve an RMSE of 1.93 on daily values. To place this in context, the method of Frei (2013) interpolates station data to model non-linearities in the vertical thermal profile together with a distance weighting scheme to account for terrain effects. Given this, TopoSCALE compares quite favourably given that only vertical profiles are modelled explicitly. In addition there are advantages of the gridded ERA dataset over interpolated station data. It should be noted that TAIR at the majority of stations are modelled at considerably lower RMSE but the overall value is affected by 4 key outliers (RMSE is sensitive to outliers), which degrade the overall result (Figs. 4 and 8).

3. Uncertainty of the sub-grid scheme (TopoSUB) has two main sources: (1) the resolution of the base DEM and, (2) the description of surface cover. The resolution of the DEM defines the range of parameter space, irrespective of number of samples computed. For example, the base DEM of 30 m in this study produces in several cases a steepest sample of under 60° whereas in reality vertical slopes exist. This has an effect on both mass and energy balance computed at such sites. In this study, surface cover is prescribed as an average value of surface characteristics within a TopoSUB sample, which are derived from a simple landcover dataset. Landcover could however be used as a predictor in sample formation if this was deemed to be important e.g vegetation mosaics that significantly affect soil moisture, wind-drift or energy balance at the surface. While surface cover is often a function of topographic predictors in the study domain, samples with complex surface characteristics e.g. vegetation, boulder, bedrock matrix will exhibit a degree of uncertainty due to the fact that all members are modelled as

TCD
7, 5853–5887, 2013

the modal surface type. Due to the significance of surface (and prescribed sub-surface) characteristics in simulating surface (sub-surface) processes, this may constitute an important source of uncertainty.

4. Uncertainties due to the LSM have three main source: (i) process description (or omission), (ii) parametrisation of processes not explicitly modelled and (iii) values given to sensitive parameters. This topic has been well discussed in the literature (e.g. Gupta et al., 2005; Beven, 1995) and so here we focus on parameter values that are sensitive and therefore have a large influence on the final result. Parameter values were fixed and taken from a distribution described by Gubler et al. (2013). The exception to this is sub-surface properties (Table 1) which vary as a function of surface type. Gubler et al. (2013) provides a thorough analysis of sensitivities and uncertainties related to parameter values used in the model GEOtop and their analysis is likely applicable to many other LSM's that have similar process description as GEOtop. In this study the authors found a total parametric uncertainty based on intensive Monte Carlo simulation of 0.1–0.5 for clay silt and rock and 0.1–0.8 for peat sand and gravel. The higher values being related to higher hydraulic conductivity of these surface classes. Therefore a portion of the error statistics given in Fig. 4 could be explained by LSM uncertainty alone.

While addressing all these sources of uncertainty within the analysis is beyond the scope of the paper, an important outcome of this work is that through the improved efficiency of simulation by several orders of magnitude, intensive simulation based uncertainty analysis starts to become feasible.

6.2 Scale issues

While scale issues are a central topic of this work in scaling between atmospheric forcing and surface simulation, another important aspect of scale mismatch arises in validation. Evaluation exercises are often carried out in the literature where model results representing cells with side lengths of 10's–100's and in extreme cases, 1000's of

TCD
7, 5853–5887, 2013

metres, are compared to point-scale measurements. In this study the PERMOS dataset is point-scale in both measurements and topographic properties upon which modelled results are based. The IMIS dataset are point-scale measurements. The iButton measurements are aggregated to a footprint mean, representing a 10 m pixel. Modelled results of both datasets are based on properties derived from a 10 m DEM. While this seems a reasonable comparison, Gubler et al. (2011) demonstrated large differences in surface conditions within such a scale domain. Smoothing of slope angles by DEM resolution, localised shading of, or snow drifting at a measurement point may cause large differences in measured and modelled conditions. Additionally, as stated above, there is a limitation based on resolution of base DEM (30 m) which under-represents steep slopes in TopoSUB sampling.

6.3 Important limitations

Key limitations are discussed in terms of (a) TopoSCALE and (b) TopoSUB. TopoSCALE based limitations primarily originate from the horizontal resolution of driving field. In this study ERA-I fields at $0.75^{\circ} \times 0.75^{\circ}$ are used. This resolution is far too coarse to represent sub-grid effects such as valley temperature inversions, for example. In addition, topographic precipitation barriers are unresolved in regions like the Mattertal (SW Switzerland) which produce important rain-shadows. Finally, spatial patterns of sub-grid effects such as shallow (mainly cumulus) convective precipitation, which is important in simulating correct precipitation intensities, only start to become resolved at resolutions of around 1 km (e.g. Kendon et al., 2012). In the case of ERA-I this is because shallow convection is parameterised by a bulk mass flux scheme as described by Tiedtke (1989) which cannot resolve the level of spatial differentiation that is present in the measurements. This process is particularly significant in spring and summer months, as surface heating occurs during a typical diurnal cycle, driving convective mass fluxes. An outlook in this respect is that the presented scheme is readily scalable to higher resolution driving climate data that will likely come into the public domain in next few years. Another key limitation of the TopoSCALE scheme is that boundary

TCD
7, 5853–5887, 2013

effects are not included in the scaling of atmospheric profiles that represent the free atmosphere. This was shown in FG2013 that the diurnal amplitude of fields such as TAIR and shortwave radiation was not as great as surface-affected measurements within the boundary layer. This may have important implications for processes which are driven by strong daily amplitudes, such as spring melt of the snow-pack.

TopoSUB based limitations are largely derived from the scale of the base DEM on which sampling is based (as previously discussed), together with the description of surface cover. However, an important limitation comes from the inherent 1-D structure of samples as simulation units. This means that all lateral processes can only be parameterised and not modelled explicitly. For example, in computing horizon angles that are important for cast-shadow calculations (Dubayah and Rich, 1995), we apply a mean horizon angle derived from the sky-view factor and local slope. This has obvious problems when the horizon is highly asymmetric i.e. consider a steep south facing mountain slope overlooking a plain. In this situation the horizon angle would be artificially raised in the southerly direction to a mean level and therefore reducing radiation inputs and consequently introducing a cold bias. This effect may also give biases in terms of reduced radiation in westerly directions under convective systems (e.g. Marty et al., 2002) or under strongly anisotropic local horizons. Another example of neglected 2-D effects is illustrated well by the PERMOS2 results (steep rock-walls, Fig. 4). The results here are negatively biased indicating that part of the energy balance is missing or poorly described. GEOtop computes the emissivity (LW) and albedo (SW) of its hypothetical surroundings as identical to that of the point itself – in this case a steep rock wall. This will reduce the SW radiation reflected from surrounding terrain, which can be a significant energy input to steep rock-walls when a winter snow-pack is present. From a mass balance perspective we do not model redistribution of snow by wind or avalanche. This has an important effect on the surface energy balance where melt dates can be several weeks later due to heavy accumulations at bases of avalanche slopes (Harris et al., 2009) or earlier on wind eroded slopes (Bernhardt et al., 2010).

TCD
7, 5853–5887, 2013

We model the loss of snow on steep slopes as a function of slope angle, however this is not a mass-conservative method as the removed mass is not redistributed.

6.4 Applications and outlook

In this study we provide a large scale permafrost model estimate as a test case. However, the scheme is generic in that it is able to generate surface fields of any variable the LSM outputs. The scheme can be used to simulate high resolution maps of current conditions as well as recent dynamics which can be used to generate estimates of near future trajectories of change. In longer term planning applications the scheme, when driven by suitable climate model data can be used to produce scenarios of site-specific future conditions. A core strength of the scheme is computation reduction which means that multiple repeat simulations are more likely to be possible. This can be utilised by producing a range of outcomes that consider significant uncertainties in the model chain and therefore a range of scenarios that should be considered in a given study. In terms of required site specific knowledge these scenarios could be interpreted to provide an improved quality of result, or in terms of uncertainty related to future conditions or other unknowns, a range of outcomes to be planned for. In terms of model evaluation the scheme has two important contributions, it provides model data at an appropriate scale for validation measurements (e.g. site of measurement). Secondly, by utilising an LSM, the scheme can generate a wide range of variables, in order to maximise use of all available evaluation data and therefore provide more robust evaluations.

7 Conclusions

This study has shown that the presented scheme is able to simulate GST and TAIR reasonably well over large areas in heterogeneous terrain, using global datasets. The poorer performance of simulated SD results is, in part, due to biases in driving data. However, we have presented a simple method that enables correction of winter precip-

TCD
7, 5853–5887, 2013

itation inputs and thus greatly improving simulation of SD, where data is available. As a test application, an estimate of permafrost distribution in Switzerland has been computed with the scheme, which is comparable to published statistical model results. However, the scheme described in this study is additionally capable of producing transient simulations, results in remote areas as well as many more useful variables besides the simple variable of distribution, such as changing sub-surface material properties (e.g. ground ice loss). This underscores a key strength of the scheme, that through efficiency gains, it allows for application of LSM's at high resolutions over large areas with transient simulation possible. This opens a great number of new possibilities in the field of land-surface change assessments in heterogeneous environments. In addition it allows model results to be validated at an appropriate scale by a wide range of measurement types due to the comprehensive set of physically-consistent outputs that are generated. We summarize the main contributions and insights of this work as the following:

- the system works well in large area simulation of the tested variables due to an efficient subgrid sampling of surface heterogeneity and scaling of driving climate.
- the system produces an estimate of permafrost area in the Swiss Alps that is comparable to statistical methods.
- all inputs are derived from global datasets, allowing for consistent application globally in heterogeneous and/or remote terrain.

Acknowledgements. We would like to thank the SLF for the IMIS dataset together with M. Dall'Amico and P. Pogliotti for compiling the IMIS stations meta-data. We would like to thank ECMWF for availability of the reanalysis dataset ERA-Interim. We thank PERMOS for GST logger dataset and work done within the TEMPS project to make this data accessible. This work was also supported by GC3: Grid Computing Competence Centre (www.gc3.uzh.ch) with customized libraries (gtsub_control and GC3Pie) and user support. This project was funded by the Swiss National Science Foundation projects CRYOSUB and X-Sense.

TCD
7, 5853–5887, 2013

References

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303–309, 2005. 5854

Bernhardt, M., Liston, G. E., Strasser, U., Zängl, G., and Schulz, K.: High resolution modelling of snow transport in complex terrain using downscaled MM5 wind fields, The Cryosphere, 4, 99–113, doi:10.5194/tc-4-99-2010, 2010. 5870

Betts, A. K.: Land-surface-atmosphere coupling in observations and models, Journal of Advances in Modeling Earth Systems, 2, 4, doi:10.3894/JAMES.2009.1.4, 2009. 5855

Beven, K.: Linking parameters across scales: subgrid parameterizations and scale dependent hydrological models, Hydrol. Process., 9, 507–525, 1995. 5868

Boberg, F., Berg, P., Thejll, P., Gutowski, W. J., and Christensen, J. H.: Improved confidence in climate change projections of precipitation evaluated using daily statistics from the PRUDENCE ensemble, Clim. Dynam., 32, 1097–1106, doi:10.1007/s00382-008-0446-y, 2008. 5866

Boeckli, L., Brenning, A., Gruber, S., and Noetzli, J.: A statistical approach to modelling permafrost distribution in the European Alps or similar mountain ranges, The Cryosphere, 6, 125–140, doi:10.5194/tc-6-125-2012, 2012a. 5861

Boeckli, L., Brenning, A., Gruber, S., and Noetzli, J.: Permafrost distribution in the European Alps: calculation and evaluation of an index map and summary statistics, The Cryosphere, 6, 807–820, doi:10.5194/tc-6-807-2012, 2012b. 5865, 5886

Dai, A.: Precipitation characteristics in eighteen coupled climate models, J. Climate, 19, 4605–4630, doi:10.1175/JCLI3884.1, 2006. 5859, 5866

Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A. S., Dirmeyer, P. A., Houser, P. R., Niu, G., Oleson, K. W., Schlosser, C. A., and Yang, Z.-L.: The common land model, B. Am. Meteorol. Soc., 84, 1013–1023, doi:10.1175/BAMS-84-8-1013, 2003. 5858

Dubayah, R. and Rich, P.: Topographic solar radiation models for GIS, Int. J. Geogr. Inf. Sci., 9, 405–419, 1995. 5870

Endrizzi, S., Gruber, S., Dall'Amico, M., and Rigon, R.: GEOtop 2.0: simulating the combined energy and water balance at and below the land surface accounting for soil freezing, snow cover and terrain effects, Geosci. Model Dev. Discuss., 6, 6279–6341, 2013. 5859

TCD
7, 5853–5887, 2013

Etzelmüller, B.: Recent advances in mountain permafrost research, Permafrost Periglac., 24, 99–107, doi:10.1002/ppp.1772, 2013. 5854
Fiddes, J. and Gruber, S.: TopoSUB: a tool for efficient large area numerical modelling in complex topography at sub-grid scales, Geosci. Model Dev., 5, 1245–1257, doi:10.5194/gmd-5-1245-2012, 2012. 5856
Fiddes, J. and Gruber, S.: TopoSCALE: deriving surface fluxes from gridded climate data, Geosci. Model Dev. Discuss., 6, 3381–3426, doi:10.5194/gmdd-6-3381-2013, 2013. 5856
Frei, C.: Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances, Int. J. Climatol., doi:10.1002/joc.3786, 2013. 5867
Gisnås, K., Etzelmüller, B., Farbrot, H., Schuler, T. V., and Westermann, S.: CryoGRID 1.0: permafrost distribution in Norway estimated by a spatial numerical model, Permafrost Periglac., 24, 2–19, doi:10.1002/ppp.1765, 2013. 5855
Gruber, S.: Derivation and analysis of a high-resolution estimate of global permafrost zonation, The Cryosphere, 6, 221–233, doi:10.5194/tc-6-221-2012, 2012. 5854
Gubler, S., Fiddes, J., Keller, M., and Gruber, S.: Scale-dependent measurement and analysis of ground surface temperature variability in alpine terrain, The Cryosphere, 5, 431–443, doi:10.5194/tc-5-431-2011, 2011. 5862, 5869
Gubler, S., Endrizzi, S., Gruber, S., and Purves, R. S.: Sensitivities and uncertainties of modeled ground temperatures in mountain environments, Geosci. Model Dev., 6, 1319–1336, doi:10.5194/gmd-6-1319-2013, 2013. 5859, 5861, 5863, 5868
Gupta, H. V., Beven, K. J., and Wagener, T.: Model Calibration and Uncertainty Assessment, John Wiley & Sons, Ltd, New York, 2005. 5868
Harris, C., Arenson, L. U., Christiansen, H. H., Etzelmüller, B., Frauenfelder, R., Gruber, S., Haeberli, W., Hauck, C., Hölzle, M., Humlum, O., Isaksen, K., Kääb, A., Kern-Lütschg, M. A., Lehning, M., Matsuoka, N., Murton, J. B., Nötzli, J., Phillips, M., Ross, N., Seppälä, M., Springman, S. M., and Vonder Mühll, D.: Permafrost and climate in Europe: monitoring and modelling thermal, geomorphological and geotechnical responses, Earth-Sci. Rev., 92, 117–171, doi:10.1016/j.earscirev.2008.12.002, 2009. 5870
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, doi:10.5194/tc-6-613-2012, 2012. 5855

TCD
7, 5853–5887, 2013

Kendon, E. J., Roberts, N. M., Senior, C. A., and Roberts, M. J.: Realism of rainfall in a very high-resolution regional climate model, J. Climate, 25, 5791–5806, doi:10.1175/JCLI-D-11-00562.1, 2012. 5869

Koster, R. D. and Suarez, M. J.: Modeling the land surface boundary in climate models as a composite of independent vegetation stands, J. Geophys. Res., 97, 2697, doi:10.1029/91JD01696, 1992. 5858

Leung, L. R. and Ghan, S. J.: Parameterizing subgrid orographic precipitation and surface cover in climate models, Mon. Weather Rev., 126, 3271–3291, 1998. 5859

Liston, G. and Haehnel, R.: Instruments and methods simulating complex snow distributions in windy environments using SnowTran-3D, J. Glaciol., 53, 241–256, 2007. 5855

Manders, A. M. M., van Meijgaard, E., Mues, A. C., Kranenburg, R., van Ulft, L. H., and Schaap, M.: The impact of differences in large-scale circulation output from climate models on the regional modeling of ozone and PM, Atmos. Chem. Phys., 12, 9441–9458, doi:10.5194/acp-12-9441-2012, 2012. 5859

Marty, C., Philipona, R., Fr, C., and Ohmura, A.: Altitude dependence of surface radiation fluxes and cloud forcing in the alps: results from the alpine surface radiation budget network, Theor. Appl. Climatol., 72, 137–155, 2002. 5870

Meek, D. and Hatfield, J.: Data quality checking for single station meteorological databases, Agr. Forest Meteorol., 69, 85–109, 1994. 5862

Piani, C., Haerter, J. O., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, Theor. Appl. Climatol., 99, 187–192, doi:10.1007/s00704-009-0134-9, 2009. 5859

Pitman, A. J.: Review: the evolution of, and revolution in, land surface schemes, Int. J. Climatol., 510, 479–510, doi:10.1002/joc.893, 2003. 5855

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, doi:10.5194/tc-6-1127-2012, 2012. 5860, 5862

Smith, M. W. and Riseborough, D. W.: Climate and the limits of permafrost: a zonal analysis, Permafrost Periglac., 15, 1–15, doi:10.1002/ppp.410, 2002. 5855

Tachikawa T., Hato, M., Kaku, M., and Iwasaki, A.: The characteristics of ASTER GDEM version 2, Proc. IGARSS 2011 Symposium, 24–29 July 2011, Vancouver, Canada, 3657–3660, 2011. 5861

TCD
7, 5853–5887, 2013

Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, Mon. Weather Rev., 117, 1779–1800, 1989. 5869

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: meeting a grand challenge for monitoring Earth's terrestrial water, Water Resour. Res., 47, 1–10, doi:10.1029/2010WR010090, 2011. 5854

TCD
7, 5853–5887, 2013

Table 1. Description of surface and sub-surface parameters used in this study.

Parameter	Unit	Bedrock	Debris	Vegetation
Residual water content	–	0	0.055	0.056
Saturated water content	–	0.05	0.374	0.431
Van Genuchten parameter, α	–	0.001	0.1	0.002
Van Genuchten parameter, n	–	1.2	2	2.4
Hydraulic conductivity	$mm\,s^{-1}$	10^{-6}	1	0.044

TCD
7, 5853–5887, 2013

Table 2. Description of evaluation datasets used in this study.

Dataset	Stations/sites	Type	Variables	Period	Coverage (ERA boxes)
IMIS	81	Station	GST/TAIR/SD	1996–2011	12
PERMOS	77	Logger	GST	1995–2012 (variable)	4
iButtons	40	Logger	GST	2010–11	1

TCD
7, 5853–5887, 2013

Table 3. Comparison of PE ($km^2 \times 10^3$) obtained by this study compared to other methods in the literature.

Author	Value	Method	Relevance
This study	1.97	numerical	global
Gruber (2012)	0.7–2.5	statistical	global
Boeckli et al. (2012a)	2.2	statistical	regional
Keller et al. (1998)	1.7–2.5	Statistical	regional

TCD
7, 5853–5887, 2013

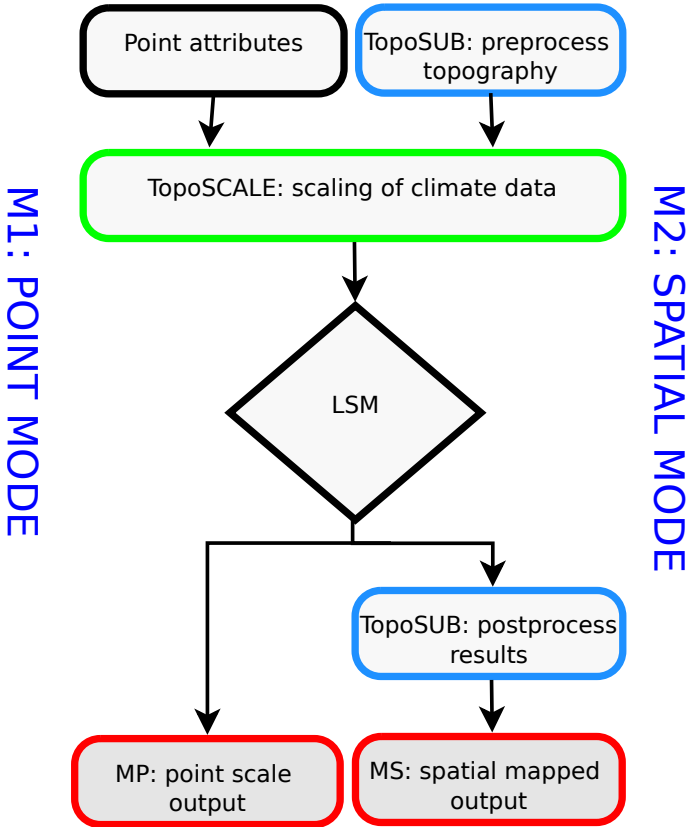

Fig. 1. Overview of how the model-chain of TopoSUB, TopoSCALE and LSM operate together. Two main modes of operation (MP) Point and (MS) spatial, are shown.

TCD
7, 5853–5887, 2013

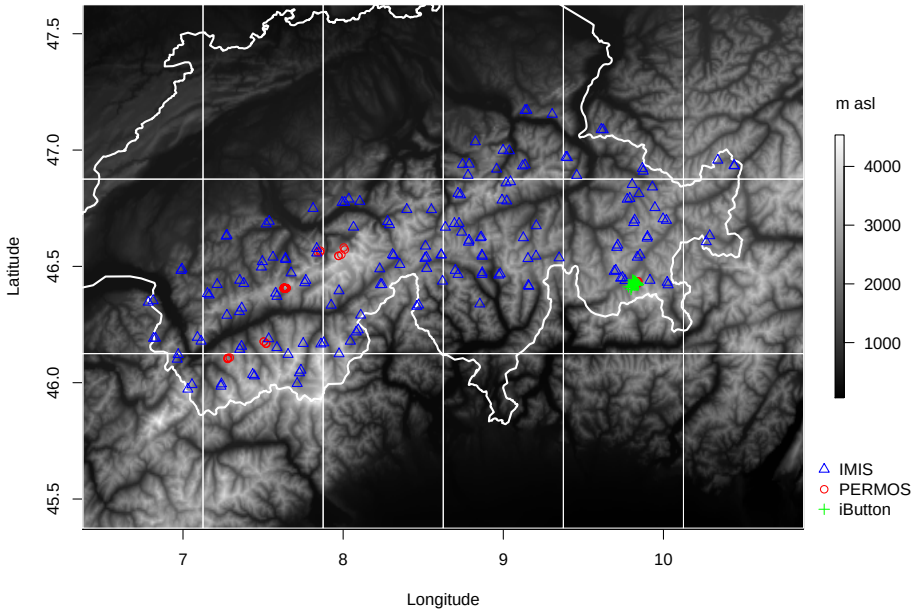

Fig. 2. Experiment domain centred on the Swiss Alps together with evaluation dataset locations. The ERA-I grid is overlaid in white.

TCD
7, 5853–5887, 2013

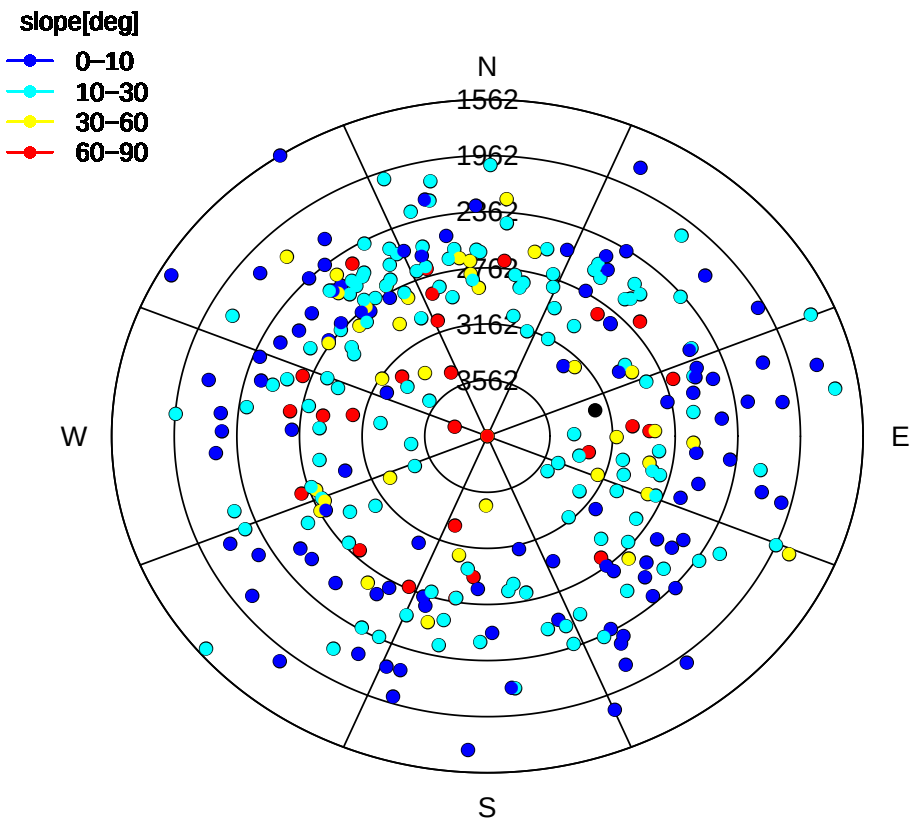

Fig. 3. Polar plots describing topographic distribution of validation sites. Elevation range 1560–3750 m a.s.l., slopes 0–> 90° and all aspects are represented.

TCD
7, 5853–5887, 2013

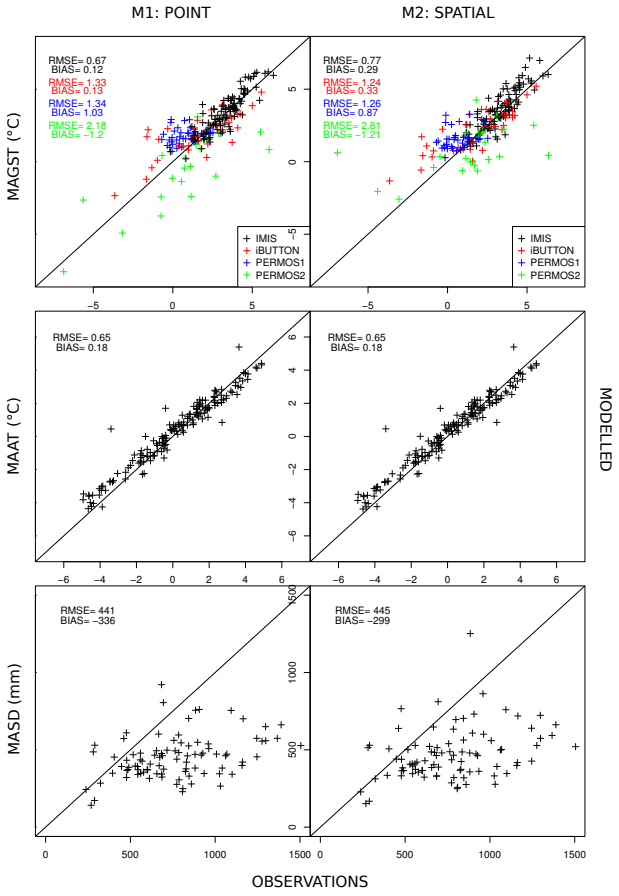

Fig. 4. A comparison of MP and MS results. Modelled MAGST, MAAT and MASD evaluated against IMIS sites (2006–2011) together with statistics.

TCD
7, 5853–5887, 2013

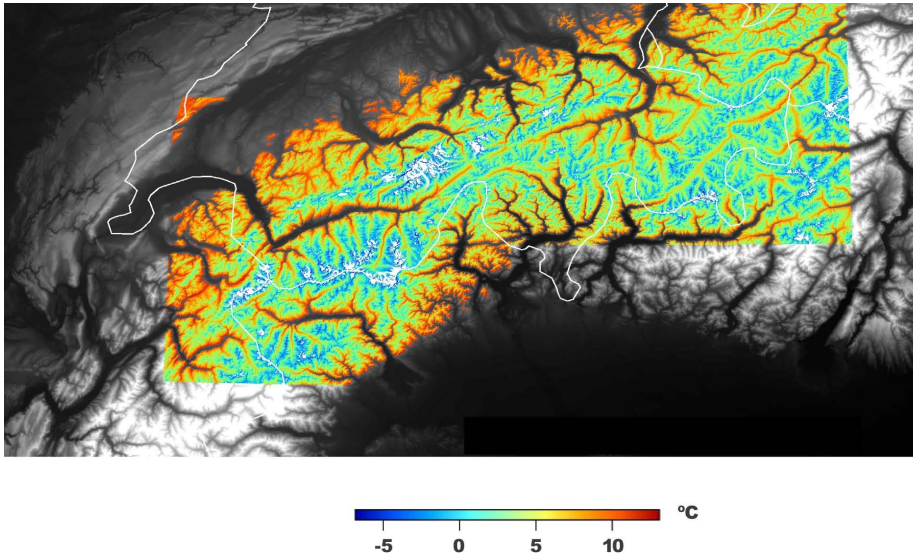

Fig. 5. A visual impression of MS spatialised results: a section of large area simulation of MAGST with glacier mask for areas above 1000 m a.s.l. (UTM zone 32° N). Switzerland's southwern border is overlaid for orientation.

TCD
7, 5853–5887, 2013

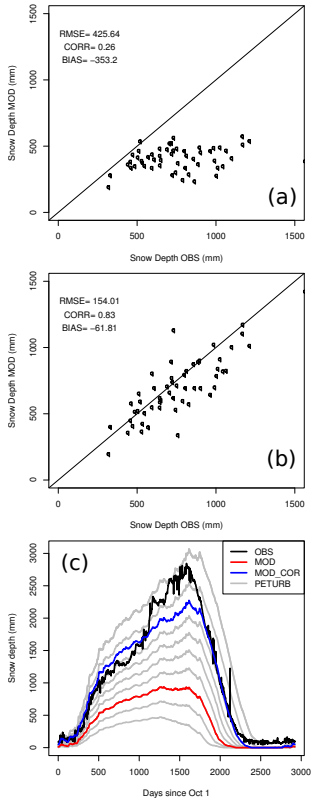

Fig. 6. Snow depth without **(a)** and with **(b)** snow correction method. **(c)** exemplifies the process at one point with improvement in modelled results by fitting the snow-pack melt date.

TCD
7, 5853–5887, 2013

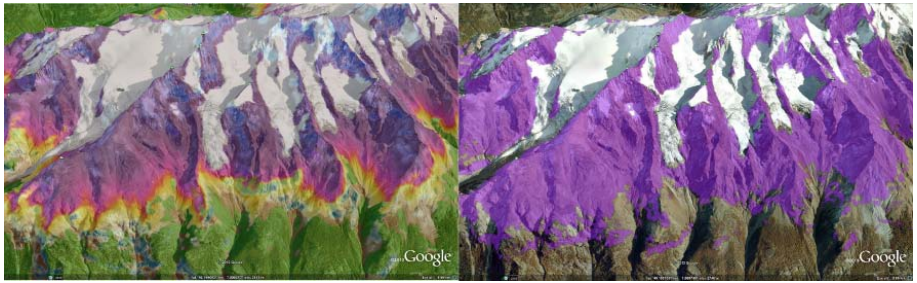

Fig. 7. Visual comparison of permafrost extent computed by **(a)** MS, this study and **(b)** state of the art statistical model (Boeckli et al., 2012b).

TCD
7, 5853–5887, 2013

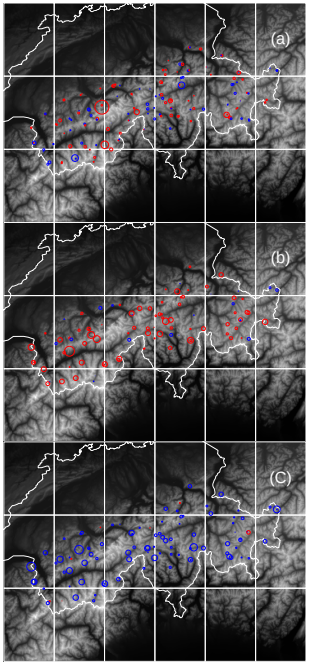

Fig. 8. Distribution of bias in **(a)** TAIR **(b)** GST **(c)** SD at the macro-climatic scale. Blue indicates negative bias (model colder/less) red indicates positive bias (model warmer/more). Size of circle indicates the relative magnitude. Sites correspond to all IMIS stations included in the analysis.

TCD
7, 5853–5887, 2013