The Cryosphere, 10, 1859–1881, 2016
www.the-cryosphere.net/10/1859/2016/
doi:10.5194/tc-10-1859-2016

The Cryosphere Open Access EGU

Multilevel spatiotemporal validation of snow/ice mass balance and runoff modeling in glacierized catchments

Florian Hanzer[1,2], **Kay Helfricht**[3], **Thomas Marke**[2], **and Ulrich Strasser**[2]

[1]alpS – Centre for Climate Change Adaptation, Innsbruck, Austria
[2]Institute of Geography, University of Innsbruck, Innsbruck, Austria
[3]Institute for Interdisciplinary Mountain Research, Austrian Academy of Sciences, Innsbruck, Austria

Correspondence to: Florian Hanzer (florian.hanzer@uibk.ac.at)

Received: 2 March 2016 – Published in The Cryosphere Discuss.: 5 April 2016
Revised: 13 July 2016 – Accepted: 7 August 2016 – Published: 25 August 2016

Abstract. In this study, the fully distributed, physically based hydroclimatological model AMUNDSEN is set up for catchments in the highly glacierized Ötztal Alps (Austria, 558 km^2 in total). The model is applied for the period 1997–2013, using a spatial resolution of 50 m and a temporal resolution of 1 h. A novel parameterization for lateral snow redistribution based on topographic openness is presented to account for the highly heterogeneous snow accumulation patterns in the complex topography of the study region. Multilevel spatiotemporal validation is introduced as a systematic, independent, complete, and redundant validation procedure based on the observation scale of temporal and spatial support, spacing, and extent. This new approach is demonstrated using a comprehensive set of eight independent validation sources: (i) mean areal precipitation over the period 1997–2006 derived by conserving mass in the closure of the water balance, (ii) time series of snow depth recordings at the plot scale, (iii–iv) multitemporal snow extent maps derived from Landsat and MODIS satellite data products, (v) the snow accumulation distribution for the winter season 2010/2011 derived from airborne laser scanning data, (vi) specific surface mass balances for three glaciers in the study area, (vii) spatially distributed glacier surface elevation changes for the entire area over the period 1997–2006, and (viii) runoff recordings for several subcatchments. The results indicate a high overall model skill and especially demonstrate the benefit of the new validation approach. The method can serve as guideline for systematically validating the coupled components in integrated snow-hydrological and glacio-hydrological models.

1 Introduction

Assessing the amount of water resources stored in mountain catchments as snow and ice as well as the timing of meltwater production and the resulting runoff is of high interest for glaciological and hydrological investigations and hydropower production. Climate-change-induced shifts in snow and ice melt will alter the hydrological regimes in glacierized catchments in terms of both timing and magnitude of streamflow discharge. Longer periods of negative glacier mass balances first result in increased runoff due to the enlarged contribution of glacier melt later on, followed by a decline of runoff amounts as a consequence of the reduced glacier area (e.g., Jansson et al., 2003; Beniston, 2003; Collins, 2008; Bliss et al., 2014). For some regions, this moment of "peak water" has already passed, while for others it is expected over the course of the current century (Salzmann et al., 2014; Bliss et al., 2014). Consequently, these catchments undergo a regime shift from ice-melt-dominated towards snowmelt-dominated runoff cycles, accompanied by a shift of the monthly maximum runoff amounts towards earlier periods of the year due to accelerated snowmelt (e.g., Horton et al., 2006).

When using hydroclimatological simulation models to investigate past, current, or future water resources, it can generally be distinguished between conceptual and physically based models. Conceptual models often use simple transfer functions to translate input data (e.g., meteorological data) to the desired output variables (e.g., runoff). They tend to be comparatively parsimonious both in terms of input data and the number of parameters; however, as the parameters often

have no direct physical meaning their values usually have to be determined by calibration, i.e., by tuning them until the model produces the desired output. As conceptual models are usually calibrated to optimize results for a certain period of time, their performance under changing conditions (e.g., climate change) is uncertain. With physically based models on the contrary, it is aimed to explicitly simulate individual processes by following the laws of physics such as conservation of energy, mass, and momentum. In a fully physically based model, all parameters have a physical interpretation and can, in principle, be derived from field measurements. Determination of these parameters, however, is highly scale-dependent, both with respect to the investigated processes and the model itself (e.g., Blöschl and Sivapalan, 1995). In practice, often few direct measurements are available and a number of required input parameters has to be inter- and extrapolated in space and time.

When validating hydrological models, commonly only runoff records are applied as direct field measurements. It is, however, well established that validating hydrological models by only comparing observed and simulated runoff at the catchment outlet is not sufficient, as multiple parameter sets may yield the same results (the equifinality problem Beven, 1993) – for example, in glacierized catchments underestimations in simulated precipitation volumes may be compensated by increased ice melt contributions. This might be acceptable for short-term applications such as operational flood forecasting, where the main aim is to acquire accurate discharge simulations, but when applying such model parameterizations for long-term simulations (e.g., to determine future runoff changes due to climate change) the errors due to the misrepresentation of specific processes can accumulate and the model provides misleading results. Hence, whenever possible, multiple independent data sets should be used for model calibration and validation (e.g., Refsgaard, 1997; Grayson et al., 2002; Schaefli et al., 2005; Finger et al., 2015).

For this purpose, for snow-hydrological applications frequently satellite-derived snow extent observations are additionally used in the validation process (e.g., Parajka and Blöschl, 2008; Finger et al., 2011). As they are available operationally in high temporal resolution (up to daily, but always with the constraint of frequent cloud obstructions), they allow us to spatially assess snow accumulation and melt processes in the model. Additionally, incorporating measured glacier mass balances has been shown to lead to more realistic process representations in models (but not necessarily in terms of improved runoff results) (e.g., Konz and Seibert, 2010; Finger et al., 2011; Schaefli et al., 2005; Schaefli and Huss, 2011). Glacier mass balance is an integral measure of the accumulation and ablation processes over the glacier in a defined period, but observations are either available only for very few glaciers with direct measurements or are acquired using digital elevation model (DEM) differencing covering multiyear periods.

Since several studies have shown that especially accurately capturing the amount and distribution of snow accumulation during the winter is a prerequisite for reliable long-term runoff simulations (e.g., Huss et al., 2014; Magnusson et al., 2011), much work has been put in parameterizing models to better capture winter snow accumulation in terms of spatial distribution and volume. For this purpose, commonly point measurements of snow depth, snow water equivalent (SWE), or precipitation are inter- and extrapolated, but information on the total water volume stored in entire mountain catchments is rare (e.g., Jonas et al., 2009). Measurements of solid precipitation are impaired with errors (e.g., Sevruk, 1986; Goodison et al., 1998), and the representativeness of point measurements for entire mountain catchments is uncertain (e.g., Grünewald and Lehning, 2011). A relatively recent technology is to use lidar-derived surface elevation differences to obtain snow depth maps in high spatial resolution (e.g., Deems et al., 2006; Grünewald et al., 2010; Grünewald and Lehning, 2011; Grünewald et al., 2013; Schöber et al., 2014; Helfricht et al., 2012, 2014a, b). Schöber et al. (2014) demonstrated that using lidar-derived SWE maps in the calibration of a spatially distributed snow-hydrological model significantly improved the results for simulated snow accumulation compared to the assimilation of optical remote sensing data of the snow-covered area only.

In our study, we set up the physically based hydroclimatological model AMUNDSEN (Strasser, 2008) for a study region in the highly glacierized Ötztal Alps (Austria). By introducing a concept for systematic model validation using the "observation scale" (Blöschl and Sivapalan, 1995) of temporal and spatial support, spacing, and extent, we validate the model results against (i) mean areal precipitation over the period 1997–2006, (ii) time series of point-based snow depth recordings at several locations, (iii)–(iv) multitemporal snow extent maps acquired from Landsat and MODIS imagery, (v) the snow accumulation distribution for the winter season 2010/2011 acquired using airborne laser scanning (ALS) surveys, (vi) glacier-averaged annual surface mass balances for three glaciers in the study area, (vii) spatially distributed glacier surface elevation changes for the entire area over the period 1997–2006, and (viii) hourly runoff records for several subcatchments.

2 Study site and model input data

The study site is located in the mountain region of the Ötztal Alps (Tyrol, Austria, Fig. 1) and comprises the headwater catchments of the valleys Ötztal, Pitztal, and Kaunertal, which contribute to the streamflow of the river Inn. The combined area of the investigated catchments is 558 km^2, of which 480 km^2 is gauged. Elevations range between 1760 m a.s.l. at the lowest elevated runoff gauge and 3770 m a.s.l. at the top of Wildspitze, the highest summit of Tyrol. The Ötztal Alps are the most glacierized moun-

Table 1. Area, minimum/maximum/mean elevation, and glacierization (as of the year 1997) of the investigated catchments. The catchment IDs correspond to the labels in Fig. 1.

ID	Catchment	Area (km^2)	z_{min} (m a.s.l.)	z_{max} (m a.s.l.)	z_{mean} (m a.s.l.)	Glacierization 1997 (%)
1	Rofenache	98.6	1891	3762	2890	38.3
2	Am Barst/Gurgler Ache	72.4	1883	3496	2780	31.8
3	Niedertalbach	66.7	1878	3585	2883	31.1
4	Gepatschalm	53.9	1893	3526	2821	39.7
5	Gepatschspeicher	52.3	1744	3445	2480	1.7
6	Taschachbach	44.7	1789	3754	2737	26.1
7	Pitze	27.0	1812	3519	2830	48.2
8	Radurschlbach	24.0	1794	3349	2601	1.5
9	Tscheybach	16.4	1800	3055	2417	0.0
10	Rifflsee Ausfluss	15.9	2232	3386	2777	17.8
11	Venter Ache	15.0	1836	3498	2680	13.5
12	Verpeilbach	12.1	1807	3370	2562	9.3
13	Fissladbach	11.4	1797	3137	2576	2.1
14	Poschach/Gurgler Ache	8.1	1826	3217	2435	1.6
15	Platzertal	7.9	2298	3104	2667	2.8
16	Watzebach	6.7	1819	3525	2718	10.2
17	Ferwallbach	6.3	1851	3254	2608	4.5
18	Königsbach	5.9	1876	3077	2607	0.0
19	Rostizbach	4.8	1842	3389	2628	0.7
20	Gsallbach	3.9	1843	3250	2560	9.1
21	Madatschbach	3.8	1847	3339	2605	13.6
22	Bloabachl	0.2	1775	2638	2323	0.0
23	Poschach	0.2	1828	2328	2004	0.0
	Total	558.1	1744	3762	2740	24.6

tain region of the Eastern Alps, accounting for almost one-third of the glacier area in Austria. According to the glacier outlines of the second Austrian Glacier Inventory (Fischer et al., 2015), in 1997 137 km^2 (24 %) of the investigated area was covered by glaciers. Only 11 km^2 (2 %) of the investigated area is covered by forests and another 3 km^2 (0.5 %) by shrubs (source: Land Tirol – http://data.tirol.gv.at). Main surface types besides the ice-covered areas are alpine grass vegetation, debris cover, and rock walls. Due to the touristic development and the production of hydropower in the region, a large number of weather stations and discharge measurements ensure an extensive basis of meteorological and hydrological data. Approx. half of the study area (277 km^2) is catchments which supply water to the Gepatsch hydropower reservoir (see Fig. 1). Table 1 lists the area, minimum/maximum/mean elevation, and glacierization (as of the year 1997) of the investigated catchments.

An ALS-derived DEM from the year 2006 (source: Land Tirol – http://data.tirol.gv.at) resampled to 50 m resolution was used as input for the model and the calculation of derived terrain variables (e.g., slope, aspect, sky-view factor, and catchment boundaries). Meteorological records (air temperature, precipitation, relative humidity, global radiation, and wind speed) in hourly resolution from 14 automatic weather stations in and surrounding the study region were used to drive the model. Initial ice thickness distribution of all glaciers in the Ötztal Alps was calculated using the approach by Huss and Farinotti (2012). The method is based on glacier mass turnover and ice flow mechanics and requires glacier outlines and a DEM. Mass balance gradients and constants recommended by M. Huss (personal communication, 2011) were used to calculate volumetric balance fluxes of the individual glaciers. Required glacier surface elevations and glacier outlines of the year 1997 exist from the second Austrian Glacier Inventory (Fischer et al., 2015).

3 Methods

3.1 The hydroclimatological model AMUNDSEN

The modular, physically based, distributed modeling system AMUNDSEN (Strasser, 2008) was applied for the simulation of the snow and ice surface mass balance. AMUNDSEN has been designed to specifically address the requirements of snow modeling in mountain regions under climate change conditions and has already been extensively validated in various Alpine sites (Strasser, 2004; Pellicciotti et al., 2005;

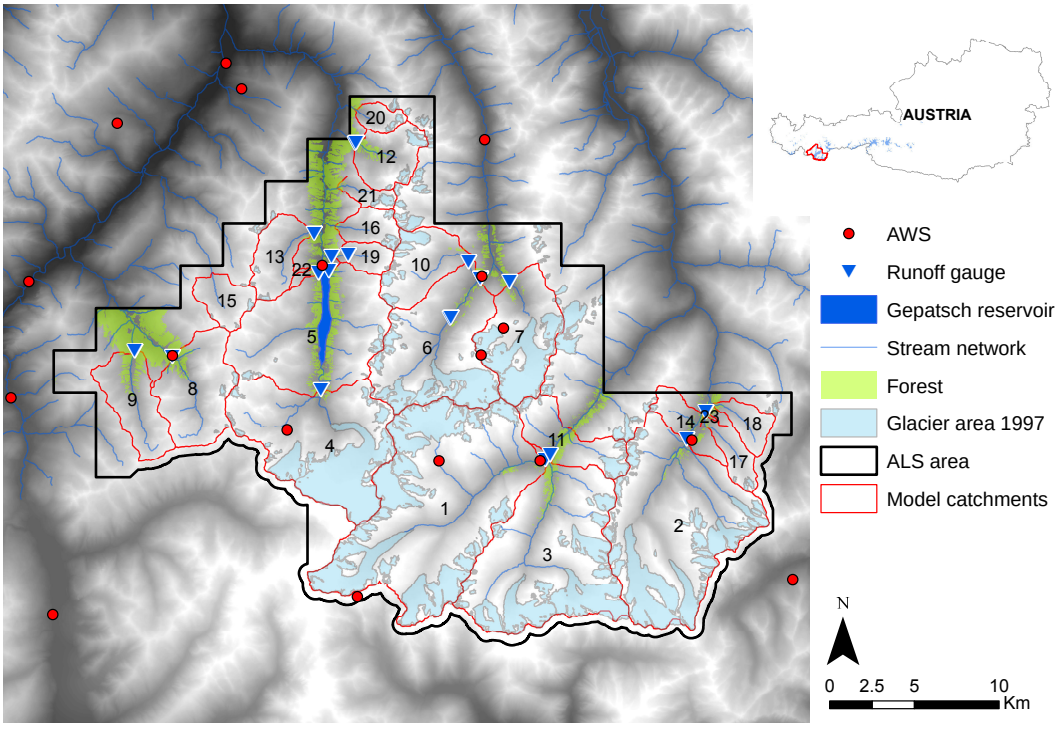

Figure 1. Location of the study site in the Ötztal Alps showing glacierized and forested areas, catchment boundaries, and the locations of the meteorological stations and runoff gauges. Numbers indicate catchment IDs (spatial statistics of the individual catchments are listed in Table 1).

Strasser et al., 2008; Strasser, 2008; Hanzer et al., 2014; Marke et al., 2015).

As input data for the model, a DEM of the model domain with a spatial resolution typically on the order of tens to hundreds of meters (the comparatively high resolution is necessary for adequately capturing the small-scale processes shaping the snow cover in complex terrain) as well as hourly to 3-hourly recordings of the meteorological variables air temperature, relative humidity, precipitation, global radiation, and wind speed are required. Several derived topographic parameters (slope, aspect, sky-view factor, openness) can either be preprocessed or calculated during runtime. In order to enable specific submodules (canopy module, evapotranspiration, runoff), various other spatial input fields (land cover, soil, catchment boundaries) have to be prescribed. The calculations presented in this study have been performed on a 50 m grid and with hourly meteorological recordings.

Interpolated fields from the scattered point measurements are – in the case of temperature, precipitation, humidity, and wind speed – obtained using a combined lapse rate–inverse distance weighting scheme, either using automatically calculated lapse rates for each time step or using prescribed monthly lapse rates. The latter approach has been chosen for this study – for temperature and humidity (dew point temperature), the lapse rates derived by Marke (2008) were used, while for precipitation monthly lapse rates were derived from time series recorded by accumulative precipitation gauges in the study area. In the case of radiation, first clear-sky global radiation is calculated following Corripio (2002), taking into account hill shading, transmission losses due to scattering (Rayleigh and aerosol scattering) and absorption (by water vapor, ozone, and other trace gases), transmission gains due to multiple reflections between the atmosphere and the ground, and reflections from surrounding terrain. Subsequently, actual global radiation is obtained by correcting the clear-sky radiation with interpolated cloud factor fields (obtained using radiation recordings at the meteorological stations). Incoming longwave radiation is also derived following Corripio (2002) using parameterizations for the radiation fractions coming from the clear sky, from clouds, and from surrounding slopes.

Precipitation phase is then determined using a wet-bulb temperature threshold of $T_w = 2\,°C$. Wet-bulb temperature is calculated by iteratively solving the psychrometric equation. Four types of snow/ice layers are distinguished in the model, namely new snow, old snow, firn, and ice, with each layer having distinct properties in terms of water equivalent, density, and albedo. Fresh snowfall is always added to the new snow layer. New snow is converted to old snow when reaching a transition density of $200\,kg\,m^{-3}$, old snow to firn always on 30 September, and firn to ice when reaching a transition density of $900\,kg\,m^{-3}$. Fresh snow density ρ_{ns} is cal-

culated as a function of air temperature T_a (°C) following Anderson (1976) and Jordan (1991):

$$\rho_{ns} = \begin{cases} 50 & T_a \leq -15\,^{\circ}\mathrm{C} \\ 50 + 1.7(T_a + 15)^{1.5} & T_a > -15\,^{\circ}\mathrm{C}. \end{cases} \quad (1)$$

Snow compaction for the new snow and old snow layer is calculated following Anderson (1976) and Jordan (1991), taking into account the effects of compaction and metamorphism:

$$\frac{\mathrm{d}\rho_s}{\mathrm{d}t} = \rho_s \left(c_1 W^* \mathrm{e}^{-c_2(T^* - T_s)} \mathrm{e}^{-c_3 \rho_s} \right), \quad (2)$$

$$\frac{\mathrm{d}\rho_s}{\mathrm{d}t} = \rho_s \left(c_4 \mathrm{e}^{-c_5(T^* - T_s)} c_6 \right), \quad (3)$$

$$c_6 = \begin{cases} \mathrm{e}^{-c_7(\rho_s - \rho_d)} & \rho_s > \rho_d \\ 1 & \rho_s \leq \rho_d, \end{cases} \quad (4)$$

with ρ_s (kg m^{-3}) being the layer (new snow or old snow) density, W^* (kg m^{-2}) the load of snow water equivalent (snow in the layer above and 50 % of the snow in the current layer), $c_1 = 0.01\,\mathrm{m}^{-1}\,\mathrm{h}$ (new snow), $c_1 = 0.001\,\mathrm{m}^{-1}\,\mathrm{h}$ (old snow), $c_2 = 0.08\,^{\circ}C^{-1}$, $c_3 = 0.021\,\mathrm{m}^3\,\mathrm{kg}^{-1}$, $c_4 = 0.01\,\mathrm{m}^{-1}\,\mathrm{h}$, $c_5 = 0.04\,^{\circ}C^{-1}$, $c_7 = 0.046\,\mathrm{m}^3\,\mathrm{kg}^{-1}$, $\rho_d = 150\,\mathrm{kg\,m}^{-3}$, and $T^* = 0\,^{\circ}C$. For the firn layer, a linear transition from firn to ice in 10 years is assumed, while ice density is kept constant at 900 kg m^{-3}.

Snow surface albedo α is parameterized following Rohrer (1992) taking into account snow age and temperature:

$$\alpha = \alpha_{min} + (\alpha_{t-1} - \alpha_{min})\,\mathrm{e}^{-c_T/24}, \quad (5)$$

where α_{min} is the (prescribed) minimum albedo, α_{t-1} the albedo in the previous time step, and c_T a temperature-dependent recession factor (implemented by prescribing two factors $c_{T \geq 0}$ and $c_{T < 0}$ for positive and negative air temperatures, respectively). For the present study, fresh snow albedo was set to 0.85, while α_{min}, $c_{T \geq 0}$, and $c_{T < 0}$ for new snow and old snow were set to 0.55, 0.12, and 0.05, respectively. Firn and ice albedo were held constant with $\alpha_{firn} = 0.4$ and $\alpha_{ice} = 0.2$.

In forested areas, a canopy submodule optionally modifies the meteorological variables for inside-canopy conditions (Strasser, 2008) and accounts for the forest snow processes of interception, sublimation, and melt unload following Liston and Elder (2006). Evapotranspiration over vegetated areas is calculated using the FAO Penman–Monteith approach (Allen et al., 1998).

The snow and ice surface energy balance is calculated as

$$Q_N + Q_H + Q_E + Q_A + Q_B + Q_M = 0, \quad (6)$$

with Q_N being the shortwave and longwave radiation balance, Q_H the sensible heat flux, Q_E the latent heat flux, Q_A the advective energy supplied by solid or liquid precipitation, Q_B the soil heat flux, and Q_M the energy potentially available for melt. For a detailed description of the calculation of the individual energy fluxes see Strasser (2008).

For the application in this study, the original model setup of AMUNDSEN was adapted to the mountain region of the Ötztal Alps by adding modules which enable a more realistic simulation of the catchment precipitation, of the timing of snowmelt (by considering cold content and liquid water content of the snowpack), and of runoff concentration. These are described in the following sections.

3.2 Precipitation correction

With respect to the undercatch of solid precipitation by common rain gauges (e.g., Sevruk, 1986; Goodison et al., 1998), a number of previous studies showed that the measured values of solid precipitation have to be corrected for systematic errors due to wetting loss, evaporation loss, and wind-induced undercatch (e.g., Rohrer and Braun, 1994; Farinotti et al., 2011; Schöber et al., 2014).

A common and straightforward method to apply a correction for snow undercatch is to introduce a fixed snow correction factor (SCF) which is applied to the fraction of precipitation identified as snow. However, it has been shown that errors due to wind-induced undercatch are especially large at lower temperatures, where snowfall mainly consists of smaller particles which are blown away more easily (e.g., Sevruk, 1983; Goodison et al., 1998). Using fixed SCF values thus tends to result in either underestimations of winter snowfall amounts or overestimations of snowfall events during spring and fall. A more robust method of snow correction is hence to introduce a variable correction factor derived from the meteorological variables (most importantly wind speed and temperature) measured at the gauge site.

For our study, we used the empirical correction for the Hellmann-type precipitation gauge presented in Goodison et al. (1998). Catch ratio (CR; i.e., the fraction of the actual precipitation amount that is captured by the gauge) in percent is thereby calculated as a function of air temperature T_a (°C) and wind speed (WS; m s^{-1}):

$$\mathrm{CR} = 96.63 + 0.41\mathrm{WS}^2 - 9.84\mathrm{WS} + 5.95T_a. \quad (7)$$

Adjusted precipitation values are then obtained by dividing the original values by CR.

This correction is applied to the station measurements at each model time step prior to the spatial interpolation of precipitation. However, since model results indicated that the simulated snowfall amounts were still underestimated, a post-interpolation adjustment of solid precipitation using a fixed SCF (for details see Sect. 5.1) was additionally implemented.

3.3 Snow redistribution

With the standard method for the spatial interpolation of point precipitation measurements implemented in AMUNDSEN, obtained interpolation values are influenced by the respective grid cell elevation and its distance to surrounding

weather stations. However, it is well known that snow accumulation patterns in complex terrain are fundamentally influenced by topographic controls beyond elevation alone, most importantly being attributed to redistribution of snow by wind and gravitational forces (McKay and Gray, 1981; Blöschl and Kirnbauer, 1992; Grünewald et al., 2014; Bernhardt et al., 2010; Warscher et al., 2013). Considering these processes is a prerequisite for reliable long-term mass balance simulations; hence the distribution of solid precipitation in AMUNDSEN has been updated using an empirical relation between SWE and topographic parameters.

Numerous studies have applied statistical models to explain snow cover variability using multiple regressions of topographic parameters such as elevation, slope, aspect, curvature, viewshed, and terrain roughness (e.g., Elder et al., 1991; Chang and Li, 2000; Pomeroy et al., 2002; Winstral et al., 2002; Lehning et al., 2011; Grünewald et al., 2013). However, Grünewald et al. (2013) showed that statistical relations between snow depth and topography are site-specific and performance decreases considerably when applying calibrated regression formulas to snow depth distributions in other catchments. Additionally, the topographic derivatives depend distinctly on the spatial scale used for calculation. Helfricht et al. (2014b) showed that the spatial variability of snow depth in a glacierized catchment is caused by a short-range variability based on small-scale terrain roughness, and a long-range variability with respect to the glacierized and wind-sheltered cirques and valleys in contrast to wind-exposed mountain ridges.

In this study we used topographic openness (Yokoyama et al., 2002) for the parameterization of the spatial snow distribution according to Helfricht (2014). Openness is a parameter originally developed to visualize topographic character and features in images, as it expresses the degree of dominance or enclosure of a location on an irregular surface. Topographic openness has two viewer perspectives in terms of positive and negative openness. Positive openness for a DEM grid point is obtained by averaging the zenith angles calculated for all eight compass directions from the grid point, while negative openness is obtained by averaging the respective nadir angles. The latter was used in this study to parameterize the spatial snow distribution, as it yields low values for convex forms and high values for concave forms. The openness values depend on the length scale L, which is the maximum distance considered for calculation: low L values result in a high spatial variability of openness, while high L values display large-scale topography of ridges and valley floors and hence highlight the overdeepening of, e.g., the surface elevations of glacier tongues compared to the surrounding ridges and peaks (Helfricht, 2014).

Negative openness Ψ_L was calculated for the entire Ötztal mountain range based on a 50 m DEM for $L = 50\,\mathrm{m}$ and $L = 5000\,\mathrm{m}$, according to the definitions given in Yokoyama et al. (2002) (resulting values for a subset of the area are shown in Fig. 2, top). A linear relation was applied between the minimum and the maximum threshold of negative openness to derive the snow redistribution factor (SRF), shown in Fig. 2 (bottom):

$$\widetilde{\Psi}_{50} = 3 \cdot (\Psi_{50} - 1.2), \tag{8}$$

$$\widetilde{\Psi}_{5000} = 3 \cdot (\Psi_{5000} - 1), \tag{9}$$

$$\mathrm{SRF} = \frac{1}{2}\left(\overset{\top 1.6}{\underset{\perp 0.1}{}}\widetilde{\Psi}_{50} + \overset{\top 1.6}{\underset{\perp 0.1}{}}\widetilde{\Psi}_{5000}\right), \tag{10}$$

where $\overset{\top 1.6}{\underset{\perp 0.1}{}}\widetilde{\Psi}$ denotes $\widetilde{\Psi}$ clipped to values between 0.1 and 1.6. Hence, at least 10 % of the initial atmospheric solid precipitation can be stored even in almost vertical slopes and in very exposed areas for the time of the precipitation event, while wind-sheltered areas can hold a maximum of 1.6 times the initial amount. At each time step, solid precipitation of each raster cell is multiplied with the corresponding redistribution factor. The new total amount of solid precipitation over the entire area is related to the initial precipitation amount in order to keep the total precipitation volume constant (mass conservation). Consequently, simulated accumulation is reduced in areas of low negative openness (i.e., exposed ridges, sheer rock faces) and increased in sheltered areas (i.e., cirques and low elevated valley floors). The two length scales and the factors in Eqs. (8)–(10) were determined by manual optimization for the best fit of the redistribution factor field with the ALS-derived surface elevation differences (interpreted as snow accumulation over the winter season).

To summarize, in total three steps of snow adjustments are applied: (i) the wind speed and temperature-dependent correction of measured precipitation at the meteorological stations (Eq. (7); variable in time and space), (ii) the additional post-interpolation snowfall adjustment using a fixed SCF (constant in time and space), and (iii) the adjustment using the SRFs acquired using Eq. (10) (constant in time, variable in space). Whereas the first two steps are required to correct precipitation input towards a realistic precipitation volume, the latter does not change the total volume but rather redistributes the solid precipitation with respect to the terrain.

3.4 Cold content and liquid water content

To account for temperature changes inside the snow cover, a parameterization for cold content and liquid water content based on the work of Braun (1984) has been added to the model. Meltwater is thereby not immediately removed from the snowpack, but a certain amount of liquid water (originating from surface melt or rain) can be retained in the snowpack. In the case of negative surface energy balances, this liquid water can refreeze. Further heat loss is used to build up a cold content, which needs to be depleted again before actual melt can occur. The cold content corresponds to the energy required to warm the entire (dry) snowpack to 0 °C. By relating this energy with the latent heat of fusion of ice ($333.7\,\mathrm{kJ\,kg^{-1}}$), cold content can alternatively also be expressed in units of water equivalent (corresponding to the

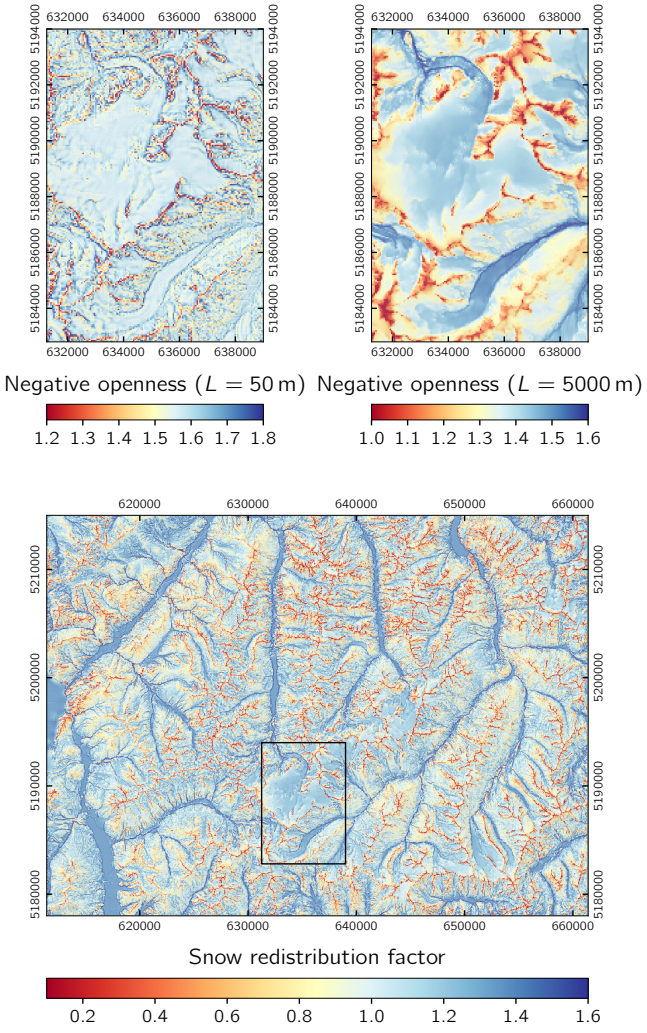

Figure 2. Negative openness of a subset of the study area for length scales $L = 50$ m and $L = 5000$ m (top), and the snow redistribution factor derived by a combination of both data sets (bottom; black frame shows the boundaries of the upper plots).

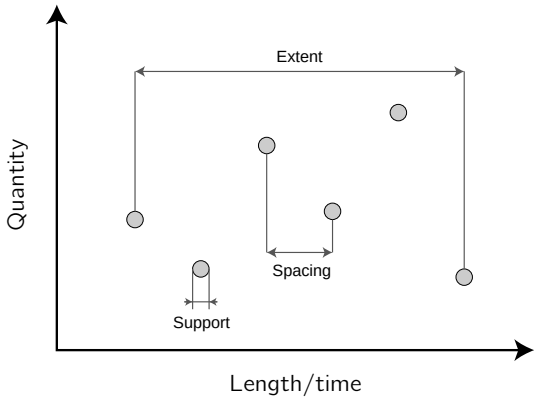

Figure 3. Illustration of the observation scale of a set of measurements (based on Blöschl and Sivapalan, 1995). Support is the integration volume or time of a single sample, spacing is the distance or time between individual samples, and extent is the total coverage in space or time of the entire data set.

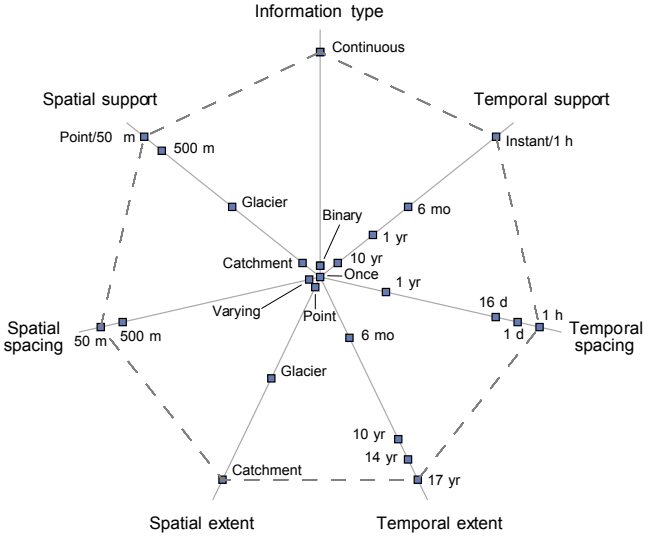

Figure 4. Radar chart axes for the visualization of the observation scale of the used validation data sets.

equivalent amount of liquid water needed to release the required amount of energy by freezing), which is the formulation we use within this study. The module for cold content and liquid water content is described in more detail in the Supplement.

Following Blöschl and Kirnbauer (1991), for the new snow and old snow layers the liquid water retention capacity of the snowpack was set to 10 % of the total snowpack weight, and the maximum possible cold content to 3 % of the total snowpack weight for this study. The refreezing factor (the amount of heat loss that is used to build up the cold content) was set to 0.65. This value, slightly higher than the literature value of 0.5, was determined by comparison of observed and simulated snow depth recordings.

3.5 Runoff concentration

For runoff concentration, a linear reservoir cascade approach (Nash, 1960) following Asztalos et al. (2007) was implemented in the model. Runoff originating from rainfall and from meltwater released from glaciers and the snowpack is thereby cumulated in each time step and catchment and routed through four parallel linear reservoir cascades for unglacierized areas, bare ice areas, firn-covered areas on glaciers, and snow-covered areas on glaciers (the latter consisting of the AMUNDSEN "new snow" and "old snow" layers, which are not treated separately in terms of runoff concentration). A constant fraction $f_{\text{glacierized}}$ of the inflow into the snow, firn, and ice reservoirs, as well as a fraction

Table 2. Parameter ranges for the runoff module calibration.

Parameter	Lower limit	Upper limit
k_{snow} (h)	5	20
k_{firn} (h)	5	20
k_{ice} (h)	1	5
$k_{\text{unglacierized}}$ (h)	5	20
k_{soil} (h)	10	100
n_{snow} (–)	1	20
n_{firn} (–)	2	6
n_{ice} (–)	1	4
$n_{\text{unglacierized}}$ (–)	1	7
n_{soil} (–)	2	10
$f_{\text{glacierized}}$ (–)	0.05	0.50
$f_{\text{unglacierized}}$ (–)	0.05	0.80

$f_{\text{unglacierized}}$ of the inflow into the unglacierized reservoir, is diverted into an additional soil reservoir. The parameters of the linear reservoir model (for each cascade the number of parallel reservoirs n and the storage constant k, as well as $f_{\text{glacierized}}$ and $f_{\text{unglacierized}}$) are determined by calibration separately for each catchment. Calibration is performed using an automatic optimization routine with the aim to maximize the objective function

$$\text{NSE}_V = \text{NSE} - 0.1|V_{\text{E}}|, \tag{11}$$

i.e., maximizing the Nash–Sutcliffe efficiency (NSE) and minimizing the relative volume error V_{E} (following Lindström (1997)). Table 2 shows the allowed ranges for the parameters during the automatic calibration.

4 Validation approach and data

In this study, the individual model components are validated against (i) mean areal precipitation values for the period 1997–2006 for the gauged catchments, (ii) daily to hourly snow depth recordings at five locations in the study area, (iii)–(iv) snow-covered area maps acquired by Landsat and MODIS imagery, respectively, (v) the snow accumulation distribution for the winter season 2010/2011 acquired using ALS surveys, (vi) glacier-averaged annual surface mass balances for three glaciers in the study area, (vii) spatially distributed glacier surface elevation changes for the entire area over the period 1997–2006, and (viii) hourly runoff records for eight of the investigated catchments. Besides the satellite-derived snow extent maps which only give binary information (snow yes/no), all other data sets also include volumetric information about the water resources in the study area (in varying spatial and temporal scales).

According to Blöschl and Sivapalan (1995), any finite set of observations is accompanied by an "observation scale" in space and time that can be defined by a "scale triplet" of support, spacing, and extent. Support is the integration volume or time of a single sample, spacing is the distance or time between individual samples, and extent is the total coverage in space or time of the entire data set. These concepts are illustrated in Fig. 3. Table 3 lists the spatial and temporal support, spacing, and extent of the used validation data sets, as well as the associated information type (binary vs. continuous information).

However, as the tabular representation of the scales makes it rather difficult to interpret and compare them, a visual method to display the observation scale of a data set is proposed: the seven dimensions (support, spacing, and extent in both space and time, as well as information type) are arranged as axes on a radar chart, where the left and right halves of the chart represent the spatial and temporal dimensions, respectively. The ranges of the individual axes and their ordering (from low to high or from high to low values) are designed such that a "perfect" validation data set (i.e., having the lowest possible support and spacing, the highest possible extent, and continuous information) is represented by a regular heptagon with the maximum possible diameter (i.e., extending to the maximum axis extent in each dimension). This implies that the ordering of the axis values is not consistent (extent ranges from low to high values, while support and spacing range from high to low values), but it allows for an easier visual interpretation (bigger is always better). Figure 4 shows the resulting axes including the possible values for our particular case study, while Fig. 5 shows the resulting charts for the eight validation data sets used in this study. As the charts are supposed mainly for a qualitative interpretation and comparison between the different data sets, they are shown without axis and tick labels in the latter.

The various data sets and the respective validation strategies are outlined in the following sections.

4.1 Areal precipitation

As a method to estimate long-term mean annual catchment precipitation for all gauged catchments to use as a validation data set for the respective AMUNDSEN simulation results, the OEZ approach (Kuhn and Batlogg, 1998; Kuhn, 2000) was used. This method calculates catchment-scale precipitation P from measured runoff Q, evaporation E, and glacier mass changes ΔS as the remainder from the water balance equation,

$$P = Q + E + \Delta S, \tag{12}$$

and has proven to be very robust for the simulation of decadal mean values (Kuhn, 2000). Evaporation E, typically comparatively small in glacierized catchments, is approximated using constant values depending on the surface type (ranging between 15 mm month^{-1} for snow and ice surfaces and 60 mm/month for forested areas). Liquid water storage within the snowpack and the ground is accounted for by the residuals of measured runoff vs. calculated monthly runoff originating from snow melt, ice melt, and precipitation mi-

Table 3. Information type (binary or continuous) as well as spatial and temporal support, spacing, and extent of the validation data sets used in this study.

		Space			Time		
Data set	Information type	Support	Spacing	Extent	Support	Spacing	Extent
Snow depth	continuous	point	varying	point	instant	1 d	17 years
Runoff	continuous	catchment	varying	catchment	1 h	1 h	17 years
MODIS snow-covered area	binary	500 m	500 m	catchment	instant	1 d	14 years
Landsat snow-covered area	binary	50 m	50 m	catchment	instant	16 d	17 years
Glacier mass balance	continuous	glacier	varying	glacier	1 year	1 year	17 years
Long-term glacier mass balance	continuous	50 m	50 m	glacier	10 years	once	10 years
ALS-based snow accumulation	continuous	50 m	50 m	catchment	6 months	once	6 months
Areal precipitation	continuous	catchment	varying	catchment	10 years	once	10 years

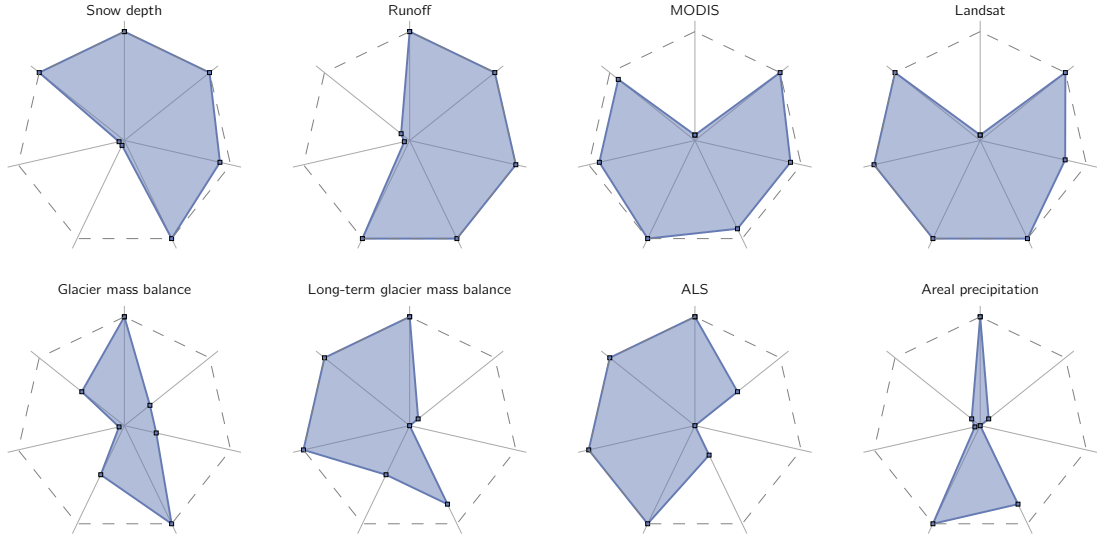

Figure 5. Radar charts (using the axis scales from Fig. 4) showing the observation scale of the validation data sets used in this study. The dashed gray heptagons represent an "optimal" validation data set.

nus evaporation (Kuhn et al., 2016). Interannual changes in ground water storage are neglected, as the method aims at mean annual values over longer periods.

To calculate areal precipitation values for our study catchments, for runoff Q the respective measurements at the catchment outlets were used, while glacier volume changes ΔS were derived from the surface elevation changes between 1997 and 2006 according to the glacier inventories performed in these years (Abermann et al., 2009).

4.2 Snow depth

Comparisons with snow depth measurements at the locations of meteorological stations allow to evaluate model performance in terms of realistic representation of accumulation (adequate correction for gauge undercatch, snow/rain separation) and ablation (surface albedo evolution, cold content and liquid water content, surface energy balance) at the point scale, as well the conversion of SWE to snow depth. The latter has, however, already been evaluated in previous studies (e.g., Marke et al., 2015) and has proven to be robust. Snow depth is evaluated at five locations covering an elevation range of almost 2000 m: Prutz (871 m a.s.l.), Nauders (1330 m a.s.l.), Obergurgl (1942 m a.s.l.), Weisssee (2480 m a.s.l.), and Pitztaler Gletscher (2864 m a.s.l.).

4.3 Snow distribution

4.3.1 Satellite-derived snow distribution

Satellite-derived snow cover images allow us to spatially validate simulation results in comparatively high spatial and temporal resolutions. For this study, we used a comprehensive set of Landsat (5/7) and MODIS scenes to derive snow extent maps for the study area. Landsat products are available in 30 m spatial resolution and a 16-day revisit time, while MODIS snow products are available daily from two satellites (Aqua and Terra), but with a coarser spatial resolution of

500 m. However, both Landsat and MODIS images are subject to frequent cloud obstructions, and hence only a limited subset of the available scenes was usable for this study.

With regard to Landsat data, a set of 26 suitable scenes covering the period 1998–2012 was manually selected (a list of the scenes can be found in the Supplement). Besides the requirement of no or minimum cloud coverage over the study area, images acquired in the early morning hours especially during winter were not taken into account, as illumination effects from low sun elevation angles often pose problems in terms of sensor saturation on sun-facing slopes and topographic shadowing, making it difficult to retrieve snow cover.

To derive snow maps from the raw Landsat bands, first the digital numbers of the individual bands were converted to top-of-atmosphere reflectances using the `i.landsat.toar` module from GRASS GIS (GRASS Development Team, 2012). Subsequently, the reflectances were topographically corrected using `i.topo.corr`. Then, the normalized difference snow index (NDSI) (Hall et al., 1995) was calculated using the band ratio

$$\mathrm{NDSI} = \frac{\mathrm{green} - \mathrm{SWIR}}{\mathrm{green} + \mathrm{SWIR}}, \tag{13}$$

where green and SWIR (shortwave infrared) correspond to TM bands 2 and 5, respectively, for both Landsat 5 and 7. Several NDSI thresholds for the binary snow/no snow classification were subsequently evaluated, but the commonly chosen value of 0.4 (Hall et al., 1995) provided adequate results and was thus used for all 26 scenes. Following Hall et al. (1998), additional to the NDSI threshold a threshold of the near-infrared (NIR) channel (TM band 4) was used to avoid misclassifying water bodies as snow – pixels with NIR < 11 % are thereby never classified as snow.

The applied NDSI threshold generally also classifies ice surfaces as snow. To discriminate snow from glacier ice, for all glacier pixels (according to the used glacier mask) broadband albedo α was calculated from Landsat TM bands 2 and 4 (green and NIR) following the relation by Knap et al. (2010):

$$\alpha = 0.726\alpha_2 - 0.322\alpha_2^2 - 0.051\alpha_4 + 0.581\alpha_4^2, \tag{14}$$

where $\alpha_2 = \mathrm{TM}\ 2$ and $\alpha_4 = \mathrm{TM}\ 4$. Then, for each scene an α threshold for discriminating snow from ice was manually derived using a histogram analysis of the α values over the glaciers. For 22 of the 26 scenes a threshold of $\alpha = 0.3$ was found to be applicable, while for the remaining scenes values of 0.25, 0.35, and 0.4 were used.

In contrast to Landsat, precomputed MODIS snow cover products are readily available for download (Hall et al., 2002). For our study we used the binary snow cover products MYD10A1 and MOD10A1 (for the Aqua and Terra satellites, respectively), which are available in 500 m resolution on a daily basis. To reduce the influence of clouds and misclassifications, Aqua and Terra scenes for each day were merged into a composite image: if a pixel in one scene was classified as cloud covered in one scene and cloud free in the other, the respective value of the cloud-free scene was taken, while if one pixel was considered snow covered in one scene and snow free in the other, the pixel was classified as snow. This approach is similar to methods applied in other studies (e.g., Xie et al., 2009; Gafurov and Bárdossy, 2009).

Table 4. Two-by-two contingency table for the comparison of binary snow cover observations and simulations (after Zappa (2008)). O and S denote observations and simulations, while the subscripts 0 and 1 correspond to snow-free and snow-covered situations, respectively.

	S_1	S_0	Σ
O_1	n_{11}	n_{01}	n_{x1}
O_0	n_{10}	n_{00}	n_{x0}
Σ	n_{1x}	n_{0x}	n_{xx}

MODIS snow products are calculated using a NDSI thresholding approach and hence also do not discriminate between snow and ice surfaces. Some studies have used the 250 m visible and near-infrared MODIS bands to classify snow and ice surfaces (e.g., Shea et al., 2013), but we found this method not applicable for our study, since the coarse resolution of the MODIS scenes makes it challenging to differentiate between snow and ice facies for the majority of (small) glaciers in the study area. As a pragmatic solution we excluded all MODIS scenes taken in the months July–September from the analyses, since it can be assumed that outside of this period the glaciers are snow covered and MODIS "snow" pixels are actually snow rather than ice.

To compare simulation results with the satellite products, the daily AMUNDSEN SWE maps were converted into binary snow cover images using an SWE threshold of 1 mm (this amount of snowfall is required to turn a summer landscape into "white winter" landscape). For the Landsat validation, the Landsat 30 m pixels were resampled to the 50 m model resolution, while for the MODIS validation the AMUNDSEN snow maps were resampled to the MODIS resolution using mode resampling, i.e., a 500 m pixel was classified as snow if at least 50 % of the 50 m pixels it comprises were snow covered.

For the comparison and evaluation of observed and simulated snow cover patterns, we use the contingency table-based efficiency criteria ACC, BIAS, and CSI (Zappa, 2008):

$$\mathrm{ACC} = \frac{n_{11} + n_{00}}{n_{xx}}, \tag{15}$$

$$\mathrm{BIAS} = \frac{n_{1x}}{n_{x1}}, \tag{16}$$

$$\mathrm{CSI} = \frac{n_{11}}{n_{xx} - n_{00}}. \tag{17}$$

For the definitions of n_{00} to n_{xx} see Table 4. Accuracy ACC is the number of correct predictions divided by the total num-

ber of samples (values between 0 and 1, with a perfect score being 1). This score, however, tends to be comparatively optimistic, as it usually yields high values both during winter (where most of the pixels are snow covered) and summer (where most of the pixels are snow free). A more sensitive score is the critical success index CSI, which is the number of correctly predicted snow events divided by the number of times where snow is predicted in the model and/or observed. Finally, BIAS corresponds to the frequency of correct snow predictions, i.e., the number of times where snow is present in the simulations divided by the number of times where it is observed. Again, a value of 1 is a perfect match, while here values below 1 indicate that snow cover is underrepresented in the model, and values above 1 indicate that the model overestimates the snow cover.

4.3.2 ALS-derived snow distribution

While satellite-derived snow cover maps can be used to analyze binary snow coverage (snow yes/no) in high temporal resolution, lidar-derived surface elevation fields allow us to obtain snow depth maps in very high spatial resolution. For this purpose, two lidar surveys are required: one for mapping the snow-free terrain and one for recording the snow-covered terrain. The difference between both surface elevations can be interpreted as snow depth (in case of glacier surfaces, the vertical component of the ice flow has to be evaluated and considered for error analysis; e.g., Sold et al., 2013; Helfricht et al., 2014a).

ALS surveys of a 746 km^2 large area encompassing all investigated catchments were performed for the winter 2010/2011 to determine the spatial snow distribution and depth at the end of the accumulation season. The two ALS surveys were performed from 7 to 12 October 2010 and 20 to 23 April 2011, respectively. An Optech ALTM Gemini device was used for both surveys, with a mean flight speed of 65 $m\,s^{-1}$ and a mean flight altitude above ground of 1000 m. Mean point densities of 3.6 and 3.8 points per square meter, respectively, were achieved. DEMs with 1 m resolution were processed and surface elevation changes were calculated in the original 1 m resolution. The DEM differences were then resampled to the grid size of 50 m according to the model resolution. On the basis of the determined distribution of snow depth, SWE was calculated using an empirical relation between snow depth SD_{ALS} (m) and snow density ρ_{ALS} ($kg\,m^{-3}$) derived by Schöber et al. (2014):

$$\rho_{ALS} = SD_{ALS} \cdot 14.8 + 347. \quad (18)$$

This relation has already been applied for the investigation area in previous studies and has proven to be robust, with average errors on the order of 10 % (Schöber, 2014).

4.4 Glacier mass balance

For three glaciers in the study region (Hintereisferner, Kesselwandferner, and Vernagtferner), long-term annual mass balance series are available. For these glaciers, the simulated specific annual mass balance is compared with the observations, as well as the cumulative values over the entire simulation period 1997–2013. In addition, using two DEMs of the study region acquired in 1997 and 2006 spatially distributed glacier surface elevation changes during this period were calculated, which allows the assessment of the performance of the multiannual simulated mass balance for all glaciers in the study region over this period. The respective snow/ice volume changes as simulated by AMUNDSEN are thereby converted to surface elevation changes using the respective layer densities as calculated by the model (see Sect. 3.1).

4.5 Runoff

Hydrological data from eight runoff gauges with adequate hourly records are used. Runoff is modeled in an hourly resolution for the time period 1997–2013. The period 1998–2006 is used for model calibration, while the period 2007–2013 is used for the validation of simulated runoff. Model performance is evaluated using the efficiency criteria percent bias (PBIAS) and NSE, as well as the benchmark efficiency (BE) (Schaefli and Gupta, 2007):

$$\mathrm{PBIAS} = 100 \cdot \frac{\sum_t \left(Q_{\mathrm{sim},t} - Q_{\mathrm{obs},t}\right)}{\sum_t Q_{\mathrm{obs},t}}, \quad (19)$$

$$\mathrm{NSE} = 1 - \frac{\sum_t \left(Q_{\mathrm{obs},t} - Q_{\mathrm{sim},t}\right)^2}{\sum_t \left(Q_{\mathrm{obs},t} - \overline{Q}_{\mathrm{obs}}\right)^2}, \quad (20)$$

$$\mathrm{BE} = 1 - \frac{\sum_t \left(Q_{\mathrm{obs},t} - Q_{\mathrm{sim},t}\right)^2}{\sum_t \left(Q_{\mathrm{obs},t} - Q_{\mathrm{bench},t}\right)^2}. \quad (21)$$

For BE, the benchmark model runoff Q_{bench} is calculated as the multiannual mean observed runoff per calendar day and hour.

5 Results and discussion

5.1 Precipitation

Mean areal precipitation for the gauged catchments in the region for the period 1997–2006 as derived by closing the water balance using the OEZ method (see Sect. 4.1) was compared to the respective AMUNDSEN simulation results.

As the applied correction of the precipitation recordings using wind speed and temperature (Eq. 7) still resulted in an underestimation of precipitation amounts (as indicated by comparisons with (i) the catchment precipitation simulated by closing the water balance (Table 5), (ii) the temporal

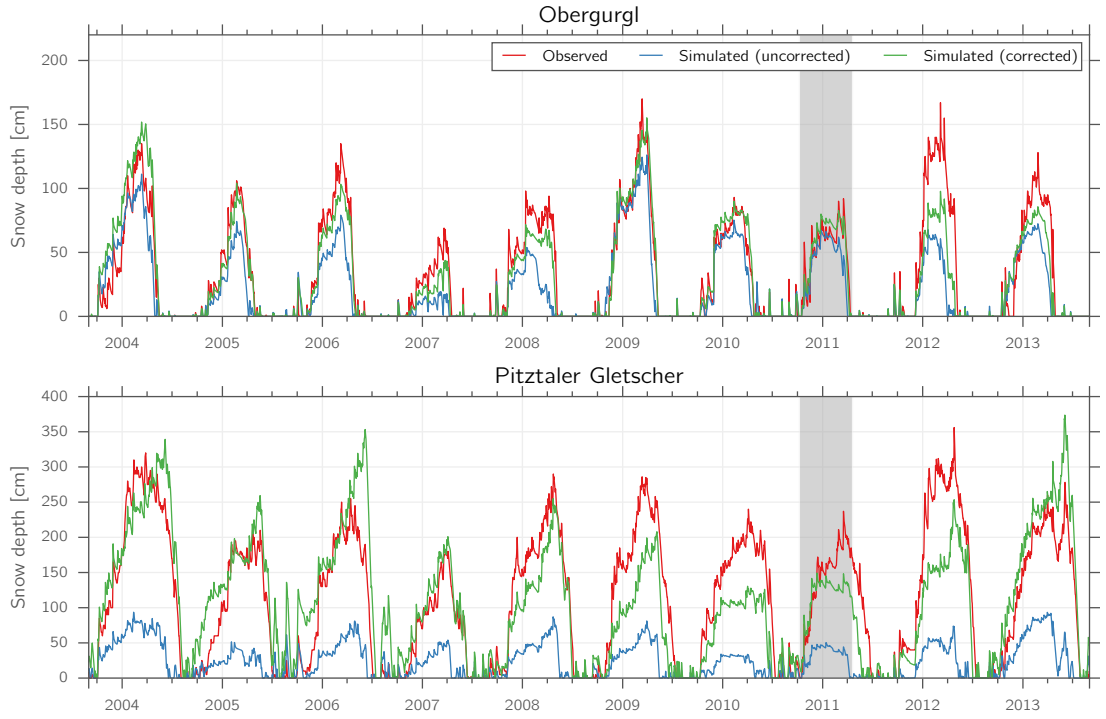

Figure 6. Observed and simulated (using uncorrected and corrected precipitation, respectively) snow depth for stations Obergurgl (1942 m a.s.l.) and Pitztaler Gletscher (2864 m a.s.l.). The gray shaded area corresponds to the period in between the two ALS surveys.

Table 5. Mean annual areal precipitation for the gauged catchments in the study area as calculated by closing the water balance (OEZ method) and AMUNDSEN (using uncorrected and corrected precipitation, respectively) for the period 1997–2006.

ID	Catchment	Area (km^2)	Precipitation (mm)	Precipitation AMUNDSEN uncorrected (mm)	Difference uncorrected (%)	Precipitation AMUNDSEN corrected (mm)	Difference corrected (%)
1	Rofenache	98.6	1507	1056	−30.0	1627	8.0
2	Am Barst/Gurgler Ache	72.4	1745	1137	−34.8	1617	−7.3
3	Niedertalbach	66.7	1410	1111	−21.2	1593	13.0
4	Gepatschalm	53.9	1799	1079	−40.0	1606	−10.7
6	Taschachbach	44.7	1632	1034	−36.6	1523	−6.7
7	Pitze	27.0	1552	1056	−32.0	1706	9.9
8	Radurschlbach	24.0	1402	882	−37.1	1113	−20.6
9	Tscheybach	16.4	1309	905	−30.9	1130	−13.6
11	Venter Ache	15.0	1253	1030	−17.8	1473	17.5
12	Verpeilbach	12.1	1093	1135	3.8	1360	24.4
13	Fissladbach	11.4	1481	1147	−22.5	1473	−0.6
14	Poschach/Gurgler Ache	8.1	1584	1019	−35.6	1441	−9.0
15	Platzertal	7.9	1395	1073	−23.1	1399	0.3
17	Ferwallbach	6.3	1537	1094	−28.8	1489	−3.1
18	Königsbach	5.9	1554	1089	−30.0	1490	−4.1
19	Rostizbach	4.8	1413	1135	−19.6	1465	3.7
	Total	475.2	1545	1067	-30.9	1543	−0.2

progress of snow depth (Fig. 6), (iii) the spatial distribution of accumulation at the end of the winter season (Fig. 10), and (iv) the long-term glacier mass changes in the Ötztal Alps; Fig. 15), an additional post-interpolation SCF (i.e., a constant factor with which the interpolated snowfall field is multiplied in each time step) was applied. By comparing precipitation amounts derived by closing the water balance with the respective AMUNDSEN results over all gauged catch-

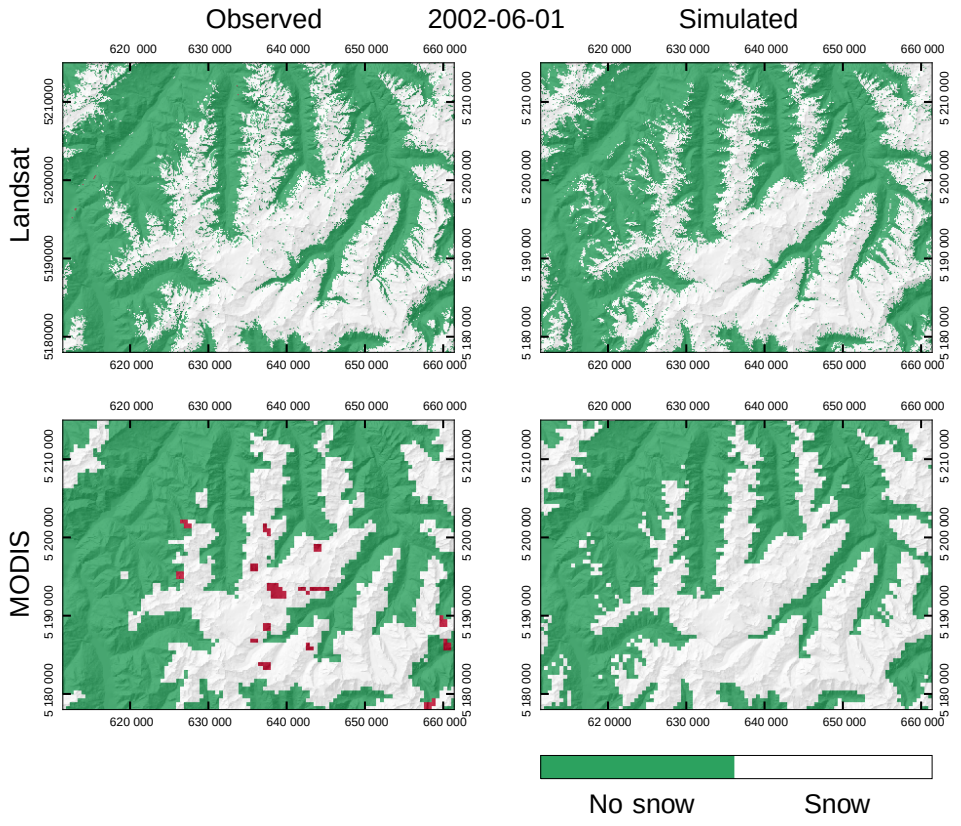

Figure 7. Observed and simulated snow cover distribution for 1 June 2002, using Landsat (top, 50 m resolution) and MODIS (bottom, 500 m resolution) scenes.

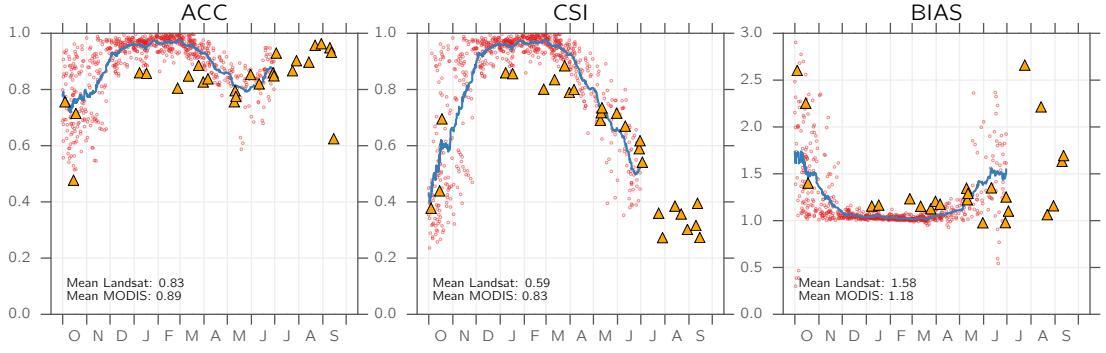

Figure 8. ACC, CSI, and BIAS for all selected Landsat (triangles) and MODIS (red circles; blue line is the 30-day running mean) scenes, sorted by calendar date.

ments, this factor was set to a value of 1.15 (i.e., 15 % increase of snowfall amounts). Using this additional adjustment, AMUNDSEN mean areal precipitation for all gauged catchments deviates less than 1 % from the respective values as acquired by closing the water balance, as compared to a difference of −31 % when using uncorrected values (Table 5). Especially considering that precipitation amounts for the largest catchments are reproduced very well, these results can be considered satisfactory. The largest differences occur in the medium-sized catchments, with the maximum deviation being approx. 25 %, while the smaller catchments generally show the smallest deviations.

In several of the following sections, results obtained using uncorrected and corrected precipitation are compared and evaluated. "Uncorrected" in this terminology corresponds to the elevation-dependent remapping of unaltered precipitation recordings alone, while "corrected" refers to the combination of the adjustment of snowfall amounts (Eq. (7) and the addi-

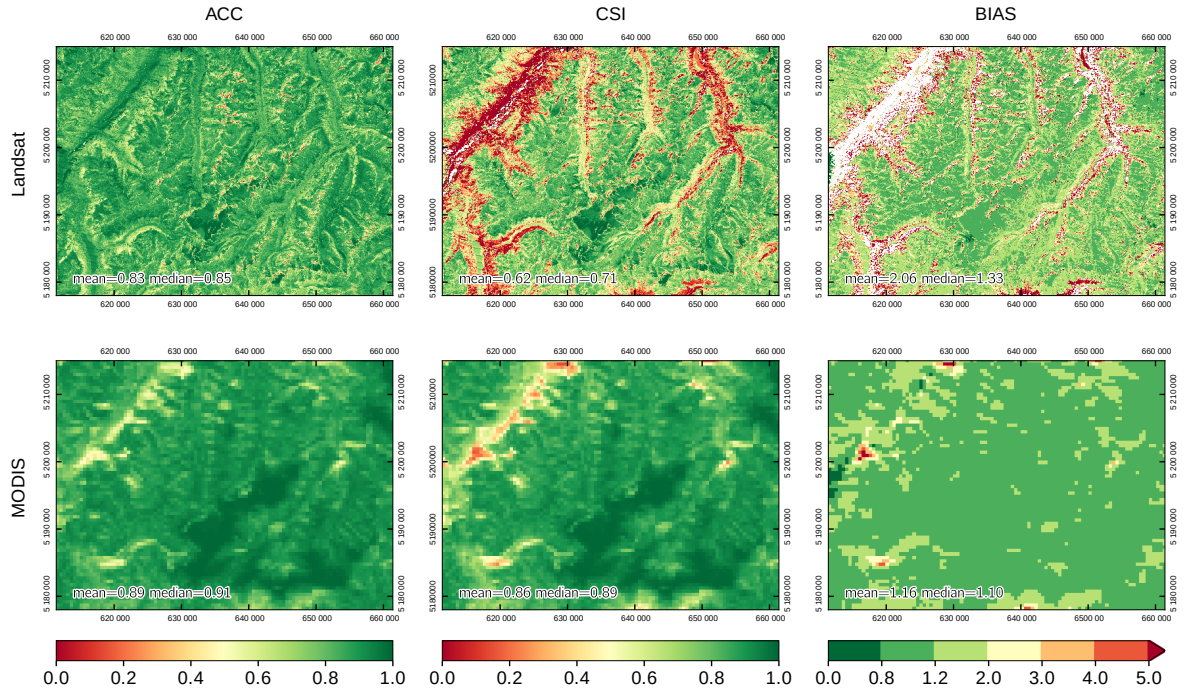

Figure 9. Pixel-based statistics of ACC, CSI, and BIAS over all selected Landsat and MODIS scenes.

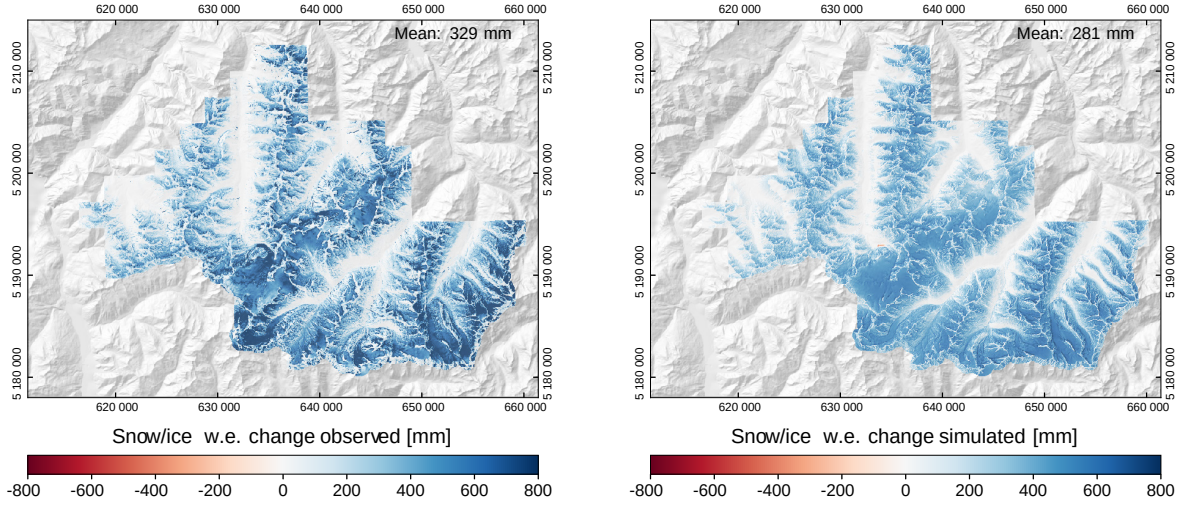

Figure 10. Observed (i.e., surface elevation differences converted to SWE using Eq. (18)) and simulated end-of-season snow distribution for the winter 2010/2011 (8 October to 22 April).

tional increase by 15 %) and the topographic snow redistribution; see Eq. (10).

5.2 Snow depth

Table 6 lists the performance measures R^2, NSE, and PBIAS for the comparison of daily snow depth observations at the point scale to the respective simulation results for five stations located in elevations between 871 and 2864 m a.s.l. for the model runs using uncorrected and corrected precipitation, respectively.

At the stations Obergurgl, Weisssee, and Pitztaler Gletscher, using corrected precipitation leads to significantly improved results in terms of all three performance measures. At the low-elevated stations Prutz and Nauders, the applied precipitation correction leads to a severe overestimation of snow depth of 223 and 120 %, respectively, on average. However, while these stations are within the simulation area (i.e., the rectangle surrounding the study catchments shown in

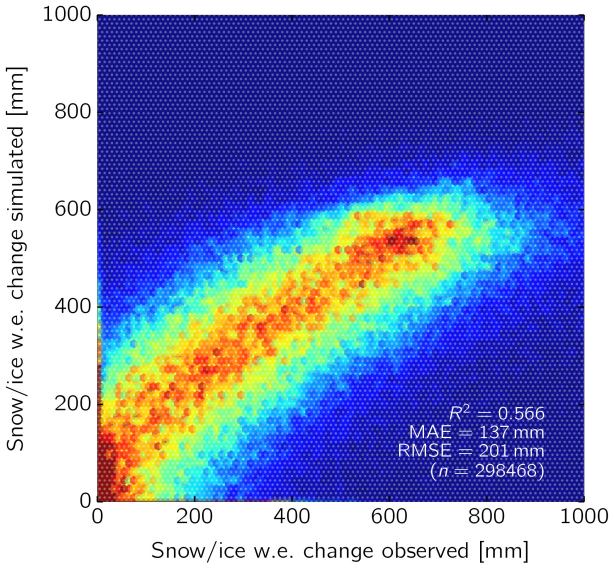

Figure 11. Binned scatter plot (colors represent point density) of observed vs. simulated water equivalent differences for the winter 2010/2011 (8 October to 22 April).

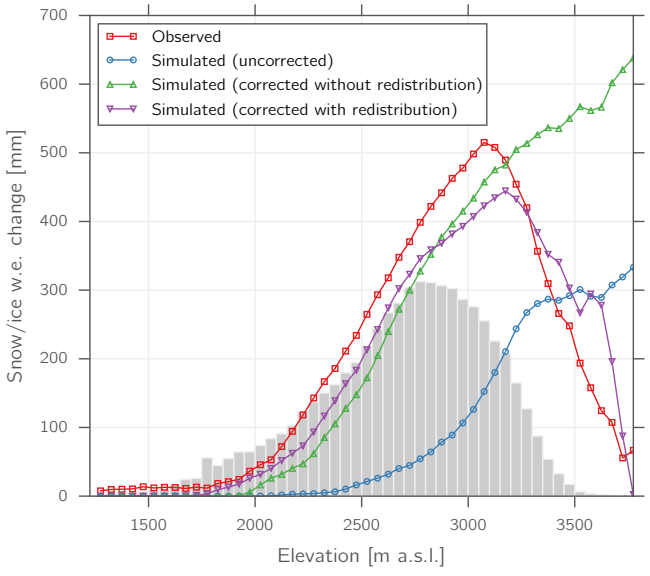

Figure 12. Observed vs. simulated water equivalent differences for the winter 2010/2011 (8 October to 22 April) by elevation (50 m bands). Simulation results are displayed for using uncorrected precipitation (elevation-dependent remapping only), corrected precipitation without topographic redistribution (Eq. (7) plus constant snow correction factor (15 %)), and corrected precipitation with topographic redistribution (Eq. (10)). The gray bars represent the relative area distribution of the elevation bands.

Fig. 1), they are located outside of the investigated catchments and are significantly lower elevated (minimum elevation of the study catchments is 1760 m).

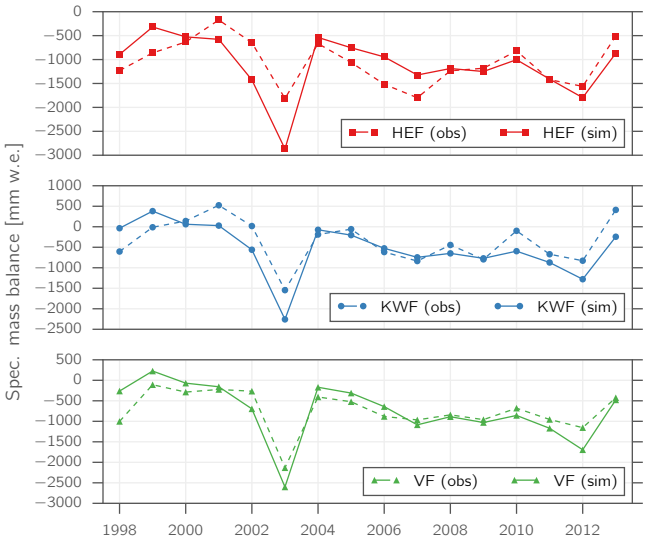

Figure 13. Observed (dashed) and simulated (solid) specific mass balance for Hintereisferner (HEF, top), Kesselwandferner (KWF, center), and Vernagtferner (VF, bottom) in the period 1997/1998–2012/2013.

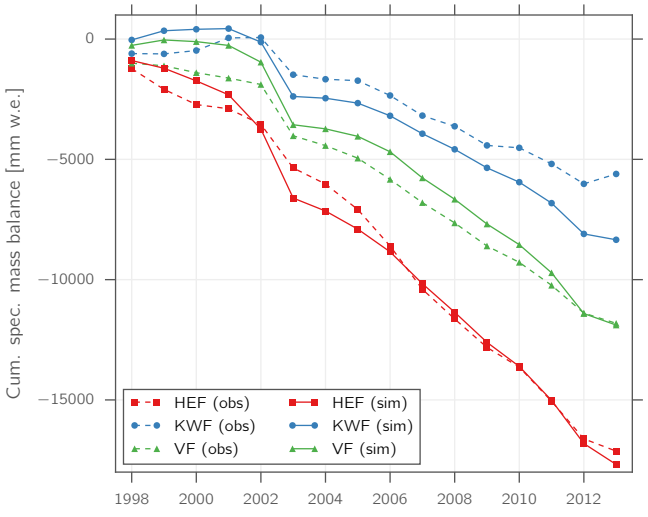

Figure 14. Observed (dashed) and simulated (solid) cumulative specific mass balance for Hintereisferner (HEF), Kesselwandferner (KWF), and Vernagtferner (VF) in the period 1997/1998–2012/2013.

Figure 6 shows observed and simulated (using both uncorrected and corrected precipitation, respectively) snow depth for the stations Obergurgl and Pitztaler Gletscher and the period 2003–2013. While some uncertainty also can be attributed to the applied snow densification parameterization for the conversion from water equivalent to snow depth, these results show that the applied precipitation corrections (correction for undercatch as well as topography-based correction) considerably improve results especially at high elevations, as can be seen for the station Pitztaler Gletscher. Nev-

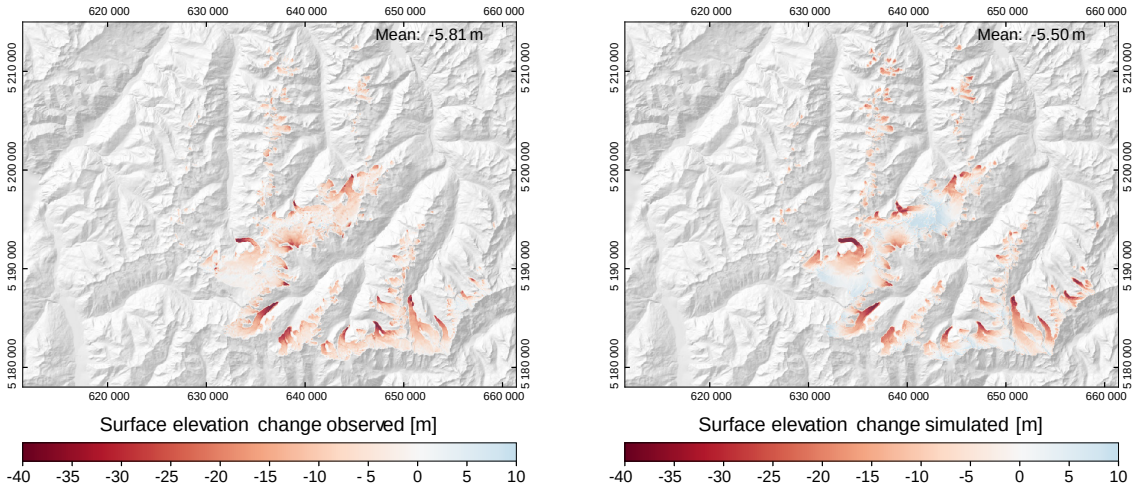

Figure 15. Observed and simulated glacier surface elevation change for the period 1997–2006.

Table 6. R^2, Nash–Sutcliffe efficiency (NSE), and percent bias (PBIAS) for observed vs. simulated snow depth obtained using uncorrected and corrected precipitation, respectively.

Station	Elevation (m a.s.l.)	Period	Uncorrected			Corrected		
			R^2	NSE	PBIAS (%)	R^2	NSE	PBIAS (%)
Prutz	871	2005–2013	0.62	0.60	10.3	0.66	−2.21	222.5
Nauders	1330	2004–2013	0.84	0.74	−29.2	0.74	−0.35	119.5
Obergurgl	1942	1999–2013	0.65	0.60	−28.8	0.85	0.85	−6.0
Weisssee	2480	2006–2013	0.39	−0.13	−69.6	0.76	0.52	28.7
Pitztaler Gletscher	2864	1997–2013	0.57	−0.52	−76.8	0.71	0.68	7.1

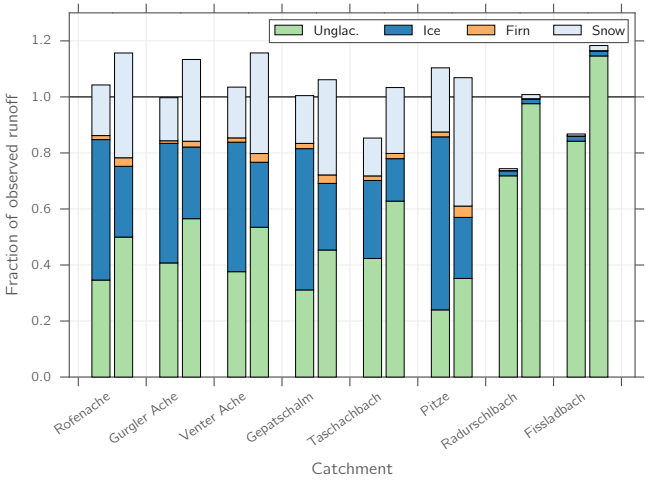

Figure 16. Runoff components (prior to the diversion into the soil reservoir) as simulated by AMUNDSEN for the calibration period 1998–2006 shown as fractions of the mean observed runoff in the same period. The two bars for each catchment show the results for the model runs using uncorrected (left) and corrected (right) precipitation, respectively.

ertheless, at this station snow depth is still underestimated in most seasons despite using corrected precipitation. For the medium-elevated station Obergurgl, while the differences between uncorrected and corrected precipitation are less distinct, the applied corrections result in a very satisfying representation of snow depth over the course of the seasons.

5.3 Snow distribution

5.3.1 Comparison with satellite data

For the comparison of simulated snow distributions, 26 cloud-free Landsat scenes as well as daily MODIS scenes (starting in the year 2000) were obtained. With regard to the latter, as outlined in Sect. 4.3.1 images acquired in the months July–September were discarded due to difficulties in discriminating snow and ice surfaces. Additionally, only scenes with less than 5 % total cloud coverage over the study area were considered. This left 733 usable MODIS scenes, corresponding to approx. one image every 5 days (in the period October–June) on average.

Figure 7 exemplarily shows the observed and simulated snow cover distributions for the Landsat and MODIS scenes from 1 June 2002. With regard to the observations, the ef-

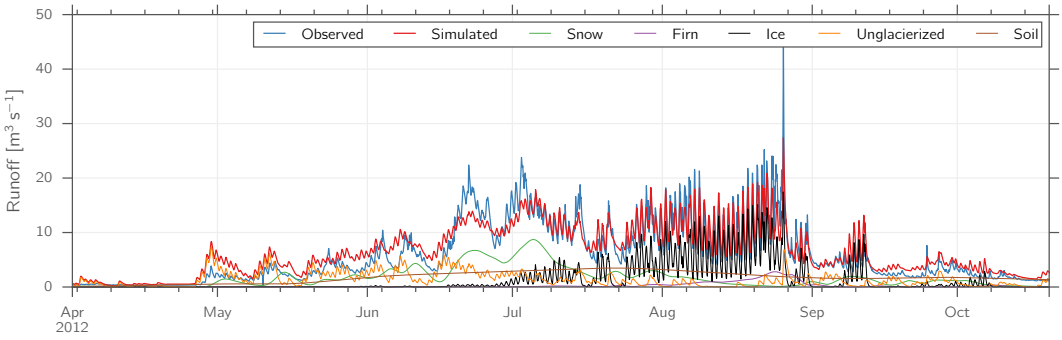

Figure 17. Observed vs. simulated runoff (including individual runoff components) for gauge Gepatschalm (1893–3526 m a.s.l.) in the period April–October 2012.

fect of the reduced resolution of the MODIS product is clearly visible – while the Landsat image also highlights small-scale variations in snow cover (e.g., snow-free ridges, small snow patches), a large amount of detail is smoothed out in the MODIS image. This is also illustrated in the total snow-covered area – in the Landsat image, 45 % of the total area are classified as snow covered, as compared to 37 % for the MODIS image. Simulation results agree well with the observations for this date, with regard to both the visual comparison and the performance measures; for Landsat (MODIS), ACC = 0.85 (0.87), CSI = 0.72 (0.73), and BIAS = 0.98 (1.18).

Figure 8 shows the obtained values for ACC, CSI, and BIAS for all selected Landsat and MODIS scenes sorted by calendar date, as well as the 30-day running means for the MODIS values (blue lines). Generally, the average performance measures (displayed in the bottom left of each figure) show better results for MODIS, but it has to be noted that the average values are not directly comparable, since in the case of Landsat they are calculated from significantly fewer samples and additionally – in contrast to MODIS – also include scenes acquired during the summer months. As expected, for all three measures performances are very good during the winter months (particularly January to March), where most of the area is snow covered. However, they deteriorate during spring and fall. ACC shows the highest performance values, with on average approx. 90 % of all pixels being correctly classified for MODIS data and 83 % for Landsat data. Values decrease during spring and fall, but are rarely (in 9 % of the MODIS scenes and 8 % of the Landsat scenes) below 0.7. The range of CSI values in contrast is much larger, with – when looking at the MODIS 30-day running mean – an approximately linear increase/decrease during fall and spring from values around 0.4 in late spring to approx. 0.95 during winter. Lowest values (between 0.25 and 0.4) are obtained for the eight Landsat scenes during the summer months. With regard to BIAS, values are very close to the optimal value of 1 during winter. Outside of this period, larger deviations occur. It is notable, that – with very few exceptions – BIAS values are always above 1, meaning that the model tends to overestimate the snow cover. The particularly high values during fall might indicate that rainfall events are frequently misinterpreted as snowfall events in the model, as well as early season snowfall events that rapidly melt again in reality, while in the model the snow lasts on the ground. However, if only little snow is present in the observation, even comparatively moderate overestimations in the model can lead to high BIAS values, which is also an explanation for the particularly high values of this parameter for some of the Landsat scenes during summer.

The performance measures can be calculated not only space integrated (i.e., for each scene) but also time integrated, i.e., calculating them on a pixel-by-pixel basis over all scenes. The results are shown in Fig. 9 for Landsat at the top and MODIS at the bottom. For all three performance measures, best results are obtained on the glaciers and high-elevated regions (where a snow cover is present most of the time). Largest deviations occur in forested areas, which becomes particularly visible for the CSI and BIAS scores. Satellite-based snow cover mapping in forests is notoriously difficult and associated with larger errors than in the open due to the obstruction of snow on the ground by the trees. While in the MODIS snow mapping algorithm in forested areas the NDSI threshold is lowered in order to increase the classification accuracy (Klein et al., 1998), no such distinction was made for the Landsat snow classification performed for this study, which might be one reason for the lower model skill for Landsat as compared to MODIS in these areas. For most non-forested pixels, ACC and CSI values are above 0.7 for both Landsat and MODIS, underlining the satisfying model performance. BIAS is in the range of 20 % over- or underestimation for most MODIS pixels, while in the Landsat comparison larger areas fall into the class of BIAS values up to 2. Once again it has to be noted, however, that most of the Landsat scenes were acquired during spring or summer, where a correct representation of snow/rain distinction is especially crucial; hence the lower performance scores for Landsat are not unexpected.

To summarize, these results show that the total snow-covered area is frequently overestimated, on average by 20 % (60 %) for MODIS (Landsat). Whereas in only 1 % (0 %) of all investigated scenes snow cover is underestimated by more than 5 %, it is overestimated by more than 5 % in 51 % (92 %) of all scenes. The largest mismatches between the observations and the simulations occur during the summer months (where accurately reproducing the snow cover is the most challenging), followed by fall and spring. Elevation-wise the largest errors occur below the lowest-elevated runoff gauge (further analyses are shown in the Supplement). A likely explanation for this effect (which is also observed in the snow depth comparisons) is that snow correction and snow redistribution factors are overestimated in these low elevations.

5.3.2 Comparison with ALS data

Figure 10 shows the observed and simulated snow distributions (i.e., observed surface elevation differences between the ALS acquisition dates converted to SWE using Eq. (18) and simulated SWE differences for the same period) for the winter 2010/2011. A 746 km^2 large area was covered by the ALS acquisitions, resulting in a sample size of 298 468 pixels (with 50 m resolution) available for model validation. As seen in Fig. 10, despite the applied precipitation corrections the total snow volume is still underrepresented in the model by approx. 15 % in this particular period; however, the applied snow redistribution parameterization leads to a significantly improved representation of the accumulation patterns (snow-free ridges, increased accumulation on glacierized areas) as compared to model results using elevation-dependent precipitation fields only. Figure 11 shows the high correlation of simulated with observed snow water equivalent values with an R^2 value of 0.57. This is a considerable improvement in model performance considering the results using uncorrected precipitation (elevation-dependent remapping only) ($R^2 = 0.07$) and using undercatch-corrected precipitation without topographic redistribution ($R^2 = 0.23$) (plots are shown in the Supplement). Figure 12 shows the SWE differences by elevation (50 m bands) for the ALS data, the simulation results without any precipitation adjustment (elevation-dependent remapping only), the simulation results derived using corrected precipitation but no topographic redistribution (Eq. (7) and constant SCF; 15 %), and the simulation results derived using corrected precipitation and topographic redistribution (Eq. 10). From this figure it is evident that the snowfall adjustments strongly increase model performance in the lower and medium elevations (where snow accumulation is dramatically underestimated in the simulation run without precipitation correction), but only the simulation run including topographic redistribution is able to reproduce the observed decline in SWE following the peak at approx. 3000 m a.s.l.

The obtained correlations are in the range of the results by Schöber et al. (2014), who obtained $R^2 = 0.52$ in a 166 km^2 large subcatchment of our study region, also using a cell size of 50 m. Grünewald et al. (2013) obtained R^2 values between 0.30 and 0.91 for different investigation areas between 1.5 and 28 km^2, but using a coarser cell size of 400 m, and Schirmer et al. (2011) obtained $R^2 = 0.42$ for a site in Switzerland using 10 m resolution (however, both of these studies used statistical models based on topographic parameters alone to model snow depth distribution and did not employ snow cover models). Hence, considering the size of the investigated area and the high spatial resolution, our results can be considered very satisfactory. However, again it has to be emphasized that – as also shown in other studies (e.g., Grünewald et al., 2013) – the relations between snow depth and topography are site-specific and hence have to be calibrated for different study regions individually. In the case of the openness-based correction applied in this study, this concerns the choice of the scale parameters L for negative openness, the value range of negative openness for the different scale parameters (Eqs. 8 and 9), and the final relation to precipitation correction (Eq. 10).

5.4 Glacier mass balance

Figures 13 and 14 show the annual and cumulative mass balances for Hintereisferner (HEF), Kesselwandferner (KWF), and Vernagtferner (VF). The individual annual mass balances of the three glaciers are reproduced reasonably well by the model, with $R^2 = 0.46$ for HEF, $R^2 = 0.64$ for KWF, and $R^2 = 0.81$ for VF. The cumulative mass balance over the period 1997–2013 is captured very well for VF and HEF, with differences of only 39 mm (0.3 %) for VF and −377 mm (2.1 %) for HEF, while KWF shows the lowest performance with a deviation of −2633 mm (32 %). For all three glaciers, results improve with increasing simulation time, possibly due to the improved representation of the interpolated meteorological fields due to increased station density (cf. Schöber, 2014, p. 40 f.). A meteorological station is located in direct vicinity to VF, which is a possible explanation for the distinctively good performance of the model with regard to this glacier. The largest errors in mass balance simulations for all three glaciers occur in the year 2003. Further investigation reveals that this can at least partly be traced back to the meteorological forcing data, specifically two precipitation events in October and November 2002 which have not been captured by the precipitation recordings at the station Vernagtbach. Excluding the precipitation recordings from this station in the period October–November 2002 distinctly improves the simulated mass balances for the glaciological year 2003 (a more detailed analysis and further results are shown in the Supplement).

The generally satisfying model performance for the three glaciers indicates that the model setup is suitable for glacier mass balance simulations at the regional scale, as no glacier-specific model calibration has been performed. This is further underlined when looking at the long-term glacier sur-

Table 7. NSE, BE, and PBIAS of simulated vs. observed runoff obtained using uncorrected and corrected precipitation. Calibration and validation periods are – if not otherwise stated – 1998–2006 and 2007–2013, respectively. The subscripts c and v denote calibration and validation periods, respectively.

		Uncorrected						Corrected					
ID	Catchment	NSE_c	NSE_v	BE_c	BE_v	$PBIAS_c$ (%)	$PBIAS_v$ (%)	NSE_c	NSE_v	BE_c	BE_v	$PBIAS_c$ (%)	$PBIAS_v$ (%)
1	Rofenache	0.81	0.75	0.11	0.30	6.1	0.1	0.81	0.78	0.56	0.35	12.1	20.0
2	Am Barst/Gurgler Ache[1]	0.81	0.77	0.25	0.28	−3.0	0.3	0.82	0.80	0.46	0.26	5.6	20.8
3	Niedertalbach[2]	0.79	–	0.07	–	7.6	–	0.79	–	0.47	–	14.1	–
4	Gepatschalm	0.86	0.84	0.11	0.23	3.6	−3.5	0.88	0.87	0.57	0.50	4.9	7.3
6	Taschachbach[3]	0.85	0.82	0.28	0.18	−14.2	−18.7	0.90	0.88	0.49	0.45	3.4	2.2
7	Pitze[3]	0.85	0.84	−0.60	−0.24	10.3	2.9	0.88	0.87	0.61	0.53	6.3	4.4
8	Radurschlbach[4]	0.46	0.56	−1.09	−0.80	−31.9	−10.9	0.66	0.67	−0.26	−0.43	−8.4	22.7
13	Fissladbach[4]	0.60	0.68	−0.46	−0.10	−17.4	−3.0	0.67	0.63	−0.16	−0.23	12.8	31.9
	Mean	0.75	0.75	-0.17	-0.02	-4.86	-4.69	0.80	0.79	0.34	0.20	6.4	15.6

[1] No runoff observations for 2013.
[2] No runoff observations for 2003–2013.
[3] No runoff observations for 2009–2013.
[4] No runoff observations for 2011–2013.

face elevation changes in the period 1997–2006. The respective observations are derived from the glacier inventories of the respective years and allow to spatially assess the model performance in simulating the surface mass balance for all glaciers in the study region over this period. The mean surface elevation change derived from the glacier inventory was −5.81 m between 1997 and 2006, while surface elevation changes simulated with AMUNDSEN amount to −5.50 m for the same period (Fig. 15, $R^2 = 0.44$). Hence, the mean value of simulated multiannual snow and ice water resources loss is approx. 0.3 m (5 %) more positive compared to the observed value. This relatively low value is on the same order of magnitude as the uncertainties arising from glacier volume change measurements. Hence, the model can be considered capable of realistically simulating the runoff contribution by glacier melt. Possible explanations for the deviation – beside the measurement uncertainty – are processes such as basal melt or collapsing glacier tongues, which are not represented in the model. Additionally, the albedo of the dirty glacier tongues covered with mineral dust and cryoconite might be overestimated by the model. Since no ice flow is considered in AMUNDSEN up to now, the spatial distribution of glacier thickness changes is biased at the low elevated glacier tongues and in high elevated accumulation areas. Although glacier ice flow velocities slowed down in the past decades and the contribution of vertical ice flow is low at most of the glaciers in the study region (e.g., Helfricht et al., 2014a), surface elevation changes caused by ice flow accumulate to considerable values in multiannual time periods. Hence, in order to prevent systematic model biases in long-term applications it is crucial to integrate a dynamic tracking of the glacier surface, which is proposed to be implemented in a future version of AMUNDSEN.

5.5 Runoff

For the runoff simulations, calibration and validation periods were generally 1998–2006 and 2007–2013, respectively, but for some catchments only shorter time series of runoff measurements were available. Table 7 shows the results (NSE, BE, and PBIAS) of the simulations for the catchments in which respective hourly runoff observations were available, while Fig. 16 shows the fractions of the individual runoff components as simulated by AMUNDSEN using uncorrected and corrected precipitation, respectively, for the calibration period (1998–2006). Figure 17 exemplarily shows the observed and simulated runoff as well as the individual runoff components for the gauge Gepatschalm (1893–3526 m a.s.l.) and the year 2012.

By means of the applied precipitation correction, NSE values for almost all catchments slightly improve as compared to the model run using uncorrected precipitation, with values of around or above 0.8, with the exception of two small catchments with very low glacierization (because the runoff module is specifically designed for glacierized catchments it is expected to perform worse in catchments with no or low glacierization). However, for most catchments even the use of uncorrected precipitation leads to comparatively high NSE values. This is typical for catchments with a strong runoff seasonality: the fact that NSE uses the mean observed runoff as reference makes it a rather inefficient skill score in these cases – high NSE values can be obtained as soon as the seasonality is represented in the model, regardless of whether smaller-scale fluctuations are misrepresented (Schaefli and Gupta, 2007). Looking at BE, variations between catchments as well as the effects of the applied precipitation correction are much more pronounced. Using corrected precipitation significantly increases BE in almost all catchments in both the calibration and validation period. Despite using corrected precipitation, in the two smallest and least glacierized gauged catchments BE values are below zero, indicating that

the multiannual mean per calendar day and hour as predictor would be better than the model. When looking not only at the respective single best-performing parameter set (with regard to Eq. 11) for each catchment but also at a range of similarly performing parameter sets, it shows that similarly performing parameter sets result in a remarkably large spread of BE (results shown in the Supplement). However, the best-performing parameter sets in terms of Eq. (11) are also among the best-performing sets in terms of BE.

The total runoff volume bias PBIAS is (absolutely) smaller for the model runs with uncorrected precipitation (with runoff on average underestimated by approx. 5 % in both the calibration and the validation period) than when corrected precipitation is used, where runoff is on average overestimated by 6.4 and 15.6 %, respectively, with even higher values for the two largest catchments. This can also be seen in Fig. 16, which also illustrates that underestimated precipitation amounts are compensated by increased ice melt amounts in the model runs using uncorrected precipitation. The calibration criterion $NSE_{\sqrt{}}$ (Eq. (11), values not shown in Table 7) shows little variation in all catchments, with mean values between 0.74 and 0.79 for both runs (uncorrected vs. corrected) and in both the calibration and validation period.

6 Conclusions

The physically based hydroclimatological model AMUNDSEN was set up and applied for a partly glacierized, 558 km^2 large region in the Ötztal Alps (Austria) on a 50 m grid. Model validation was performed by comparing mean areal precipitation, point-based snow depth recordings, remotely sensed snow extent maps from MODIS and Landsat satellites, lidar-derived surface elevation changes for the accumulation season 2010/2011, glacier-averaged annual surface mass balances for three glaciers, spatially distributed glacier surface elevation changes for the entire area over the period 1997–2006, and hourly runoff observations for eight catchments. A correction function for solid precipitation has been added to the model in order to compensate for the severe undercatch of precipitation measurements in high elevated and wind-exposed areas. This pre-interpolation correction of the recorded precipitation depths using temperature and wind speed has proven to significantly improve model results, but an additional SCF of 15 % had to be applied to the resulting snowfall fields. A parameterization for cold content and liquid water content of the snowpack has been implemented to account for the delayed start of melting of the snowpack following cold periods and nights with clear-sky conditions. Finally, a parameterization for snow redistribution using topographic openness has been added to the model, helping to significantly better represent the complex snow accumulation patterns in the study site as compared to assuming an elevation dependency alone.

The results of the study especially emphasize the importance of a systematic model validation based on multiple data sources. When looking solely at runoff, also using entirely implausible precipitation inputs for the model leads to acceptable results as the respective missing snow melt inputs are compensated by enlarged ice melt contributions, even though the spatial patterns of snow accumulation and snow depth amounts are misrepresented in the model. Hence, in order to produce realistic contributions of the individual runoff components, it is necessary to assess and validate the individual model components and processes separately: areal catchment precipitation derived from the hydrological balance can be used to derive realistic total precipitation volumes. Point-based snow depth observations are necessary for the evaluation of the simulated temporal evolution of the snowpack in terms of the correct timing and magnitude of accumulation and melt. Binary snow extent maps allow a multitemporal evaluation of the simulated snow distribution, while lidar-derived surface elevation differences add detailed information about the amount of snow resources stored in the area during the winter. Finally, annual glacier mass balances from field surveys as well as long-term mass balances derived from DEM differencing allow the evaluation of model performance in reproducing the long-term surface mass balance over the glaciers, which is essential prior to applying the model for scenario studies. Currently, an ice flow parameterization is being implemented in the model, which will allow us to apply the model setup for climate change scenarios in order to estimate future runoff conditions and glacier retreat in the study site.

While the value of using multiple data sets for model validation has already been investigated in previous studies, our concept of multilevel spatiotemporal validation represents a new approach for assessing glacio-hydrological simulation accuracy on a fully comprehensive level. Usually, exemplary comparisons of single simulated model variables are utilized for validation, either (i) with time series of local recordings of the respective variable, (ii) with discrete (in time) shots of the spatial distribution of the variable, or (iii) in which runoff is compared to gauge recordings. The latter is a measure that summarizes all relevant processes (of accumulation, redistribution and ablation of snow and ice in our case) but does not allow for separating these processes. Multilevel spatiotemporal validation is a systematic, independent, complete, and – here – even fully redundant validation procedure for the simulated key variables in the model. The spatial distribution of the simulation result is validated in all six dimensions of the observation scale – temporal as well as spatial support, spacing, and extent – by means of two independent satellite data products. Likewise, the water volume (i.e., the mass balance) of the simulation results is validated by at least two independent validation data sources for each dimension, e.g., spatial spacing (in high model resolution) by long-term glacier mass balances and ALS. Redundancy of the validation data sources will only seldom be available in most ap-

plications and in practice. It may serve here as additional proof of robustness and reliance for both the simulation results and the validation concept. This concept may serve as a next piece for a new generation of integrated environmental system models with coupled components, separately validated, and defined interface design in between them (Strasser et al., 2014).

The Supplement related to this article is available online at doi:10.5194/tc-10-1859-2016-supplement.

Acknowledgements. This work was carried out within the framework of the projects "H03 MUSICALS A – Multiscale Snow/Ice Melt Discharge Simulation for Alpine Reservoirs", carried out in the research programme of alpS – Centre for Climate Change Adaptation in Innsbruck, and "hiSNOW – High resolution monitoring and modelling under climate change conditions – Combining ALS and TLS data acquisition with energy and mass balance modelling at Hochjochferner/Val Senales, Italy", financed by the Autonomous Province of Bozen/Bolzano. The authors want to thank the COMET research programme of the Austrian Research Promotion Agency (FFG), the TIWAG – Tiroler Wasserkraft AG, and the Autonomous Province of Bozen/Bolzano. Meteorological and hydrological data were provided by the Central Institution for Meteorology and Geodynamics (ZAMG), the Hydrographic Service of Tyrol, the TIWAG – Tiroler Wasserkraft AG, the Commission for Glaciology of the Bavarian Academy of Sciences and Humanities, and the Autonomous Province of Bozen/Bolzano. We also thank Johannes Schöber for valuable discussions during the preparation of the manuscript. Finally, we gratefully acknowledge the valuable and constructive comments by Bettina Schaefli and one anonymous referee which helped to improve the manuscript.

Edited by: V. Radic
Reviewed by: B. Schaefli and one anonymous referee

References

Abermann, J., Lambrecht, A., Fischer, A., and Kuhn, M.: Quantifying changes and trends in glacier area and volume in the Austrian Ötztal Alps (1969–1997–2006), The Cryosphere, 3, 205–215, doi:10.5194/tc-3-205-2009, 2009.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration – Guidelines for computing crop water requirements, Tech. rep., 1998.

Anderson, E. A.: A point energy and mass balance model of a snow cover, Tech. Rep. NWS 19, NOAA, 1976.

Asztalos, J., Kirnbauer, R., Escher-Vetter, H., and Braun, L.: A distributed energy balance snowmelt model as a component of a flood forecasting system for the Inn river, in: Alpine Snow Workshop, Munich, 9–17, 2007.

Beniston, M.: Clim. Change in Mountain Regions: A Review of Possible Impacts, Clim. Change, 59, 5–31, 2003.

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.

Beven, K.: Prophecy, reality and uncertainty in distributed hydrological modelling, Adv. Water Resour., 16, 41–51, 1993.

Bliss, A., Hock, R., and Radić, V.: Global response of glacier runoff to twenty-first century climate change, J. Geophys. Res.-Earth, 119, 717–730, 2014.

Blöschl, G. and Kirnbauer, R.: Point snowmelt models with different degrees of complexity – internal processes, J. Hydrol., 129, 127–147, 1991.

Blöschl, G. and Kirnbauer, R.: An analysis of snow cover patterns in a small alpine catchment, Hydrol. Process., 6, 99–109, 1992.

Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling: A review, Hydrol. Process., 9, 251–290, 1995.

Braun, L. N.: Simulation of snowmelt-runoff in lowland and lower alpine regions of Switzerland, Ph.D. thesis, ETH Zurich, 1984.

Chang, K.-T. and Li, Z.: Modelling snow accumulation with a geographic information system, Int. J. Geograph. Inf. Sci., 14, 693–707, 2000.

Collins, D. N.: Climatic warming, glacier recession and runoff from Alpine basins after the Little Ice Age maximum, Ann. Glaciol., 48, 119–124, 2008.

Corripio, J.: Modelling the energy balance of high altitude glacierised basins in the Central Andes, Ph.D. thesis, University of Edinburgh, 2002.

Deems, J. S., Fassnacht, S. R., and Elder, K. J.: Fractal Distribution of Snow Depth from Lidar Data, J. Hydrometeorol., 7, 285–297, 2006.

Elder, K., Dozier, J., and Michaelsen, J.: Snow accumulation and distribution in an alpine watershed, Water Resour. Res., 27, 1541–1552, 1991.

Farinotti, D., Usselmann, S., Huss, M., Bauder, A., and Funk, M.: Runoff evolution in the Swiss Alps: projections for selected high-alpine catchments based on ENSEMBLES scenarios, Hydrol. Process., 26, 1909–1924, 2011.

Finger, D., Pellicciotti, F., Konz, M., Rimkus, S., and Burlando, P.: The value of glacier mass balance, satellite snow cover images, and hourly discharge for improving the performance of a physically based distributed hydrological model, Water Resour. Res., 47, W07519, doi:10.1029/2010WR009824, 2011.

Finger, D., Vis, M., Huss, M., and Seibert, J.: The value of multiple data set calibration versus model complexity for improving the performance of hydrological models in mountain catchments, Water Resour. Res., 51, 1939–1958, 2015.

Fischer, A., Seiser, B., Stocker Waldhuber, M., Mitterer, C., and Abermann, J.: Tracing glacier changes in Austria from the Little Ice Age to the present using a lidar-based high-resolution glacier inventory in Austria, The Cryosphere, 9, 753–766, doi:10.5194/tc-9-753-2015, 2015.

Gafurov, A. and Bárdossy, A.: Cloud removal methodology from MODIS snow cover product, Hydrol. Earth Syst. Sci., 13, 1361–1373, doi:10.5194/hess-13-1361-2009, 2009.

Goodison, B. E., Louie, P., and Yang, D.: WMO solid precipitation measurement intercomparison, Tech. Rep. WMO/TD 872, Geneva, 1998.

GRASS Development Team: Geographic Resources Analysis Support System (GRASS GIS) Software, Open Source Geospatial Foundation, 2012.

Grayson, R. B., Blöschl, G., Western, A. W., and McMahon, T. A.: Advances in the use of observed spatial patterns of catchment hydrological response, Adv. Water Resour., 25, 1313–1334, 2002.

Grünewald, T. and Lehning, M.: Altitudinal dependency of snow amounts in two small alpine catchments: can catchment-wide snow amounts be estimated via single snow or precipitation stations?, Ann. Glaciol., 52, 153–158, 2011.

Grünewald, T., Schirmer, M., Mott, R., and Lehning, M.: Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment, The Cryosphere, 4, 215–225, doi:10.5194/tc-4-215-2010, 2010.

Grünewald, T., Stötter, J., Pomeroy, J. W., Dadic, R., Moreno Baños, I., Marturià, J., Spross, M., Hopkinson, C., Burlando, P., and Lehning, M.: Statistical modelling of the snow depth distribution in open alpine terrain, Hydrol. Earth Syst. Sci., 17, 3005–3021, doi:10.5194/hess-17-3005-2013, 2013.

Grünewald, T., Bühler, Y., and Lehning, M.: Elevation dependency of mountain snow depth, The Cryosphere, 8, 2381–2394, doi:10.5194/tc-8-2381-2014, 2014.

Hall, D. K., Riggs, G. A., and Salomonson, V. V.: Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data, Remote Sens. Environ., 54, 127–140, 1995.

Hall, D. K., Foster, J. L., Verbyla, D. L., Klein, A. G., and Benson, C. S.: Assessment of snow-cover mapping accuracy in a variety of vegetation-cover densities in central Alaska, Remote Sens. Environ., 66, 129–137, 1998.

Hall, D. K., Riggs, G. A., Salomonson, V. V., DiGirolamo, N. E., and Bayr, K. J.: MODIS snow-cover products, Remote Sens. Environ., 83, 181–194, 2002.

Hanzer, F., Marke, T., and Strasser, U.: Distributed, explicit modeling of technical snow production for a ski area in the Schladming region (Austrian Alps), Cold Reg. Sci. Technol., 108, 113–124, 2014.

Helfricht, K.: Analysis of the spatial and temporal variation of seasonal snow accumulation in Alpine catchments using airborne laser scanning, Ph.D. thesis, Innsbruck, 2014.

Helfricht, K., Schöber, J., Seiser, B., Fischer, A., Stötter, J., and Kuhn, M.: Snow accumulation of a high alpine catchment derived from LiDAR measurements, Adv. Geosci., 32, 31–39, 2012.

Helfricht, K., Kuhn, M., Keuschnig, M., and Heilig, A.: Lidar snow cover studies on glaciers in the Ötztal Alps (Austria): comparison with snow depths calculated from GPR measurements, The Cryosphere, 8, 41–57, doi:10.5194/tc-8-41-2014, 2014a.

Helfricht, K., Schöber, J., Schneider, K., Sailer, R., and Kuhn, M.: Interannual persistence of the seasonal snow cover in a glacierized catchment, J. Glaciol., 60, 889–904, 2014b.

Horton, P., Schaefli, B., Mezghani, A., Hingray, B., and Musy, A.: Assessment of climate-change impacts on alpine discharge regimes with climate model uncertainty, Hydrol. Process., 20, 2091–2109, 2006.

Huss, M. and Farinotti, D.: Distributed ice thickness and volume of all glaciers around the globe, J. Geophys. Res.-Earth, 117, F04010, doi:10.1029/2012JF002523, 2012.

Huss, M., Zemp, M., Joerg, P. C., and Salzmann, N.: High uncertainty in 21st century runoff projections from glacierized basins, J. Hydrol., 510, 35–48, 2014.

Jansson, P., Hock, R., and Schneider, T.: The concept of glacier storage: a review, J. Hydrol., 282, 116–129, 2003.

Jonas, T., Marty, C., and Magnusson, J.: Estimating the snow water equivalent from snow depth measurements in the Swiss Alps, J. Hydrol., 378, 161–167, 2009.

Jordan, R.: A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM, 89,, Tech. rep., Hanover, NH, 1991.

Klein, A. G., Hall, D. K., and Riggs, G. A.: Improving snow cover mapping in forests through the use of a canopy reflectance model, Hydrol. Process., 12, 1723–1744, 1998.

Knap, W. H., Reijmer, C. H., and Oerlemans, J.: Narrowband to broadband conversion of Landsat TM glacier albedos, Int. J. Remote Sens., 20, 2091–2110, 2010.

Konz, M. and Seibert, J.: On the value of glacier mass balances for hydrological model calibration, J. Hydrol., 385, 238–246, 2010.

Kuhn, M.: Verification of a hydrometeorological model of glacierized basins, Ann. Glaciol., 31, 15–18, 2000.

Kuhn, M. and Batlogg, N.: Glacier Runoff In Alpine Headwaters In A Changing Climate, IAHS publication no. 248, p. 79, 1998.

Kuhn, M., Helfricht, K., Ortner, M., Landmann, J., and Gurgiser, W.: Liquid water storage in snow and ice in 86 Eastern Alpine basins and its changes from 1970–1997 to 1998–2006, Ann. Glaciol., CJO 2016, doi:10.1017/aog.2016.24, 2016.

Lehning, M., Grünewald, T., and Schirmer, M.: Mountain snow distribution governed by an altitudinal gradient and terrain roughness, Geophys. Res. Lett., 38, L19504, doi:10.1029/2011GL048927, 2011.

Lindström, G.: A Simple Automatic Calibration Routine for the HBV Model, Nord. Hydrol., 28, 153–168, 1997.

Liston, G. E. and Elder, K.: A distributed snow-evolution modeling system (SnowModel), J. Hydrometeorol., 7, 1259–1276, 2006.

Magnusson, J., Farinotti, D., Jonas, T., and Bavay, M.: Quantitative evaluation of different hydrological modelling approaches in a partly glacierized Swiss watershed, Hydrol. Process., 25, 2071–2084, 2011.

Marke, T.: Development and Application of a Model Interface to couple Land Surface Models with Regional Climate Models for Climate Change Risk Assessment, Ph.D. thesis, Ludwig-Maximilians-Universität München, 2008.

Marke, T., Strasser, U., Hanzer, F., Stötter, J., Wilcke, R. A. I., and Gobiet, A.: Scenarios of Future Snow Conditions in Styria (Austrian Alps), J. Hydrometeorol., 16, 261–277, 2015.

McKay, G. A. and Gray, D. M.: The distribution of snowcover, in: Handbook of Snow, edited by: Gray, D. M. and Male, D. H., 153–190, Pergamon Press, Toronto, 1981.

Nash, J. E.: A unit hydrograph study, with particular reference to British catchments, Proc. Inst. Civ. Eng., 17, 249–282, 1960.

Parajka, J. and Blöschl, G.: The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models, J. Hydrol., 358, 240–258, 2008.

Pellicciotti, F., Brock, B., Strasser, U., Burlando, P., Funk, M., and Corripio, J.: An enhanced temperature-index glacier melt model including the shortwave radiation balance: Development and testing for Haut Glacier d'Arolla, Switzerland, J. Glaciol., 51, 573–587, 2005.

Pomeroy, J. W., Gray, D. M., Hedstrom, N. R., and Janowicz, J. R.: Prediction of seasonal snow accumulation in cold climate forests, Hydrol. Process., 16, 3543–3558, 2002.

Refsgaard, J. C.: Parameterisation, calibration and validation of distributed hydrological models, J. Hydrol., 198, 69–97, 1997.

Rohrer, M. B.: Die Schneedecke im schweizerischen Alpenraum und ihre Modellierung, Zürcher Geographische Schriften, 49, 178 pp., 1992.

Rohrer, M. B. and Braun, L. N.: Long-Term Records of Snow Cover Water Equivalent in the Swiss Alps, in: EGS XVII General Assembly, 65–78, Edinburgh, 1994.

Salzmann, N., Huggel, C., Rohrer, M., and Stoffel, M.: Data and knowledge gaps in glacier, snow and related runoff research – A climate change adaptation perspective, J. Hydrol., 518, 225–234, 2014.

Schaefli, B. and Gupta, H. V.: Do Nash values have value?, Hydrol. Process., 21, 2075–2080, 2007.

Schaefli, B. and Huss, M.: Integrating point glacier mass balance observations into hydrologic model identification, Hydrol. Earth Syst. Sci., 15, 1227–1241, doi:10.5194/hess-15-1227-2011, 2011.

Schaefli, B., Hingray, B., Niggli, M., and Musy, A.: A conceptual glacio-hydrological model for high mountainous catchments, Hydrol. Earth Syst. Sci., 9, 95–109, doi:10.5194/hess-9-95-2005, 2005.

Schirmer, M., Wirz, V., Clifton, A., and Lehning, M.: Persistence in intra-annual snow depth distribution: 1. Measurements and topographic control, Water Resour. Res., 47, W09516, doi:10.1029/2010WR009426, 2011.

Schöber, J.: Improved snow and runoff modelling of glacierized catchments for flood forecasting, Ph.D. thesis, Innsbruck, 2014.

Schöber, J., Schneider, K., Helfricht, K., Schattan, P., Achleitner, S., Schöberl, F., and Kirnbauer, R.: Snow cover characteristics in a glacierized catchment in the Tyrolean Alps – Improved spatially distributed modelling by usage of Lidar data, J. Hydrol., 519, 3492–3510, 2014.

Sevruk, B.: Correction of Measured Precipitation in the Alps Using the Water Equivalent of New Snow, Nord. Hydrol., 14, 49–58, 1983.

Sevruk, B.: Correction of precipitation measurements: Swiss experience, in: Precipitation Measurement, edited by: Sevruk, B., 187–196, Swiss Federal Institute of Technology, Zurich, 1986.

Shea, J. M., Menounos, B., Moore, R. D., and Tennant, C.: An approach to derive regional snow lines and glacier mass change from MODIS imagery, western North America, The Cryosphere, 7, 667–680, doi:10.5194/tc-7-667-2013, 2013.

Sold, L., Huss, M., Hoelzle, M., Andereggen, H., Joerg, P. C., and Zemp, M.: Methodological approaches to infer end-of-winter snow distribution on alpine glaciers, J. Glaciol., 59, 1047–1059, 2013.

Strasser, U.: Spatial and temporal variability of meteorological variables at Haut Glacier d'Arolla (Switzerland) during the ablation season 2001: Measurements and simulations, J. Geophys. Res., 109, D03103, doi:10.1029/2003JD003973, 2004.

Strasser, U.: Modelling of the mountain snow cover in the Berchtesgaden National Park, Tech. Rep. 55, Berchtesgaden, 2008.

Strasser, U., Bernhardt, M., Weber, M., Liston, G. E., and Mauser, W.: Is snow sublimation important in the alpine water balance?, The Cryosphere, 2, 53–66, doi:10.5194/tc-2-53-2008, 2008.

Strasser, U., Vilsmaier, U., Prettenthaler, F., Marke, T., Steiger, R., Damm, A., Hanzer, F., Wilcke, R. A. I., and Stötter, J.: Coupled component modelling for inter- and transdisciplinary climate change impact research: Dimensions of integration and examples of interface design, Environ. Model. Softw., 60, 180–187, 2014.

Warscher, M., Strasser, U., Kraller, G., Marke, T., Franz, H., and Kunstmann, H.: Performance of complex snow cover descriptions in a distributed hydrological model system: A case study for the high Alpine terrain of the Berchtesgaden Alps, Water Resour. Res., 49, 2619–2637, 2013.

Winstral, A., Elder, K., and Davis, R. E.: Spatial Snow Modeling of Wind-Redistributed Snow Using Terrain-Based Parameters, J. Hydrometeorol., 3, 524–538, 2002.

Xie, H., Wang, X., and Liang, T.: Development and assessment of combined Terra and Aqua snow cover products in Colorado Plateau, USA and northern Xinjiang, China, J. Appl. Remote Sens., 3, 14 pp., 2009.

Yokoyama, R., Shirasawa, M., and Pike, R. J.: Visualizing topography by openness: a new application of image processing to digital elevation models, Photogram. Eng. Remote Ssens., 68, 257–266, 2002.

Zappa, M.: Objective quantitative spatial verification of distributed snow cover simulations – an experiment for the whole of Switzerland, Hydrol. Sci. J., 53, 179–191, 2008.