



# Ground temperature variations in a talus slope influenced by permafrost: a comparison of field observations and model simulations

B. Staub<sup>1</sup>, A. Marmy<sup>1</sup>, C. Hauck<sup>1</sup>, C. Hilbich<sup>1,2</sup>, and R. Delaloye<sup>1</sup>

<sup>1</sup>University of Fribourg, Department of Geosciences, Geography, Chemin du Musée 4, 1700 Fribourg, Switzerland

<sup>2</sup>University of Zurich, Department of Geography, Glaciology and Geomorphodynamics Group, Winterthurerstr. 190, Zurich, Switzerland

*Correspondence to:* B. Staub (benno.staub@unifr.ch)

Received: 3 June 2014 – Revised: 7 November 2014 – Accepted: 28 November 2014 – Published: 23 February 2015

**Abstract.** Variations in surface and near-surface ground temperatures (GST) dominate the evolution of the ground thermal regime over time and represent the upper boundary condition for the subsurface. Focusing on the Lapires talus slope in the south-western part of the Swiss Alps, which partly contains massive ground ice, and using a joint observational and modelling approach, this study compares and combines observed and simulated GST in the proximity of a borehole. The aim was to determine the applicability of the physically based subsurface model COUP to accurately reproduce spatially heterogeneous GST data and to enhance its reliability for long-term simulations. The reconstruction of GST variations revealed very promising results, even though two-dimensional processes like the convection within the coarse-blocky sediments close to the surface or ascending air circulation throughout the landform (“chimney effect”) are not included in the model. For most simulations, the model bias revealed a distinct seasonal pattern mainly related to the simulation of the snow cover. The study shows that, by means of a detailed comparison of GST simulations with ground truth data, the calibration of the upper boundary conditions – which are crucial for modelling the subsurface – could be enhanced.

## 1 Introduction

Permafrost, defined as ground material remaining at temperatures below or at 0°C for two or more consecutive years (Williams and Smith, 1989), is a widespread phenomenon in the Alps covering approximately 5% (Boeckli et al., 2012) of the surface area of Switzerland. It typically occurs in locations with a cold microclimate at elevations above 2500 m a.s.l. not covered by thick glaciers (Gruber and Haeberli, 2009; Noetzi and Gruber, 2005). Except for some characteristic landforms like rock glaciers or push moraines, mountain permafrost is basically invisible from the surface and direct (and especially non-invasive) observation is therefore difficult. However, with regard to ongoing and future climatic changes, a comprehensive understanding of the processes influencing these potentially very fragile environments is essential to explain causes for recent variations and

to analyse its possible future evolution (Harris et al., 2009). Therefore this study investigates the response of mountain permafrost to selected meteorological and snow cover conditions using a combined observational and modelling approach with focus on surface and near-surface ground temperatures (GST).

### 1.1 Permafrost characteristics in high-alpine terrain

Typical permafrost landscapes in the Swiss Alps are characterised by a rough topography; often heterogeneous composition of the (sub)surface regarding the structure of the sediment; the porosity and the repartition of water, air and ice within the voids (Gruber and Haeberli, 2007; Schneider et al., 2012, 2013); and a short but spatio-temporally highly variable snow-free period. As permafrost is defined thermally, its spatial distribution (cf. Boeckli et al., 2012) is largely

influenced by all factors controlling the energy balance at the ground surface (Hoelzle et al., 2001; Stocker-Mittaz et al., 2002) and additionally by lateral and hydrothermal effects (Endrizzi and Gruber, 2012; Noetzli and Gruber, 2009; Scherler et al., 2010). In addition to topo-climatic parameters like altitude, exposition, slope angle or albedo influencing the temperature and humidity at the surface, spatio-temporal dynamics of the snow cover are also of particular importance for seasonal and inter-annual variations of GST (usually measured 5–20 cm below the ground surface) and ground temperatures (GT, e.g. measured at greater depths in boreholes) (Delaloye, 2004; Gubler et al., 2011; Zhang, 2005). A snow layer of more than  $\sim 80$  cm thickness (depending on the roughness of the terrain and the inner structure of the snow cover) is able to thermally decouple the ground from the atmosphere (Goodrich, 1982; Ishikawa, 2003; Luetsch et al., 2008; Zhang, 2005). On the other hand, a thin discontinuous snow cover of less than  $\sim 5$ –20 cm (again highly depending on the terrain roughness) may intensify ground cooling (“autumn snow effect”) on a coarse-grained surface compared to snow-free conditions due to the increased albedo and the absence of the thermal insulation (Hoelzle et al., 1999; Keller and Gubler, 1993). For almost two decades, the relative influence of the timing of the onset and melt-out of the snow cover and the corresponding efficiency of ground cooling and ground warming on mean annual ground temperatures (MAGT) have been topics of discussion (Apaloo et al., 2012; Delaloye, 2004; Zhang, 2005). Results from Schneider (2014) and others indicate that ground cooling in late autumn might be more important (for inter-annual temperature variations) due to its long-lasting effect (Delaloye, 2004; Vonder Mühl et al., 1998) over the whole winter season. But regarding the long-term permafrost evolution, increasing air temperatures in summer and in the annual mean (MAAT) might play the key role (Etzelmüller et al., 2011; IPCC, 2014; Isaksen et al., 2011; Marmy et al., 2013; Scherler et al., 2013).

Initial conditions and spatio-temporal variations in the composition of the substrate (Arenson et al., 2002; Kneisel et al., 2008), especially of the ground ice and water content (Hanson and Hoelzle, 2005; Hilbich et al., 2011; Schneider et al., 2013), are important for the response of ground temperatures to warmer atmospheric conditions (Engelhardt et al., 2010; Scherler et al., 2013). Ice-rich landforms with a thick, coarse-blocky active layer showed a slower GT increase during a simulation with warming surface conditions (Scherler et al., 2013), whereas sites with small amounts of ground ice and water may react faster to ground-heating events in terms of GT increase (Hilbich et al., 2011; Luetsch et al., 2008; Zenklusen Mutter and Phillips, 2012). Therefore general conclusions about the current state and future evolution of Alpine permafrost require comprehensive data sets, sophisticated data analysis and modelling techniques, and careful interpretation of both observations and model simulations.

## 1.2 Ground temperature modelling in the Alpine domain

The heterogeneity of the surface and subsurface material as well as the complex topography complicate long-term modelling in mountainous terrain (Engelhardt et al., 2010). In addition, many permafrost model applications are faced with disparities in scale between coarse-gridded input data (e.g. from regional climate models) and observational calibration and validation time series (Fiddes and Gruber, 2014). The high cumulated uncertainty from scenarios, downscaling issues and the complexity of real-world processes are major disadvantages of models compared to direct observations (Hawkins and Sutton, 2009). Recent studies (Ekici et al., 2014; Marmy et al., 2013; Scherler et al., 2013) applied the 1-D subsurface model COUP (see Sect. 2.4) for the field sites Schilthorn (mountain peak in the northern Swiss Alps) and Murtèl (Scherler et al., 2013) in the Upper Engadine, which are part of the Swiss permafrost monitoring network, PERMOS. The experience from the very coarse-blocky and ice-rich Murtèl rock glacier showed that many processes which were not or only partially included in the model such as convection, the lateral flow of (melt) water or the thermal radiation between blocks play an important role in reality, making subsurface modelling very challenging (Scherler et al., 2010, 2014). Against this background, further experience should be gained by applying COUP to other PERMOS study sites with different landforms, topo-climatic characteristics, and surface and subsurface properties.

Marmy et al. (2013) showed with simulations for the Schilthorn that the thermal regime of the active layer may react very sensitively to systematically increasing air temperatures, especially in autumn, when repartitioning of precipitation into rain or snow is a function of air temperature, and least during the winter months, when the presence of a snow cover is independent of air temperature. Changes in mean annual precipitation showed no clear effect on GST, but the precipitation distribution among the seasons and in general the timing and duration of the snow cover does matter (Marmy et al., 2013). However, questions still remain with regard to if and how much the stability and reliability of model simulations depend on the meteorological conditions during the calibration period. This gap in research is of particular importance because the permafrost monitoring data indicate large inter-annual variations in GST, mainly related to snow cover dynamics (Delaloye, 2004; PERMOS, 2013). Moreover, GST represents the most important upper boundary calibration parameter because any systematic bias may cumulate in the simulations. Therefore COUP should also be applied for a set of GST loggers, focusing on the performance of upper boundary conditions compared to observations.

### 1.3 Research questions and aims of this article

In this study, we apply the COUP model to simulate several GST time series in the surroundings of a borehole on a talus slope characterised by rather coarse debris. The aim is to evaluate the performance of the COUP model for the reconstruction of GST time series focusing on inter-annual and seasonal patterns, and also to improve the calibration of the upper boundary conditions for the simulation of subsurface temperatures. In particular we address the following research questions:

1. How well can the one-dimensional subsurface model COUP (see Sect. 2.4) reproduce observed ground and ground surface temperature variations over spatially heterogeneous mountain permafrost?
2. How similar are observed and modelled reactions of the ground thermal regime to selected meteorological and snow cover conditions?
3. Which processes are under- or overestimated, or even missing, in the COUP model, and how can the calibration of the upper boundary parameters be validated and improved?

## 2 Data and methods

### 2.1 Study area: the Lapires talus slope

The study comprises the analysis of observed and simulated GST and GT from the Lapires site, a large north-facing talus slope at an altitude range of 2400–2700 m a.s.l. (Valais Alps, between Verbier and Nendaz; see Fig. 1). The site has been selected because so far no talus slope has been simulated using the COUP model and based on the availability of GST and GT time series at least 10 years long (Table 1) and down-scaled reconstructed meteorological data (see Sect. 2.2). The relatively regular topography of the Lapires talus slope further limits the complexity of microclimatic effects.

Four boreholes have been drilled at the Lapires site within a number of different research projects: the drilling of the oldest borehole, LAP\_0198, was done as percussion drilling, which prohibited obtaining detailed information about the internal structure and composition. The structure of the sediments around LAP\_0198 is illustrated in Fig. 2 and was estimated by compiling data from excavations done for the construction of two cable car pylons in 1998 (Delaloye et al., 2001), the cores that were drilled in 2008 (Scapozza, 2013), the geophysical measurements (Delaloye, 2004; Hilbich, 2010; Lambiel, 2006) and the GT records: the active layer of about 4–5.5 m thickness is situated on top of an ice-rich permafrost layer of ~ 15 m thickness with temperatures very close to the melting point. The porosity within the permafrost layer ranges from 30 to 60 % and is mostly sealed by ice (Scapozza, 2013). The bedrock (gneiss) was never

reached while drilling any of the boreholes and hence must be at least at 40 m depth (Scapozza, 2013). The occurrence of permafrost within the Lapires talus slope is of discontinuous nature (see Fig. 1; Delaloye 2004; Lambiel, 2006; Scapozza, 2013) and linked to a complex system of internal air circulation (Delaloye and Lambiel, 2005). The latter is also called the “chimney effect” and most effective when the temperature (and density) gradient between the air and the voids in the porous substrate is large and may be responsible for a rapid ground cooling within a short amount of time (days/hours), in particular at the bottom of the talus slope, where cold air is aspirated in winter (Gądek, 2012; Phillips et al., 2009; Wakonigg, 1996). In contrast to scree slopes at lower altitudes, where the chimney effect is dominant for the presence or absence of permafrost (Delaloye et al., 2003; Morard et al., 2008; Wakonigg, 1996), it likely plays an important but secondary role (Delaloye and Lambiel, 2005) for the permafrost evolution at the Lapires site (cf. Table 1 for a qualitative estimate of this effect). But the thermally driven convection of interstitial air in the porous, coarse-blocky substrate during winter may be of major influence, causing very effective cooling of the uppermost ground layers (Gruber and Hoelzle, 2008; Harris and Pedersen, 1998).

### 2.2 Data sets

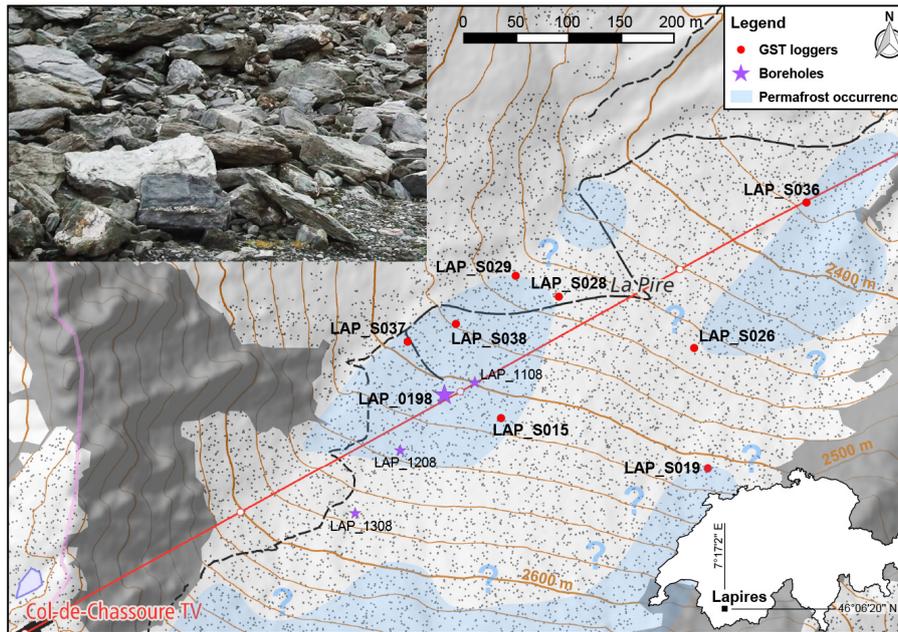
From the available data, eight GST loggers and one borehole have been selected (Table 1, Fig. 1). The GST time series used for this study were measured every 2 h with miniature temperature loggers (type UTL-1), which are characterised by an absolute measurement error of  $\pm 0.1$  K and a dynamics resolution of 0.27 K (8 bit). The loggers are placed 10–20 cm below the ground surface. During the snow melt period in spring and early summer, GSTs are close to 0 °C (“zero curtain”) due to the consumption and release of latent heat. This period is used to calibrate the sensors. GST measurements are usually representative of one specific point only or a small area and may spatially vary up to 2.5–3 °C over distances of 10–100 m (Gubler et al., 2011; Isaksen et al., 2011) because of the heterogeneity of the ground properties, snow conditions and topoclimatic effects. The samples used for this study are more homogeneous and have a mean standard deviation of 1.2 K over a spatial range of ~ 500 m (based on daily mean values from all loggers and a common period of 1717 days).

For the calibration of the COUP model runs (see Sect. 2.4), the GST time series were used at 2 h resolution. The analysis of the GST time series and the comparison with the COUP simulations are based on daily means centred at mid-day (12:00 UTC+1).

GT measurements show permafrost conditions for borehole LAP\_0198, as well as LAP\_1108 and LAP\_1208 (Scapozza, 2013). The fourth borehole, LAP\_1308, higher up on the talus slope, is permafrost-free, which confirms a local warming effect of internal air circulation (Delaloye

**Table 1.** Characteristics of the selected GST loggers and the uppermost thermistor\* of borehole LAP\_0198. The topographic parameters altitude, slope, aspect and potential incoming shortwave radiation (PISR) have been computed using SAGA GIS (Olaya, 2009). All other values are based on daily mean GST recorded during the hydrological years 2006/07, 2009/10 and 2010/11, where complete observations were available for all loggers. SD: standard deviation among all available daily mean GSTs; BTS: bottom temperature of the winter snow cover (mean GST during March); FDD: freezing degree days (sum of negative temperatures per hydrological year); TDD: thawing degree days. Chimney effect: “yes” for points with evidence of internal air circulation according to the GST records and field observations.

	LAP_S015	LAP_S019	LAP_S026	LAP_S028	LAP_S029	LAP_S036	LAP_S037	LAP_S038	LAP_0198
Sensor depth (m)	0.2	0.1	0.1	0.2	0.2	0.2	0.1	0.2	0.2*
Installed since (yr)	1999	2000	2000	2000	2000	2000	2001	2003	1999
Altitude (m a.s.l.)	2510	2505	2445	2450	2450	2375	2490	2480	2500
Slope (°)	27	30	26	23	21	15	21	19	24
Aspect (°)	10	7	29	13	56	36	37	39	13
PISR ( $\text{kWh m}^{-2} \text{yr}^{-1}$ )	905	834	1018	996	1143	1201	1129	1121	972
GST <sub>mean</sub> (°C)	0.70	1.12	1.29	1.36	1.33	-0.26	0.33	-0.45	0.55
SD <sub>mean</sub> (K)	3.70	4.22	3.91	4.17	4.14	5.92	4.29	4.50	4.19
BTS (°C)	-2.34	-2.50	-1.30	-2.29	-2.49	-6.62	-4.03	-4.03	-3.08
FDD ( $^{\circ}\text{C d yr}^{-1}$ )	-351	-361	-238	-276	-291	-942	-514	-752	-452
TDD ( $^{\circ}\text{C d yr}^{-1}$ )	606	769	709	773	778	847	633	589	655
Chimney effect	yes	yes	yes	?	?	yes	?	yes	yes
Snow-free date	18 June	26 May	12 June	20 June	11 June	02 June	30 June	17 June	20 June

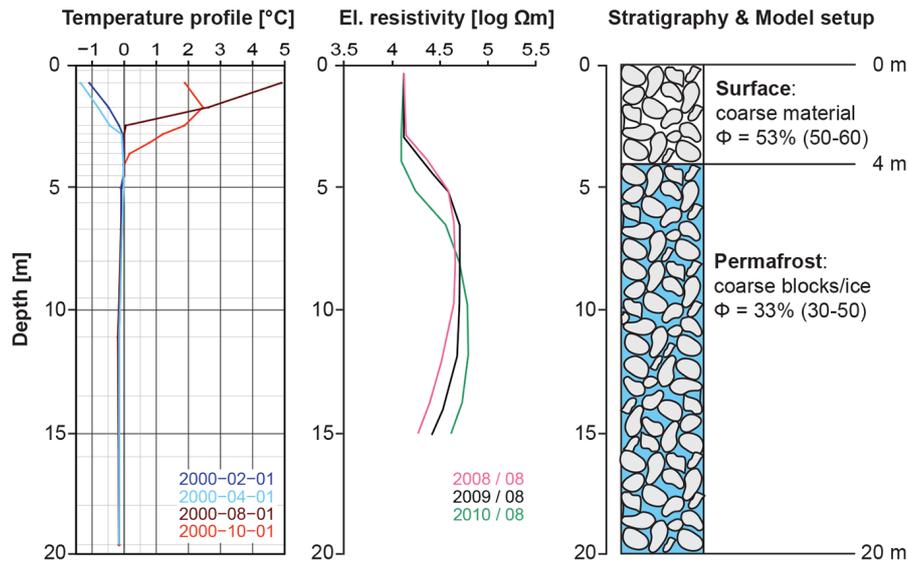


**Figure 1.** The Lapires study site is located in the Lower Valais in the south-east of Switzerland in the ski area of Nendaz/Verbiers. The map shows the positions of four boreholes (among which LAP\_0198 was used for this study) and eight GST loggers, as well as a qualitative indication of the spatial extent of permafrost occurrence according to Delaloye (2004), Lambiel (2006) and Scapozza (2013). The photograph was taken close to borehole LAP\_0198 and provides a closer look at the surface characteristics, which mainly consist of coarse blocks but also some finer material. Reproduced with permission of Swissstopo (BA14056).

and Lambiel, 2005; Scapozza, 2013). As the simulation with COUP requires long calibration time series, we focus on borehole LAP\_0198. Because the GT variations recorded within the permafrost layer of LAP\_0198 are very small (of the order of the measurement uncertainty of the old thermistor chain in place until 2009; since then, sensors of the type “YSI 44031” have been installed) and close to the melting

point, these data needed to be checked for plausibility and have been corrected or removed where necessary.

Meteorological data for this site are available through a local weather station at borehole LAP\_0198, which has been measuring air temperature and shortwave radiation since 1998, as well as wind and snow depths since 2009 (with interruptions). Precipitation, which is needed as forcing data



**Figure 2.** Vertical profile of GT and electrical resistivities according to PERMOS (2013) for borehole LAP\_0198 and the stratigraphy (Scapozza, 2013) on which basis the initial model setup for LAP\_0109 was created.

for the COUP model, is not measured on-site. In order to have continuous and physically consistent meteorological input data for the COUP model for all the above-mentioned parameters, a reconstruction has been driven in the framework of the ongoing project “The Evolution of Mountain Permafrost in Switzerland” (TEMPS, 2011–2015). Using an empirical–statistical downscaling approach (Rajczak et al., 2015), the bias correction and spatial transfer from RCM output data to the point of interest is done within two steps over a most representative station with long-term measurements (MeteoSwiss station “Grand-St. Bernard” in the case of Lapires). The reconstructed data reproduce the general behaviour of the on-site meteorological parameters well (performance for air temperature based on  $\sim 3500$  days of observations:  $R^2 = 0.97$ ,  $RMSE = 1.23$  K), but they may be less accurate when analysing single meteorological events, especially for precipitation. For qualitative comparison with the simulated snow pack, snow depth records were taken from the closest snow station of the IMIS/ENET network (ATT-2, “Les Attelas”; source: SLF), which is at the same elevation level as the GST loggers but  $\sim 1.5$  km to the west of the study site in an almost flat area (usually not influenced by avalanches, below a slope exposed to the north-west).

### 2.3 GST aggregates and indices

Daily, seasonal and inter-annual temperature variations are most prominent at the ground surface and become buffered and delayed at depth (Goodrich, 1982). Distributed and continuous GST measurements are therefore a good data source to analyse meteorological effects on the site-specific ground thermal regime at various spatial and temporal scales (Delaloye, 2004; Fiddes and Gruber, 2012; Gubler et al., 2011;

Schmid et al., 2012). The comparison between sites and years, as well as between simulated ( $GST_{sim}$ ) and observed ( $GST_{obs}$ ) GST, can be simplified by using anomaly values (inter-annual deviations from the mean during a common reference period, further identified with the annex (a). In addition, aggregates over specific time periods (e.g. monthly and annual means) and time series of the model bias (sim-obs) and running Pearson correlation are used.

To investigate inter-annual temperature variations, we calculated running annual means for GST ( $rMAGST$ , Eq. 1), air temperature ( $rMAAT$ ) and GT ( $rMAGT$ ) on the basis of daily mean values, aligned at the end (date  $t$ ) of a 365-day period, in the form

$$rMAGST_t [^\circ C] = \frac{GST_t + GST_{t-1} + \dots + GST_{t-364}}{365}. \quad (1)$$

As a proxy for the thermal decoupling and the direction of the heat flux between the ground and the atmosphere, daily ( $SO$ , Eq. 2) and annual surface offsets ( $rMASO = rMAGST - rMAAT$ ) were calculated.

$$SO_t [K] = GST_t - AirT_t \quad (2)$$

To compare the cumulative temperature evolution for different years, cumulative degree day sums (CDD) were calculated from daily mean GST for each day ( $t$ ) and hydrological year (hy, starting on 1 October) as described in Eq. (3):

$$CDD_{hy,t} [DD] = \sum_{t=Oct-01}^{Sep-30} GST_t. \quad (3)$$

CDDs have been computed for  $GST_{obs}$  and  $GST_{sim}$ , which further allow the calculation of the cumulative model bias ( $CMB = CDD_{sim} - CDD_{obs}$ ).

On the “zero curtain” (Sect. 2.2) the timing of the snow melt in spring can be determined (Schmid et al., 2012), and because of the insulating effect of the snow, the first arrival of a thermally insulating snow cover in early winter can also be approximated. For this study, GST time series have been considered as “snow-covered” as far as the weekly standard deviation undercut the dynamic resolution.

#### 2.4 COUP model calibration and parameterisation

The COUP model is a numerical one-dimensional model coupling the heat and subsurface water transfer processes using the general heat flow equation (Jansson and Karlberg, 2004; Jansson, 2012) together with several empirical equations to model various subsurface and snow parameters. The model has already been used for a variety of subsurface conditions, and has also been applied in Alpine permafrost studies (e.g. Ekici et al., 2014; Marmy et al., 2013; Scherler et al., 2010).

The COUP model settings for this study consist of 50 vertical layers to a depth of 20 m with increasing thickness of model layers with depth. For the analysis of the GST, only the first layer was considered (thickness = 20 cm), but the whole profile was simulated to take into account the influence from subsurface processes. The upper boundary conditions are calculated by a complete energy balance calculation at the ground surface (cf. Scherler et al., 2014) and are drastically influenced by the snow conditions that temporarily decouple the ground from the atmosphere. The lower boundary conditions are calculated from the analytical solution of the sine variation at the ground surface and a mean value for the damping of the temperature signal for the whole profile of the subsurface. The geothermal heat flux is negligible as only surface processes are being analysed and the timescales considered are short. The subsurface has been divided into two main horizons (see Fig. 2). The data available from the temperature measurements, geophysics and the cores that were drilled in the surroundings in 2008 (PERMOS, 2013; Scapozza, 2013) provide indications about the porosity, although their values all include a large uncertainty. Therefore we defined ranges of plausible porosities for both horizons (50–60 % for the surface and 30–50 % at depth) according to available information (Sect. 2.1) and let the model choose the most suitable values with the GLUE method (see below). As shown in Fig. 2, the resulting profile is (a) from the surface to 4 m depth with a porosity of 53 %, representing the coarse blocky material, and (b) from 4 m to the bottom (20 m) with a lower porosity (33 %).

The model is run with the reconstructed meteorological data set at daily resolution (available for 1981–2013; see Sect. 2.2), with a spin-up of 3 years for the GST simulations and a longer spin-up of 20 years for borehole LAP\_0198. This spin-up procedure is sufficient to have the model in equilibrium because the initial conditions are estimated by the model.

In addition to the meteorological forcing data set and the setup of the vertical profile of the subsurface, the COUP model simulations depend on a large number of parameter settings, which are usually not known exactly for a specific site, and are partly determined by calibration (Jansson, 2012). In our study, the calibration has been driven by inverse modelling using the general likelihood uncertainty estimation (GLUE) method (Beven and Binley, 1992). The GLUE method tests different sets of parameter values chosen stochastically among a physically reasonable range. The analysis of the equivalence between model results and observations enables the set of parameter values giving the best fit with the observed system to be selected. In this study, a setup of 20 000 (for LAP\_0198) and 10 000 (for the GST) different parameter combinations has been tested and the performance has been analysed statistically using the coefficient of determination and the mean model bias in temperature to reach the best fit with the observations. The parameters used for the calibration are summarised in Table 2. This approach may lead to the equifinality problem creating a substantial uncertainty if the model is used for prediction (Beven and Freer, 2001) when the simulation period is longer than the calibration period. For the present study, the optimal parameter combination has been selected.

The snow parameters are hereby crucial for both the GST and the borehole calibration as they influence the upper boundary conditions. Among them, CritSnowCoverDepth specifies the snow thickness at which the surface is considered fully covered by snow. Below this threshold value, the ground surface is considered only as partly snow-covered and the surface temperature and the surface albedo are then calculated as a ratio between the values for bare ground and snow-covered ground (Jansson and Karlberg, 2004). The density of the snow is dynamic and COUP simulates the snow density using the density of new snow (DensityNewSnow) and a compaction rate depending on the depth of the total snow cover. The density also influences the thermal conductivity of snow. The melting of the snow is influenced by the global radiation (MeltCoefGlobalRad), the air temperature (MeltCoefAirTemp) and the heat flow from the ground.

In addition to the snow parameters, the parameters influencing the albedo have also been used for calibration. The albedo of snow is a function of snow surface age and can be tuned by setting a differential minimal albedo value for old snow (AlbSnowMin). The albedo of the bare ground is a function of pressure head (i.e. water content) and limited by a maximum (AlbedoDry) and a minimum (AlbedoWet) value.

For the subsurface the following parameters were determined by inverse modelling: the porosity (Porosity) and the hydraulic conductivity parameter (MatrixConductivity and TotalConductivity). The parameters not used in the GLUE calibration procedure were set to literature values (Jansson and Karlberg, 2004).

**Table 2.** List of calibration parameters and their values obtained by inverse modelling. Concerning the hydraulic conductivity parameters (set for borehole LAP\_0198): (a) value assigned to the upper 4 m of the vertical profile and (b) value assigned from 4 m depth to the bottom of the vertical profile.

	LAP_S015	LAP_S019	LAP_S026	LAP_S028	LAP_S029	LAP_S036	LAP_S037	LAP_S038	LAP_0198
Snow parameters									
CritSnowCoverDepth (m)	0.15	0.17	0.21	0.28	0.13	0.21	0.25	0.30	0.30
DensityNewSnow (kg m <sup>-3</sup> )	145.59	144.79	175.52	105.07	121.52	175.77	185.12	176.90	133.90
MeltCoefGlobalRad	$6.5 \times 10^{-9}$	$2.5 \times 10^{-8}$	$1.6 \times 10^{-7}$	$1.9 \times 10^{-7}$	$1.2 \times 10^{-7}$	$2.1 \times 10^{-6}$	$1.9 \times 10^{-6}$	$1.6 \times 10^{-6}$	$7.2 \times 10^{-8}$
MeltCoefAirTemp	1.68	2.84	1.48	0.64	1.37	0.90	2.64	1.22	0.90
Albedo parameters									
AlbSnowMin (%)	20.89	55.89	24.83	42.52	45.32	49.88	55.09	43.88	26.92
AlbDry (%)	68.63	71.93	75.46	45.23	61.00	79.42	77.73	77.32	55.92
AlbWet (%)	18.29	13.24	23.29	18.04	11.23	7.89	12.17	7.35	17.61
Hydraulic conductivity parameters									
Porosity (%)									53 <sup>a</sup> /33 <sup>b</sup>
MatrixConductivity (mm day <sup>-1</sup> )									13 225 <sup>a</sup> /98 412 <sup>b</sup>
TotalConductivity (mm day <sup>-1</sup> )									13 225 <sup>a</sup> /98 412 <sup>b</sup>

### 3 Results

#### 3.1 Meteorological events and their influence on GST

We define “meteorological events” as meteorological conditions that show distinct effects on rMAGST. We propose to use the first derivative of rMAGST as a non-universal attempt at numerical detection. Rapid changes in rMAGST clearly exceeding the standard deviation of the whole time series in a moving time window of 90 days indicate specific meteorological and/or snow cover conditions. Such events can occur within the time frame of a few days (e.g. snow or rain events), weeks (e.g. summer heatwave or dry fall conditions) or even several months (e.g. air temperature anomalies, perennial snow patches from avalanche deposits), but their effect on the ground thermal regime might be of very different intensity and persistence. Meteorological events are part of the natural variability of the system with a high recurrence, but because of analogies to possible future climate scenarios, they are often discussed within the same scientific context (e.g. Marmy et al., 2013; Schär et al., 2004). One should be aware that changes in running means depend on the averaging time window and are less sensitive to consecutive anomalies of the same algebraic sign (causing, for example, the intensive ground cooling in winter 2005/06 to be masked by the precedent cold rMAGST period; see Fig. 3b).

Figure 3a shows inter-annual variations in rMAGSTa, rMAATa and running annual surface offset (rMASO; see Sect. 2.3) for the period 2000–2012 at the Lapires field site, as well as snow depth records from the nearby snow station ATT-2 (see Sect. 2.2). Although Fig. 3 shows the results for one location and the temporal evolution of the snow cover can be spatially heterogeneous, the temporal behaviour of rMAGSTa is very similar in many parts of the Swiss Alps (PERMOS, 2013).

The observed inter-annual rMAGSTa variations are in the same range as those of rMAATa, but with a significantly different pattern of alternating warmer and colder periods (Fig. 3a). The surface offset rMASOa (rMAGSTa–rMAATa, cf. Eq. 2) can be explained in large part by inter-annual variations in snow conditions (Fig. 3c). Positive rMASOa indicates a net warming influence of the snow cover (usually due to early and/or much snow in the first half of the winter, or early melt-out in spring). Negative rMASOa typically occur in years with late and/or limited snow fall or delayed snow-melt. Figure 3 also shows that meteorological and mainly snow cover conditions in autumn and early winter are of particular importance for the variations in rMAGSTa (which explains, for example, the earlier increase of rMAGSTa compared to rMAATa in event E3, cf. Table 3), because these effects may cumulate during the whole winter (Delaloye, 2004). Eleven events covering a wide range of considerably different meteorological and snow cover conditions are highlighted in Fig. 3b and further described in Table 3. The performance of the GST simulations will be analysed in the following sections regarding these events, in particular regarding potential relations between good/bad model performance and the various types of meteorological events.

For all events listed in Table 3 (except for E7 and E9), the timing of the snow cover played a major role: in the case of E1, E3, E8 and E10, early snow fall in autumn limited ground cooling, whereas the events E2, E4 and E6 tended to be cooler due to a late onset of an insulating snow cover. The date of the melt-out in summer was also partially responsible for the warmer (E3, E5, E10) and colder (E2, E4, E11) periods. Air temperature has a likewise high impact on the ground thermal regime, as it (1) influences the melt rate of the snow cover in spring, (2) directly heats (or cools) the ground during the snow-free season and (3) modifies the thermal gradient between the ground and the snow surface (E3, E6–E10).

**Table 3.** Selected meteorological events during the observation period with significant influence on GST.

Event	Type	Start	End	Meteorological and snow conditions
E1	warming	2000-09	2001-06	Early snow fall in September 2000 and heavy precipitation event in October 2000 lead to limited ground cooling during winter 2000/01
E2	cooling	2001-07	2002-06	Late snow melt in 2001 and late snow fall in 2001/02 (dry November and December 2001)
E3	warming	2002-08	2003-08	Short snow event in mid-summer 2002, early snow in 2002/03, early melt-out and heatwave in summer 2003
E4	cooling	2004-06	2005-05	Late snow melt in summer 2004 and late snow fall in winter 2004/05, relatively cold winter air temperatures
E5	warming	2005-05	2005-08	Early snow melt in June 2005
E6	warming	2006-06	2007-06	Record-breaking warm air temperatures between June 2006 and June 2007, but late snow in 2006/07
E7	cooling	2007-07	2007-12	Cold period between July and December 2007
E8	warming	2008-09	2009-09	Early snow in October 2008 limited ground cooling, but moderate snow depths in January 2009 coinciding with cold air temperatures, very hot spring and warm late summer 2009
E9	cooling	2010-09	2010-12	Cold air temperatures and little precipitation in September and October 2010
E10	warming	2011-06	2011-11	Early snow melt already by end of May 2011 and warm September and November 2011
E11	cooling	2012-05	2012-09	Large snow falls in late winter 2011/12 and delayed melt-out in 2012

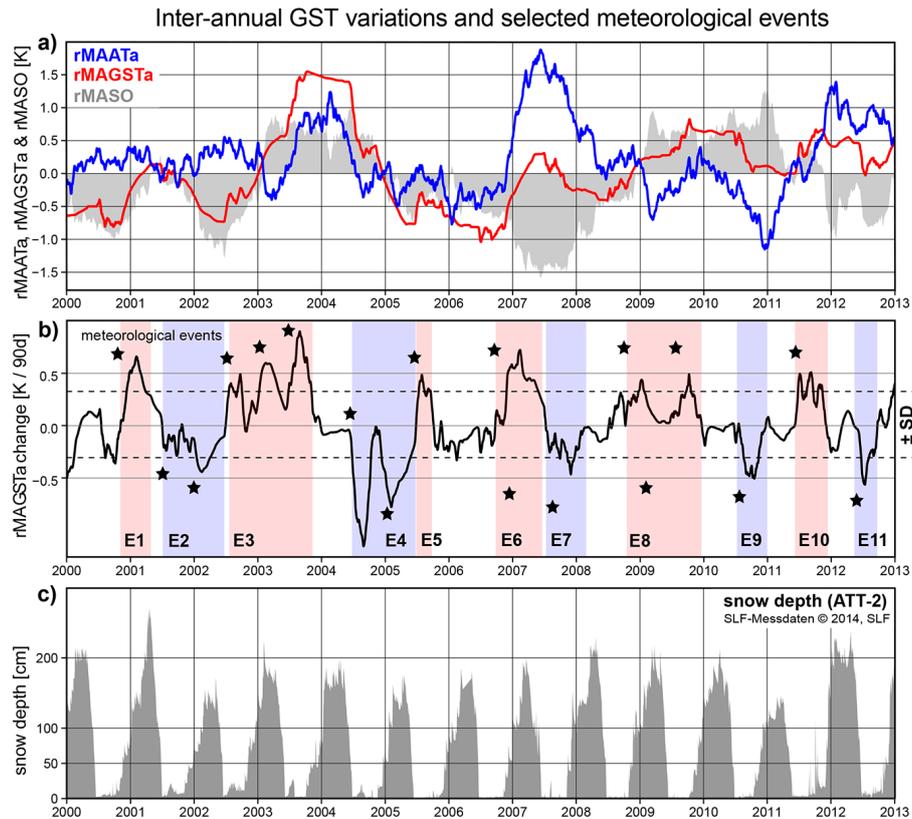
### 3.2 Comparison of observed and simulated inter-annual GST variations

For a first comparison,  $GST_{sim}$  values have been plotted against  $GST_{obs}$  for the uppermost thermistor of borehole LAP\_0198 and eight GST loggers, and completed with summary statistics (Fig. 4). Even though the results look fairly good at a first glance (mean  $R^2 \sim 0.9$ ), there are some prominent simulation errors, in particular during the snow melt phase (scatters along  $0^\circ\text{C}$ ). The causes for these errors can either be related to (1) the microclimate (differing from the reconstructed meteorological data that has been adjusted to match meteorological conditions at LAP\_0198), (2) particularities regarding the dynamics of the snow cover (e.g. modification by wind, avalanches, skiers) or (3) other local characteristics related to the before-mentioned ventilation effects or subsurface properties.

As shown in Fig. 5a for logger LAP\_S015,  $GST_{sim}$  and  $snow_{sim}$  closely follow the observations during the whole period between 1999 and 2013. At first glance, the model seems capable of reproducing strong GST variations over short timescales. Warming events (e.g. E5, E6 or E10), as well as some cooling events (e.g. E2, E4 or E11), are also well represented. Regarding running annual means (Fig. 5b),

most observed and simulated inter-annual variations are very similar, despite a slightly larger bias during the events E3 and E7. The logger LAP\_S015 shown in Figs. 5a and b was mainly selected because it is not interrupted by gaps and situated close to borehole LAP\_0198. In Figs. 5c and d, the model bias of LAP\_S015 is compared with the other loggers regarding the temporal evolution of the mean 365-day model bias and the Pearson correlation coefficient (running 365-day  $R^2$  between  $GST_{sim}$  and  $GST_{obs}$ ). For all GST loggers the mean model bias ranges between  $-1$  and  $+2$  K. The temporal behaviour is similar for many loggers, but over the years no systematic trend can be observed. Pearson correlation is generally high (mean: 0.92), but weaker between summer 2001 and 2003 (several years with delayed melt-out in a row), 2007 and 2009 (bias in the melt-out day and weak model performance in February–May 2009), and 2012 and 2013 (delayed melt-out in 2012; see Fig. 5d). However, the similarity between different loggers may be related to (1) the meteorological input data (mainly precipitation), (2) the model parameterisation or (3) effects which cannot be simulated by the model at all (such as snow redistribution).

Because the spatio-temporal variability of the snow cover is a key parameter for both the seasonal and inter-annual GST variations and has been identified as the source of se-



**Figure 3.** The evolution of air temperature, GST and snow cover at Lapieres: (a) rMAGST (mean of the two longest GST series LAP\_S015 and LAP\_S028) and rMAAT (based on the reconstructed air temperature) anomalies and their difference (rMASO) over time. (b) Meteorological events (stars E1–E11) identified by the change in rMAGSTa (running 90-day difference) are indicated by red (warming effect) and blue (cooling effect) bars and are further described in Table 3. (c) Temporal evolution of the snow cover at the nearby snow station ATT-2 (“Les Attelas”; source: SLF). Running means are aggregated at the end of the period; the date scales indicate 1 January.

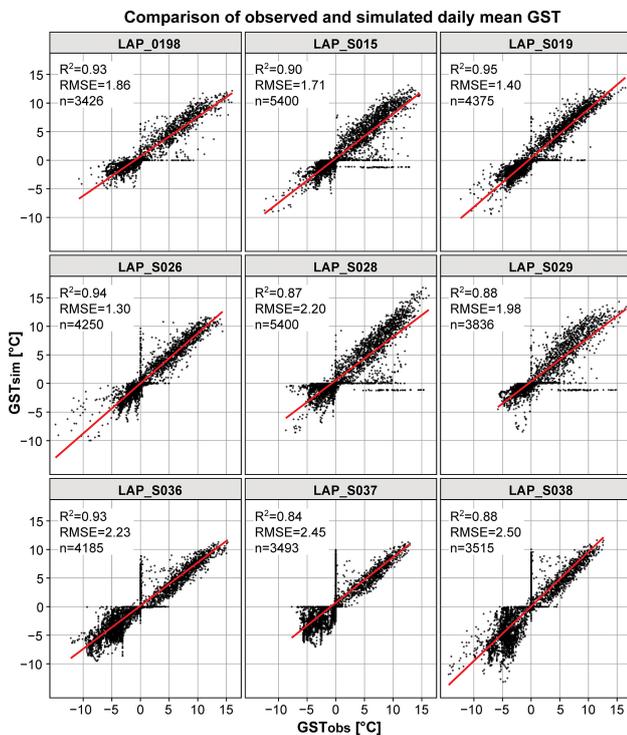
vere model biases in other modelling studies (e.g. Ekici et al., 2014), special attention is devoted to (1) the timing of the onset and melt-out of the snow cover and (2) its thermal insulation effect, which mainly depends on the thickness, density and structure of the snow layer but also on the roughness of the terrain. Simulated snow depths correlate well with the observed snow depths at Les Attelas (for LAP\_S015:  $R^2 = 0.88\text{--}0.97$ ). But because the snow depths are recorded in a place with different topography and microclimate, this comparison is of limited validity. More reliable is information about the timing and insulation effect of the snow cover directly derived from the GST time series: Table 4 compares the timing and duration of the snow-covered period (cf. Sect. 2.3) and the snow melt in spring. While for some loggers the timing of the simulated snow melt period is very close to the observations (e.g. for LAP\_0198, LAP\_S015, LAP\_S029), for others the simulated melting period ends several weeks too early (mainly for LAP\_S036, LAP\_S037, LAP\_S038). This may lead to substantially wrong GST<sub>sim</sub> in spring and early summer, which also explains the large variations in model bias and  $R^2$  (see Figs. 4, 5c and d). The causes can be related to the redistribution of snow by wind or

avalanches affecting the observations as well as the parameterisation of the model. Because the GST difference between the snow-covered state and the first days after melt-out is of the order of 5–10 K, a delayed melt-out of as little as 5 days in the simulation can cause a model bias of approximately +0.1 K with respect to the annual mean. Early and late starting of the snow melt has less effect on the annual model bias, as the mean error is usually smaller than 2–3 K each day. The duration of the snow melt period tends to be overestimated by the model (mean: +5 days), which can either be due to model parameterisation (e.g. melt coefficients, snow albedo or suboptimal model layer structure), or processes of snow redistribution and compaction (avalanches, wind, skiers) not simulated in the model. The onset of the winter snow cover is generally better reproduced by the model than the melt-out. Regarding snow depths and the amount of days with a given snow depth (Table 4), the simulations look plausible for LAP\_S015, LAP\_S019, LAP\_S026, LAP\_S028 and LAP\_S029. But for the other loggers as well as for the borehole, the simulated snow depths seem rather too shallow.

In summary, the model bias is neither constant nor increasing over time but highly depends on the simulation of the

**Table 4.** Timing and duration of the observed and simulated snow cover (obs|sim) for the three hydrological years with complete data: 2006/07, 2009/10 and 2010/11 (average values representing the day of the year starting on 1 January and the duration in days, respectively). Since no direct observations exist for the GST loggers, only simulated snow depths (average between November and May and the standard deviation) are indicated.

Snow cover characteristics	LAP_S015	LAP_S019	LAP_S026	LAP_S028	LAP_S029	LAP_S036	LAP_S037	LAP_S038	LAP_0198
Start snow melt	118 104	113 114	109 103	114 102	109 103	114 112	116 115	114 118	113 106
End snow melt	170 171	147 151	164 154	172 170	163 161	154 132	182 137	169 137	172 173
Snow melt duration (days)	53 68	35 38	56 52	59 69	55 59	41 21	67 23	56 20	60 68
Onset winter snow cover	320 317	320 326	328 326	317 325	319 318	339 326	325 318	323 325	336 318
Snow <sub>Nov–May</sub> depth (cm)	64 ± 31	56 ± 34	57 ± 32	68 ± 33	65 ± 33	27 ± 26	27 ± 26	29 ± 26	28 ± 14
Days with snow depth ≥ 30 cm	196	164	171	197	187	91	93	95	145



**Figure 4.** Comparison of  $GST_{sim}$  and  $GST_{obs}$  based on daily mean values for the uppermost thermistor at LAP\_0198 (20 cm) and eight GST loggers (cf. Table 1).

snow cover. The inter-annual pattern of the model bias is similar for most loggers but does not seem to be directly linked to specific meteorological conditions.

### 3.3 Comparison of observed and simulated seasonal GST variations

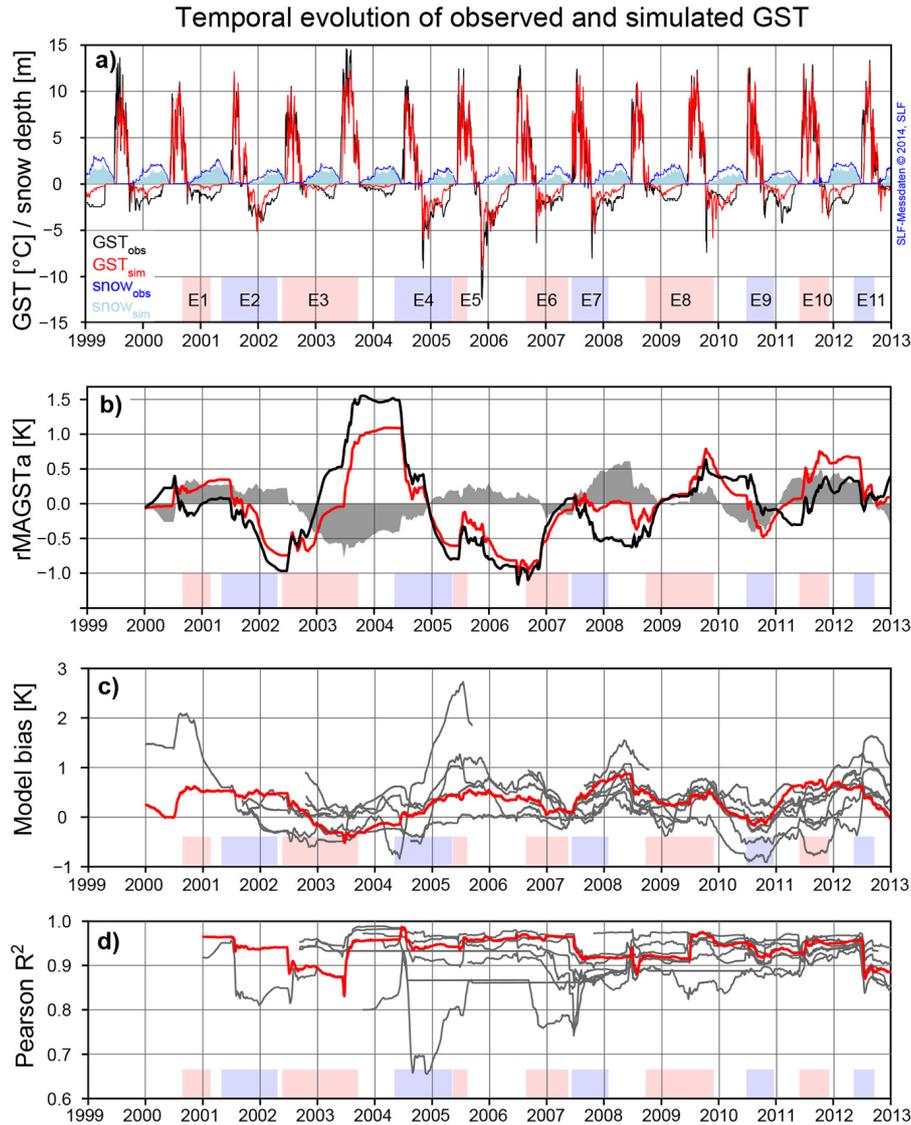
Figure 6 illustrates the seasonal variability of the model bias, and it shows that the strongest deviations occur (1) during the second part of the winter and (2) after the snow melt period. Even though the variability between different (cold/warm) years is low from February to April, the model bias remains positive between +0.5 and +2 K, whereas during the snow-free period the variability is much larger, but with a

smaller mean model bias. Loggers LAP\_S036, LAP\_S037 and LAP\_S038 show different behaviour, mainly because the timing and duration of the snow melt is not well captured by the model (see Sect. 3.2) and the snow depths seem to be underestimated (short-time variations are unrealistically large). These loggers are situated in areas of higher potential incoming radiation (cf. Table 1), which might be a first explanation for the different seasonal patterns (Fig. 6) and the larger RMSE (Fig. 4). In addition, LAP\_S036 is situated at the bottom of the talus slope in a zone of probably intense internal air circulation (Delaloye and Lambiel, 2005).

In summary,  $GST_{sim}$  values are systematically too warm between February and May, changes in the model bias are largest before and after the snow melt, and for most loggers the model bias increases between July and November. The observed seasonality of the model bias requires further investigation (see Sect. 4.2), especially regarding (a) the responsible processes and (b) a potential cumulative effect over time. The latter can be analysed in terms of the cumulative GST model bias (CMB; see Fig. 7). For most years and loggers the model bias cumulates to positive values for individual hydrological years (which equates to an overestimation of the simulated temperatures), with maxima around 500 degree days (DD) between April and July. The mostly negative model bias in summer is partly able to compensate for the positive bias during the rest of the year. For the two time series LAP\_S015 and LAP\_S028, negative CMB occurred in 2002/03 (event E3) and 2010/11, mainly due to underestimated GST during the snow-free period. A possible explanation would be overestimated subsurface water content, because the ground was very dry in summer 2003. The most positive CMB occurred in 2011, due to a strong overestimation of the winter GST and a much too early start of the snow melt (2–3 weeks).

### 3.4 Comparison of observed and simulated ground temperatures at depth

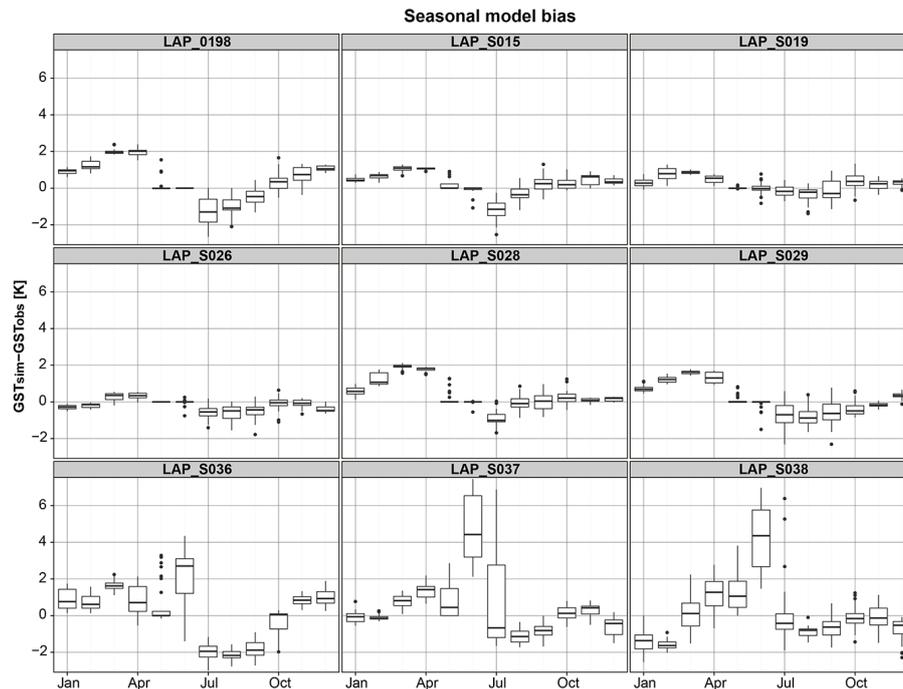
To further investigate the processes that might be responsible for the model bias observed so far in  $GST_{sim}$ ,  $GT_{sim}$  values were also compared with those of  $GT_{obs}$  for all thermistors from borehole LAP\_0198 (Table 1). In addition to



**Figure 5.** (a) Comparison of  $GST_{sim}$  (red) and  $GST_{obs}$  (black) for logger LAP\_S015; the blue area and blue line denote simulated and measured (station ATT-2, source: SLF) snow depth, respectively. (b)  $rMAGST_{sim}$  (red) and  $rMAGST_{obs}$  (black) anomalies for LAP\_S015; the shaded area represents the running 365-day model bias. (c) Running 365-day model bias for all loggers (LAP\_S015 in red). (d) Pearson correlation coefficients ( $R^2$ ) in a moving 365-day time window, based on  $R^2$  between  $GST_{sim}$  and  $GST_{obs}$  (LAP\_S015 in red). The colour bars at the bottom indicate the meteorological events described in Fig. 3 and Table 3.

the upper boundary conditions, the model for LAP\_0198 was also calibrated to fit the temperatures at depth (lower boundary condition) and the parameters listed in Table 2. Figure 8 shows that, similar to the GST simulations (Fig. 7), the winter cooling tends to be underestimated ( $GT_{sim}$  values are too warm during winter) and in general the seasonal amplitudes of  $GT_{sim}$  are smaller compared to  $GT_{obs}$ . This is interesting because the simulated snow depths seem rather shallow (Table 4), which should lead to a pronounced cooling of the ground during winter, but of course the cooling by convection between coarse blocks is not included in the model.

As Fig. 8 shows, the general pattern of different meteorological events (e.g. E3, E4 or E9) is well captured by the model, even though the freezing processes are not as fast or effective as in reality, which may be due to the missing convective and radiative heat transport through the coarse blocky layer in the model (Gruber and Hoelzle, 2008; Scherler et al., 2014). The lowering of the active layer thickness observed in the GT records and geophysical monitoring data (cf. Fig. 2) after E3 (2002/03) and E8 (2009) is also reproduced by the COUP model.



**Figure 6.** The seasonal evolution of the model bias ( $GST_{sim} - GST_{obs}$ ), aggregated for each month for the uppermost thermistor of borehole LAP\_0198 and all eight GST loggers. Dots at the end of the boxplots represent outliers (whiskers with maximum 1.5 IQR).

## 4 Discussion

### 4.1 Inter-annual variations and reaction to recent meteorological events

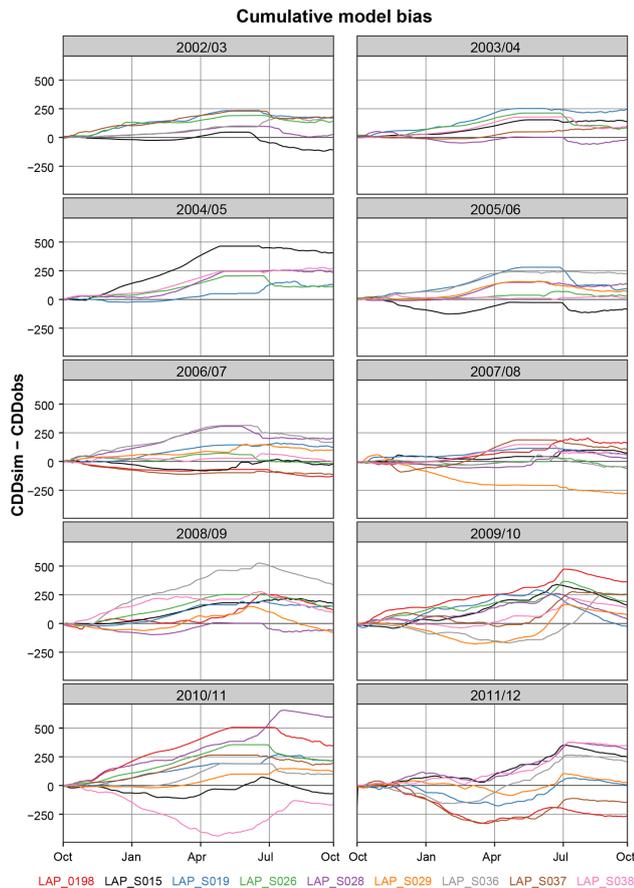
In general, the simulated variations in rMAGST closely follow the observed variations. Periods with the highest and lowest Pearson correlation between  $GST_{sim}$  and  $GST_{obs}$  were very different regarding the dominant meteorological and snow cover conditions (cf. Figs. 2 and 3), but usually coincide with large deviations of the melt-out day and the zero-curtain duration. No clear over- or underestimation of specific meteorological and/or snow cover conditions was found, and the relationship between the 365-day model bias and observed rMAGSTa is very weak (see Fig. 9a). The relationship gets even weaker for running 365-day Pearson correlation (Fig. 9b), indicating that the similarity between  $GST_{sim}$  and  $GST_{obs}$  is not influenced by inter-annual GST variations.

This implies that inter-annual GST variations and “meteorological events” are well enough simulated by the COUP model to cause no additional problems for long-term GST simulations, e.g. by additionally increasing the cumulative model bias. Even though the simulated snow depths agree astonishingly well with the recordings from ATT-2 and are also plausible regarding the variations of  $GST_{obs}$  (except for LAP\_S036–38), temporarily larger model bias and weaker  $R^2$  are likely related to snow cover dynamics and in particular with the timing of the snow melt. Cold  $GST_{obs}$  values with rather large short-time variations during winter (as the case

for LAP\_S036–38) were optimised by GLUE (Sect. 2.4) to simulate high densities of the new snow (Table 2) and in consequence shallow snow depths (Table 4). The higher thermal conductivity of the simulated snowpack explains untimely melt-out. Ekici et al. (2014) found that underestimated snow depths always lead to a cold GST model bias, whereas overestimated snow depths may cause positive or negative model bias. While this is clearly the case for some loggers (e.g. LAP\_S038 in the hydrological year 2010/11, cf. Fig. 7), in many other situations positive model biases have been found even though the simulated snow depths seemed underestimated. But since the connection between the model bias and the snow cover depends on many factors, such as the quality of the meteorological input data, effects like the redistribution of snow by wind and avalanches as well as the physics and parameterisation of the model remain important fields of current and future research.

### 4.2 Seasonal patterns and bias accumulation

A distinct seasonal pattern in the GST model bias was observed (Fig. 6). The systematic positive model bias in late winter is likely caused by (1) the ice core at depth missing in the model that acts as a heat sink (Scherler et al., 2014) and/or (2) underestimated (or missing) convective and advective processes that are known to occur in this kind of coarse-blocky substrate (Gruber and Hoelzle, 2008; Harris and Pedersen, 1998). The first effect is most prominent in winters



**Figure 7.** Cumulative GST model bias as degree days (DD) for the hydrological years (2003–2012): only loggers with complete observations are shown (cf. colour legend at the bottom).

with early ground insulation due to the snow (e.g. 2002/03, 2003/04 or 2008/09), when the cooling of the ground is limited.

The best model performance was observed during the snow-free season as well as in winters with very little snow, particularly in the colder period between autumn 2004 and late winter 2006 (event E4, cf. Figs. 3 and 5). By means of the CMB (Fig. 7) it was shown that the cumulative effect of the seasonal biases is positive in most years, although too cold  $GST_{sim}$  values in June and July (delayed melt-out, potentially underestimated radiation influence) did partially compensate for the positive bias during winter. The resulting model bias remained more or less constant during the whole observation period (Fig. 5) without clear trends, indicating that the bias of the final simulation year is not larger/worse than in the first year of the simulation. This may be different when the model is applied before or after the observation period and one should be aware of eventually cumulating errors in the near-surface layers. Having an observation period with very diverse conditions could be an advantage, first to analyse the reaction of the model during the calibration period

and second to tune the model to a large variety of possible situations. However the sensitivity of the snow and albedo parameters (Table 2) to extreme years during the calibration period requires further investigation and will give insight into the reliability of the models.

The simulation of GT using borehole LAP\_0198 was more difficult than for GST mainly because of the coarse dynamic resolution of the old thermistor chain (data until October 2009). Without a careful correction of the observations, the model would have been forced to several phase changes within the permafrost during the calibration period due to sometimes slightly positive (but erroneous) GT. The availability of geophysical measurements (which clearly showed a massive ice core which likely does not undergo complete melt processes) was very useful for plausibility checks and correcting the GT. The performance of the final simulation is comparable to those from other studies that have applied COUP to other boreholes from the PERMOS network (Marmy et al., 2013; Scherler et al., 2013, 2014). This is a surprisingly good result because multi-dimensional effects such as an up- and downslope air circulation cannot directly be simulated with a one-dimensional model like COUP, which also explains part of the model bias. Scherler et al. (2013, 2014) discussed the particularities of the energy balance within the coarse blocky surface layer of a rock glacier, using both direct observations and simulation results from the COUP model. They found a significant improvement of subsurface temperature simulations by adding a seasonally varying heat source/sink term, which parameterises the cooling effect of the coarse blocky material. Without this artificial heat sink, no permafrost would have been simulated (Scherler et al., 2014). At Lapires, this effect might be less extreme and superimposed by intra-talus ventilation, but it is likely still present. The internal air circulation in talus slopes is complex, and as it depends mainly on the temperature gradient between inside and outside the system (Delaloye and Lambiel, 2005), it likely has the strongest influence in situations with rather extreme cold or warm air temperatures. A parameterisation of these coupled processes was clearly beyond the scope of this paper but remains an open field for future research.

#### 4.3 Possibilities and limitations of COUP for GST and GT reconstruction

We assume that (a) avalanche activity, (b) a spatially varying ice content in the underlying permafrost, (c) site-specific intra-talus air circulation and (d) the disturbance of the snow cover by skiers are particularly difficult to capture with a model only calibrated by GST. Points (a) and (d) can increase or decrease the thermal conductivity of the snow cover by either purging parts of the snow pack apart from or onto the point of interest, modifying the density of the snow pack, or disturbing it. We assume that (d) is less important but leads to a slightly higher density of the snow pack, which might

be less thermally insulating towards the end of the ski season and last a bit longer until its complete melt-out. The reconstruction of the avalanche activity (a) is difficult and beyond the scope of this article, but it could eventually explain the remarkable GST bias in some years (e.g. “avalanche winter” 1999) and for some loggers located in potential deposition zones of avalanches (mainly LAP\_S026, LAP\_S028, LAP\_S029, LAP\_S036, LAP\_S037 and LAP\_S038). The systematic overestimation of the late-winter GST in years with early and intensive thermal insulation by the snow cover indicates that (b) also plays a major role, for example, in the hydrological years 2003 and 2009, when cold temperatures at greater depth (Fig. 5) were responsible for the cooling of the upper active layer before snow melt. Intra-talus ventilation (c) affects at least some of the GST loggers and the observations show that this effect mainly causes secondary temperature variations (Delaloye and Lambiel, 2005). We assume that this process is not dominant for the model bias and less important than other effects of the coarse-blocky material missing in the COUP model. As discussed in Scherler et al. (2013, 2014), convection and thermal radiation may be responsible for the systematic overestimation of the late-winter GST, even though the snow parameters (see Table 2) can partially account for this.

Since COUP is calibrated to a minimal RMSE between  $GST_{sim}$  and  $GST_{obs}$ , the depth of the sensor below the surface could be estimated in the model setup if no precise values are known. On the other hand, homogeneous monitoring conditions are of course very important for the observations. Long time series of high-quality data for the input and validation variables are essential for a stable calibration and good model performance.

In this study, we used the entire observation period of meteorological and GST observations. It would also be useful to validate the quality of the calibration with complementary data sets (e.g. water content, electrical resistivity as a proxy for the ice content or the local snow depth) if the time series are sufficiently long. In addition, some years or periods of different meteorological conditions could be excluded from the calibration to analyse this effect on the model parameters and performance. Of course, the inverse modelling approach (GLUE) also faces the issue of computing power. In order to test some parameters and to find the optimal fit, a certain number of successive simulations should be done, which is time- and computing-power-consuming.

#### 4.4 Benefits of joint observational and modelling studies

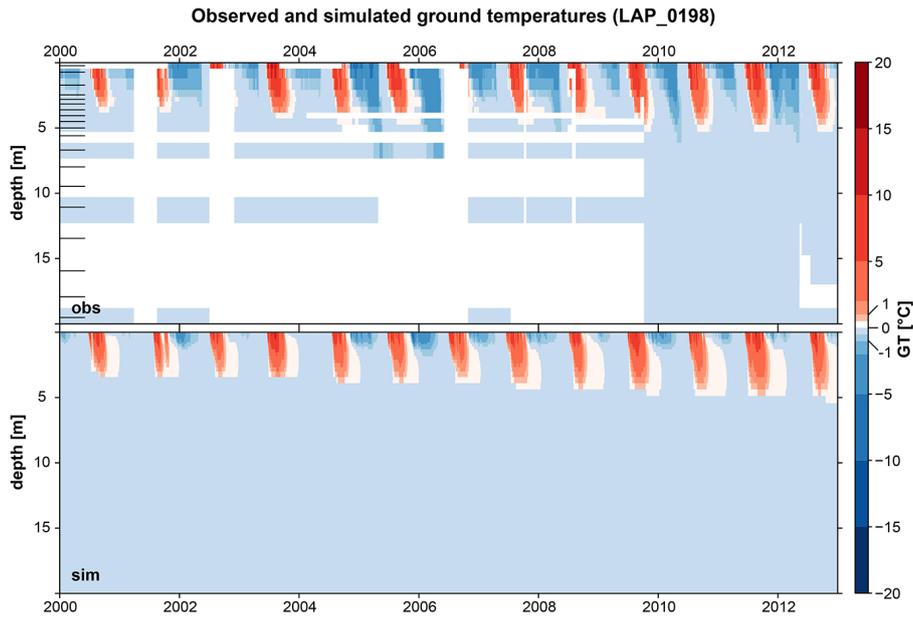
To benefit from joint research, it is important to analyse the strengths and weaknesses of the different methods. The main advantages of direct observations are (1) the comprehensive knowledge about and experience with site-specific conditions (surface characteristics, snow conditions etc.) from field work and various measurements obtained and (2) the observations themselves, which reflect the physical processes in all

their complexity of the real world. On the other hand, model studies are beneficial (1) to improve the understanding of the physical processes, (2) to simulate changes over long time periods and (3) to assess sensitivities while only changing one parameter at a time. The superposition of various influencing factors makes the quantification of physical processes based only on monitoring data very difficult. Model-based approaches, however, are faced with the need for simplification, the negligence of terrain heterogeneities, lateral and 3-D effects (at least for the COUP model used for this study) and a tendency towards over-parameterisation, which may lead to correct results but for the wrong reasons, as the observed seasonality of the model bias has shown. Furthermore, the short observation period ( $\sim 10$ – $15$  years) and remote study areas complicate both direct time series analysis (not sufficient to analyse trends on a decadal timescale) and scenario-based subsurface modelling (time series are often too short or incomplete regarding meteorological parameters for statistical downscaling of RCM output data and independent model validation).

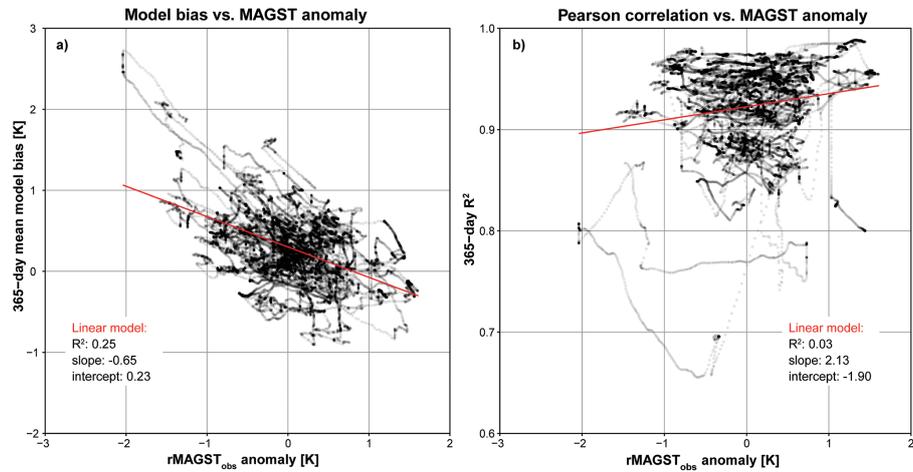
## 5 Conclusions and future perspectives

The COUP model has shown to be suitable for the reconstruction of GST time series situated on a talus slope. Not surprisingly, the best performance was found for periods without or with limited snow cover. Although the simulation of GST using only a prescribed upper boundary condition revealed surprisingly good results for the surface (which is not evident in locations with a coarse ground surface underlain by permafrost and prone to intra-talus air circulation; Delaloye and Lambiel, 2005), winters with limited ground cooling (usually after snow events in autumn) were systematically too warm in most of the simulations.

Simulated GST can serve (1) to fill gaps in the observations or to reconstruct GST of the past, (2) to investigate possible reasons for variations in the model bias and also (3) to assess the quality and uncertainty of model results. The key element to minimise the GST model bias is the reconstruction of an accurate snow cover, which is influenced by many factors and processes and particularly depends on precise meteorological input data (foremost precipitation) and a good parameterisation of the model (cf. Table 2). The meteorological input data used for this study (Rajczak et al., 2015) led to plausible results for the majority of the GST loggers and years regarding the snow cover evolution. This is remarkable because no local precipitation measurements were available. The simulation of the snow cover could possibly be further improved by separating the snowpack into several layers (for a more realistic structure with various temperature gradients), or by modelling the snowpack with software such as SNOWPACK (Lehning et al., 1999) or Alpine3D (Lehning et al., 2006).



**Figure 8.** Contour plot showing  $GT_{obs}$  (top) and  $GT_{sim}$  (bottom) over 20 m depth between 2000 and 2012 for borehole LAP\_0198 at the Lapires talus slope. The marks inside of the vertical axis on the observation panel indicate the depth of the thermistors. In white areas, observations are either missing or had to be removed because of bad quality (see Sect. 2.3).



**Figure 9.** Comparison of rMAGST anomaly for all GST loggers and the entire observation period ( $n = 31\,266$ ) with (a) the model bias ( $GT_{sim} - GT_{obs}$ ) and (b) running 365-day  $R^2$ .

The identification of specific meteorological and snow cover conditions revealed that events of very short duration (days–weeks) can largely affect variations in rMAGST. This is an issue to keep in mind when driving thermal subsurface models with downscaled RCM input data, which cannot reflect unique meteorological events but represent the mean state and seasonality of the climate. Furthermore, the influence of the calibration period should be analysed in detail and with consideration of the effects of meteorological events on the calibration parameters and model performance. At the moment it is not clear whether comparatively extreme years

should be included (to capture this natural variability in the model) or excluded from the calibration (to reduce the model bias during other periods). Further research is needed to fill this research gap. We recommend analysing inter-annual and seasonal patterns of the model bias during the calibration and validation period for every model study to understand how the model might react in different situations. This could be further combined with sensitivity assessments for the effects of single meteorological events on the snow and albedo calibration parameters by separating the validation from the calibration time window.

An important requirement for satisfying model results is having high-quality input data for long observation periods as well as precise information about the structure and composition of the subsurface which is representative of the point scale. These pre-conditions were rather good at our investigated study site, Lapires, in comparison with many other sites of the Swiss permafrost observing network PERMOS. The GST simulation at locations with permafrost evidence could be improved by prescribing a constant temperature of 0 °C at the maximal depth of the active layer (between 5.5 and 6 m for borehole LAP\_0198, cf. PERMOS 2013) and/or to improve the structure and material composition of the subsurface in the model (especially porosity, water and ice content) using, for example, the four-phase model approach (Hauck et al., 2011). Finally, more sophisticated downscaling methods of the meteorological input parameters could be applied, for instance, by correcting the influence of topography on the shortwave incoming radiation separately for each point instead of using the same meteorological input data for all GST loggers and the borehole. A detailed analysis and validation of the upper boundary conditions during the observation period clearly improves the reliability of any model for long-term simulations.

**Acknowledgements.** We greatly acknowledge the Swiss National Science Foundation for the funding of the Sinergia TEMPS project (project no. CRSII2\_136279), within which this study was conducted. Special thanks are due to our colleagues within TEMPS and PERMOS for providing the observational and reconstructed meteorological data, namely C. Lambiel from the University of Lausanne and J. Rajczak from ETH Zurich. We gratefully acknowledge the detailed and constructive comments of the two anonymous reviewers.

Edited by: M. Hoelzle

Reviewed by: two anonymous referees

## References

- Apaloo, J., Brenning, A., and Bodin, X.: Interactions between Seasonal Snow Cover, Ground Surface Temperature and Topography (Andes of Santiago, Chile, 33.5° S), *Permafrost Periglac.*, 23, 277–291, doi:10.1002/ppp.1753, 2012.
- Arenson, L., Hoelzle, M., and Springman, S.: Borehole deformation measurements and internal structure of some rock glaciers in Switzerland, *Permafrost Periglac.*, 13, 117–135, doi:10.1002/ppp.414, 2002.
- Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Process.*, 6, 279–298, doi:10.1002/hyp.3360060305, 1992.
- Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, 249, 11–29, doi:10.1016/S0022-1694(01)00421-8, 2001.
- 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, 2012.
- Delaloye, R.: Contribution à l'étude du pergélisol de montagne en zone marginale, PhD Thesis, University of Fribourg, Switzerland, 244 pp., *GeoFocus*, 2004.
- Delaloye, R. and Lambiel, C.: Evidence of winter ascending air circulation throughout talus slopes and rock glaciers situated in the lower belt of alpine discontinuous permafrost (Swiss Alps), *Norsk Geografisk Tidsskrift – Norwegian Journal of Geography*, 59, 194–203, doi:10.1080/00291950510020673, 2005.
- Delaloye, R., Reynard, E., and Lambiel, C.: Pergélisol et construction de remontées mécaniques: l'exemple des Lapires (Mont-Gelé, Valais), *Le gel en géotechnique*, Publications de la Société Suisse de Mécanique des Sols et des Roches, 141, 103–113, 2001.
- Delaloye, R., Reynard, E., Lambiel, C., Marescot, L., and Monnet, R.: Thermal anomaly in a cold scree slope, Creux du Van, Switzerland, in: *Proceedings of the 8th International Conference on Permafrost*, 21–25 July 2003, Zurich, Switzerland, 175–180, 2003.
- Ekici, A., Chadburn, S., Chaudhary, N., Hajdu, L. H., Marmy, A., Peng, S., Boike, J., Burke, E., Friend, A. D., Hauck, C., Krinner, G., Langer, M., Miller, P. A., and Beer, C.: Site-level model intercomparison of high latitude and high altitude soil thermal dynamics in tundra and barren landscapes, *The Cryosphere Discuss.*, 8, 4959–5013, doi:10.5194/tcd-8-4959-2014, 2014.
- Endrizzi, S. and Gruber, S.: Investigating the Effects of Lateral Water Flow on the Spatial Patterns of Thaw Depth, in: *Tenth International Conference on Permafrost*, Salekhard, Yamal-Nenets Autonomous District, Russia, 25 June 2012–29 June 2012, 91–96, 2012.
- Engelhardt, M., Hauck, C., and Salzmann, N.: Influence of atmospheric forcing parameters on modelled mountain permafrost evolution, *Meteorol. Z.*, 19, 491–500, doi:10.1127/0941-2948/2010/0476, 2010.
- Etzelmüller, B., Schuler, T. V., Isaksen, K., Christiansen, H. H., Farbro, H., and Benestad, R.: Modeling the temperature evolution of Svalbard permafrost during the 20th and 21st century, *The Cryosphere*, 5, 67–79, doi:10.5194/tc-5-67-2011, 2011.
- 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.
- Fiddes, J. and Gruber, S.: TopoSCALE v.1.0: downscaling gridded climate data in complex terrain, *Geosci. Model Dev.*, 7, 387–405, doi:10.5194/gmd-7-387-2014, 2014.
- Ğadek, B.: Debris slopes ventilation in the periglacial zone of the Tatra Mountains (Poland and Slovakia): The indicators, *Cold Reg. Sci. Technol.*, 74–75, 1–10, doi:10.1016/j.coldregions.2012.01.007, 2012.
- Goodrich, L. E.: The influence of snow cover on the ground thermal regime, *Can. Geotech. J.*, 19, 421–432, doi:10.1139/t82-047, 1982.
- Gruber, S. and Haeberli, W.: Permafrost in steep bedrock slopes and its temperature-related destabilization following climate change, *J. Geophys. Res.*, 112, F02S18, doi:10.1029/2006JF000547, 2007.

- Gruber, S. and Haeberli, W.: Mountain Permafrost, in *Permafrost Soils*, vol. 16, edited by: Margesin, R., 33–44, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- Gruber, S. and Hoelzle, M.: The cooling effect of coarse blocks revisited: a modeling study of a purely conductive mechanism, in: *Proceedings of the 9th International Conference on Permafrost*, Fairbanks, Alaska, 557–561, 2008.
- 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.
- Hanson, S. and Hoelzle, M.: Installation of a shallow borehole network and monitoring of the ground thermal regime of a high alpine discontinuous permafrost environment, *Eastern Swiss Alps*, *Norsk Geografisk Tidsskrift – Norwegian J. Geogr.*, 59, 84–93, doi:10.1080/00291950510020664, 2005.
- Harris, C., Arenson, L. U., Christiansen, H. H., Etzelmüller, B., Frauenfelder, R., Gruber, S., Haeberli, W., Hauck, C., Hoelzle, M., Humlum, O., Isaksen, K., Kääb, A., Kern-Lütschg, M. A., Lehning, M., Matsuoka, N., Murton, J. B., Noetzli, J., Phillips, M., Ross, N., Seppälä, M., Springman, S. M., and Vonder Mühl, 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.
- Harris, S. A. and Pedersen, D. E.: Thermal regimes beneath coarse blocky materials, *Permafrost Periglac.*, 9, 107–120, doi:10.1002/(SICI)1099-1530(199804/06)9:2<107::AID-PPP277>3.0.CO;2-G, 1998.
- Hauck, C., Böttcher, M., and Maurer, H.: A new model for estimating subsurface ice content based on combined electrical and seismic data sets, *The Cryosphere*, 5, 453–468, doi:10.5194/tc-5-453-2011, 2011.
- Hawkins, E. and Sutton, R.: The Potential to Narrow Uncertainty in Regional Climate Predictions, *B. Am. Meteorol. Soc.*, 90, 1095–1107, doi:10.1175/2009BAMS2607.1, 2009.
- Hilbich, C.: Time-lapse refraction seismic tomography for the detection of ground ice degradation, *The Cryosphere*, 4, 243–259, doi:10.5194/tc-4-243-2010, 2010.
- Hilbich, C., Fuss, C., and Hauck, C.: Automated Time-lapse ERT for Improved Process Analysis and Monitoring of Frozen Ground, *Permafrost Periglac.*, 22, 306–319, doi:10.1002/ppp.732, 2011.
- Hoelzle, M., Wegmann, M., and Krummenacher, B.: Miniature temperature dataloggers for mapping and monitoring of permafrost in high mountain areas: first experience from the Swiss Alps, *Permafrost Periglac.*, 10, 113–124, doi:10.1002/(SICI)1099-1530(199904/06)10:2<113::AID-PPP317>3.0.CO;2-A, 1999.
- Hoelzle, M., Mittaz, C., Etzelmüller, B., and Haeberli, W.: Surface energy fluxes and distribution models of permafrost in European mountain areas: an overview of current developments, *Permafrost Periglac.*, 12, 53–68, doi:10.1002/ppp.385, 2001.
- IPCC: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, 2014.
- Isaksen, K., Ødegård, R. S., Etzelmüller, B., Hilbich, C., Hauck, C., Farbrot, H., Eiken, T., Hygen, H. O., and Hipp, T. F.: Degrading Mountain Permafrost in Southern Norway: Spatial and Temporal Variability of Mean Ground Temperatures, 1999–2009, *Permafrost Periglac.*, 22, 361–377, doi:10.1002/ppp.728, 2011.
- Ishikawa, M.: Thermal regimes at the snow–ground interface and their implications for permafrost investigation, *Geomorphology*, 52, 105–120, doi:10.1016/S0169-555X(02)00251-9, 2003.
- Jansson, P. E.: CoupModel: Model Use, Calibration, and Validation, *Transactions of the ASABE*, 55, 1337–1346, doi:10.13031/2013.42245, 2012.
- Jansson, P. E. and Karlberg, L.: Coupled heat and mass transfer model for soil-plant-atmosphere systems, 365 pp., 2004.
- Keller, F. and Gubler, H.: Interaction between snow cover and high mountain permafrost, Murtèl-Corvatsch, Swiss Alps, in: *Proceedings of the 6th International Conference on Permafrost*, Beijing, China, edited by G. Cheng, 332–337, South China University of Technology Press, Guangzhou, 1993.
- Kneisel, C., Hauck, C., Fortier, R., and Moorman, B.: Advances in geophysical methods for permafrost investigations, *Permafrost Periglac.*, 19, 157–178, doi:10.1002/ppp.616, 2008.
- Lambiel, C.: Le pergélisol dans les terrains sédimentaires à forte déclivité: distribution, régime thermique et instabilités, 280 pp., *Travaux et recherches Vol. 33*, 2006.
- Lehning, M., Bartelt, P., Brown, B., Russi, T., Stöckli, U., and Zimmerli, M.: snowpack model calculations for avalanche warning based upon a new network of weather and snow stations, *Cold Reg. Sci. Technol.*, 30, 145–157, doi:10.1016/S0165-232X(99)00022-1, 1999.
- Lehning, M., Völksch, I., Gustafsson, D., Nguyen, T. A., Stähli, M., and Zappa, M.: ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology, *Hydrol. Process.*, 20, 2111–2128, doi:10.1002/hyp.6204, 2006.
- Luetschg, M., Lehning, M., and Haeberli, W.: A sensitivity study of factors influencing warm/thin permafrost in the Swiss Alps, *J. Glaciol.*, 54, 696–704, doi:10.3189/002214308786570881, 2008.
- Marmy, A., Salzmann, N., Scherler, M., and Hauck, C.: Permafrost model sensitivity to seasonal climatic changes and extreme events in mountainous regions, *Environ. Res. Lett.*, 8, 035048, doi:10.1088/1748-9326/8/3/035048, 2013.
- Morard, S., Delaloye, R., and Dorthe, J.: Seasonal thermal regime of a mid-latitude ventilated debris accumulation, in: *Proceedings of the 9th International Conference on Permafrost*, Fairbanks, Alaska, 1233–1238., 2008.
- Noetzli, J. and Gruber, S.: *Alpiner Permafrost - ein Überblick*, edited by: Lintzmeyer, K., *Jahrbuch des Vereins zum Schutz der Bergwelt*, Selbstverlag, München, 70, 111–121, 2005.
- Noetzli, J. and Gruber, S.: Transient thermal effects in Alpine permafrost, *The Cryosphere*, 3, 85–99, doi:10.5194/tc-3-85-2009, 2009.
- Olaya, V.: *Basic Land-Surface Parameters*, in *Geomorphometry – Concepts, Software, Applications*, vol. 33, 141–169, Elsevier, 2009.
- PERMOS: *Permafrost in Switzerland 2008/2009 and 2009/2010*. Noetzli, J. (ed.), *Glaciological Report (Permafrost) No. 10/11 of the Cryospheric Commission of the Swiss Academy of Sciences (SCNAT)*, Zurich, Switzerland, 2013.
- Phillips, M., Mutter, E. Z., Kern-Luetschg, M., and Lehning, M.: Rapid degradation of ground ice in a ventilated talus slope: Flüela Pass, Swiss Alps, *Permafrost Periglac.*, 20, 1–14, doi:10.1002/ppp.638, 2009.

- Rajczak, J., Kotlarski, S., Salzmann, N., and Schär, C.: Robust climate scenarios for sites with sparse observations: A two-step bias correction approach, *Int. J. Climatol.*, submitted, 2015.
- Scapoza, C.: Stratigraphie, morphodynamique, paléoenvironnements des terrains sédimentaires meubles à forte déclivité du domaine périglaciaire alpin (Géovisions no. 40), PhD Thesis, University of Lausanne, 580 pp., 2013.
- Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Haeberli, C., Liniger, M. A., and Appenzeller, C.: The role of increasing temperature variability in European summer heatwaves, *Nature*, 427, 332–6, doi:10.1038/nature02300, 2004.
- Scherler, M., Hauck, C., Hoelzle, M., Stähli, M., and Völksch, I.: Meltwater infiltration into the frozen active layer at an alpine permafrost site, *Permafrost Periglac.*, 21, 325–334, doi:10.1002/ppp.694, 2010.
- Scherler, M., Hauck, C., Hoelzle, M., and Salzmann, N.: Modeled sensitivity of two alpine permafrost sites to RCM-based climate scenarios, *J. Geophys. Res.: Earth Surface*, 118, 780–794, doi:10.1002/jgrf.20069, 2013.
- Scherler, M., Schneider, S., Hoelzle, M., and Hauck, C.: A two-sided approach to estimate heat transfer processes within the active layer of the Murtèl-Corvatsch rock glacier, *Earth Surf. Dynam.*, 2, 141–154, doi:10.5194/esurf-2-141-2014, 2014.
- 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.
- Schneider, S.: The heterogeneity of mountain permafrost – A field-based analysis of different periglacial materials, PhD Thesis, University of Fribourg, 176 pp., 2014.
- Schneider, S., Hoelzle, M., and Hauck, C.: Influence of surface and subsurface heterogeneity on observed borehole temperatures at a mountain permafrost site in the Upper Engadine, Swiss Alps, *The Cryosphere*, 6, 517–531, doi:10.5194/tc-6-517-2012, 2012.
- Schneider, S., Daengeli, S., Hauck, C., and Hoelzle, M.: A spatial and temporal analysis of different periglacial materials by using geoelectrical, seismic and borehole temperature data at Murtèl-Corvatsch, Upper Engadin, Swiss Alps, *Geogr. Helv.*, 68, 265–280, doi:10.5194/gh-68-265-2013, 2013.
- Stocker-Mittaz, C., Hoelzle, M., and Haeberli, W.: Modelling alpine permafrost distribution based on energy-balance data: a first step, *Permafrost Periglac.*, 13, 271–282, doi:10.1002/ppp.426, 2002.
- Vonder Mühl, D., Stucki, T., and Haeberli, W.: Borehole temperatures in alpine permafrost: a ten year series, in: Proceedings of the 7th International Conference on Permafrost, Yellowknife (Canada), 1089–1095, Collection Nordicana, 57, 1998.
- Wakonigg, H.: Unterkühlte Schutthalden, Beiträge zur Permafrostkartierung in Österreich, Arbeiten aus dem Inst. f. Geogr. Karl-Franzens-Universität Graz, 209–223, 1996.
- Williams, P. J. and Smith, M. W.: *The Frozen Earth: Fundamentals of Geocryology*, Cambridge University Press, Cambridge, 1989.
- Zenkhusen Mutter, E. and Phillips, M.: Active Layer Characteristics At Ten Borehole Sites In Alpine Permafrost Terrain, Switzerland, *Permafrost Periglac.*, 23, 138–151, doi:10.1002/ppp.1738, 2012.
- Zhang, T.: Influence of the seasonal snow cover on the ground thermal regime: An overview, *Rev. Geophys.*, 43, 23, doi:10.1029/2004RG000157, 2005.