Snow model comparison to simulate snow depth evolution and sublimation at point scale in the semi-arid Andes of Chile

Annelies Voordendag[1*], Marion Réveillet[2**], Shelley MacDonell[2], and Stef Lhermitte[1]

[1]Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands
[2]Centro de Estudios Avanzados en Zonas Áridas (CEAZA), ULS-Campus Andrés Bello, Raúl Britan 1305, La Serena, Chile
[*]now at: Department of Atmospheric and Cryospheric Sciences (ACINN), University of Innsbruck, Innsbruck, Austria
[**]now at: Univ. Grenoble Alpes, Université de Toulouse, Météo-France, CNRS, CNRM, Centre d'Etudes de la Neige, 38100 Grenoble, France

Correspondence: Shelley MacDonell (shelley.macdonell@ceaza.cl) and Stef Lhermitte (s.lhermitte@tudelft.nl)

Abstract. Physically-based snow models provide valuable information on snow cover evolution and are therefore key to provide water availability projections. Yet, uncertainties related to snow modelling remain large as a result of differences in the representation of snow physics and meteorological forcing. While many studies focus on evaluating these uncertainties, issues still arise, especially in environments where sublimation is the main ablation process. This study evaluates a case study in the semi-arid Andes of Chile and aims to compare two snow models with different complexities, SNOWPACK and SnowModel, at a local point, over one snow season ~~. Their sensitivity relative i) to physical calibration for albedo and fresh snow density and ii) to forcing perturbation is evaluated based on ensemble approaches~~and to evaluate their sensitivity relative to parameterization and forcing. For that purpose, the two models are forced with i) the most ideal set of input parameters, ii) an ensemble of different physical parameterizations and iii) an ensemble of biased forcing. Results indicate ~~larger uncertainty depending on the model calibration than between the two models (even though the~~ large uncertainties depending on forcing, the snow roughness length z_0, albedo parameterization and fresh snow density parameterization. The uncertainty caused by the forcing is direcly related to the bias chosen. Even though the models show significant differences in their physical complexity~~). We also confirm the importance of albedo parameterization, even though ablation is driven by sublimation. SnowModel is particularly sensitive to this choice as it strongly affects both the sublimation and the melt rates. However, the day of snow-free snow surface is not sensitive to the parameterization as it only varies every eight days. The albedo parameterization of SNOWPACK has stronger consequences on melt at the end of the season leading to a date difference of the end of the season of 41 days. However, despite these differences, the~~ , the snow model choice is of least importance, as the sensitivity of both models to the forcing data was in the same order of magnitude and highly influenced by the precipitation uncertainties. The sublimation ratio ranges are in agreement for the two models: ~~42.7-63.5~~36.4 to 80.7% for SnowModel and ~~51.3 and 64.6~~36.3 to 86.0% for SNOWPACK, and are related to the albedo ~~calibration~~ parameterization and snow roughness length choice for the two models. ~~Finally, the sensitivity of both models to the forcing data was in the same order of magnitude and highly influenced by the precipitation uncertainties.~~

1 Introduction

Snow models provide valuable information on snow cover evolution and are therefore key to quantify runoff and provide accurate water availability projections. Several models, with different complexities in the representation of different snow processes, from empirical to ~~physical~~ physically-based approaches, have been developed to simulate snow depth changes. Empirical approaches, such as degree-day models (e.g. Braithwaite and Olesen, 1989; Hock, 2003) are based on a simple statistical relationship to positive air temperatures to simulate snow melt. Comparatively, physically-based approaches consider all energy flux exchanges at the snow surface by solving the surface energy balance equation (Oke, 2002). ~~These approaches~~The use of the energy balance equation, coupled with snow models, ~~enable~~ enables a more complete understanding of snow physical processes and are essential for understanding the interaction between snow cover evolution and climate change.

Physically-based snow models have different complexities in their physical representations, from a single layer approach (e.g. Strasser and Marke, 2010), to more sophisticated multi-layer detailed models representing the evolution of snow microstructure and the layering of snow physical properties (e.g. Bartelt and Lehning, 2002; Vionnet et al., 2012), leading to a wide variety of snow models with a wide variety of parameterizations. In a snow model intercomparison study, Etchevers et al. (2004) highlighted the importance of parameterization choice, especially regarding the net longwave and albedo characterisation. After comparing 33 snow models, Rutter et al. (2009) concluded that no universal 'best' model exists and model performance ~~strongly~~ mainly depends on the study site. Furthermore, the Earth System Model - Snow Model Intercomparison Project (ESM-SnowMIP) compared several snow models to improve the models in the context of local- and global scale modelling (Krinner et al., 2018) and indicated scientific and human errors in snow model intercomparisons (Menard et al., 2021), but the study sites did not include semi-arid regions.

In addition to the development of new models, many studies have focused on model improvements offering different parameterizations in a single model ~~(e.g. Douville et al., 1995; Dutra et al., 2010)~~(e.g. Douville et al., 1995; Dutra et al., 2010; Essery, 2015). In such frameworks, often many parameters need to be calibrated and are often difficult to be set according to local measurements, such as the albedo and aerodynamic roughness length (Brock et al., 2000, 2006). To address this issue, and to consider and quantify parameter uncertainty propagation in simulated snow depth changes, recent studies have started to use ensemble approaches. Here models are evaluated based on different likely combinations of values of variables such as snow albedo, snow compaction, fresh snow density and liquid water transport (e.g. Essery et al., 2013; Lafaysse et al., 2017; Günther et al., 2019).

In addition, forcing data uncertainty has a significant influence on the simulated snow depth changes (e.g. Magnusson et al., 2015; Raleigh et al., 2015; Günther et al., 2019) and needs to be considered in model evaluations. While point scale simulations forced by direct observations generally reduce forcing uncertainties, measurement errors can be considerable due to the complexity of both measuring certain parameters as well as maintaining measurement sites (e.g. for precipitation (MacDonald and Pomeroy, 2007; Smith, 2007; Wolff et al., 2015), sensor inclination (Weiser et al., 2016) or sensor failure). Methods such as stochastic perturbation with random noise (e.g. Charrois et al., 2016) or following a uniform or normally distributed bias with different magnitudes (e.g. Raleigh et al., 2015) can be used to build an ensemble of meteorological forcing and explicitly simulate the consequence of forcing uncertainty on the simulated snow depth (e.g. Charrois et al., 2016; Zolles et al., 2019; Günther et al., 2019).

Despite past efforts to improve snow models and quantify uncertainty propagation, the uncertainties regarding snow physics representation and meteorological forcing remains (e.g. Essery et al., 2013; Raleigh et al., 2015; Günther et al., 2019); ~~especially~~ in particular in regions where sublimation is the main ablation process, due to the lack of snow modelling studies in semi-arid regions (Gascoin et al., 2013; Réveillet et al., 2020; MacDonell et al., 2013a; Mengual Henríquez, 2017).

This study aims to evaluate two physical snow models with different complexities, considering ~~calibration~~ parameterization and forcing uncertainties. ~~There we~~ We simulate snow depth changes in the semi-arid Andes of Chile using data from an automated weather station. In this ~~semi-arid~~ region, snow model uncertainty is a key concern as snow melt is an essential water resource for the population (Favier et al., 2009). Despite ~~this importance~~ the importance of snow as water resource, quantifying and understanding the snow cover evolution remains limited and challenging due to i) high sublimation rates related to ~~strong~~ high levels of incoming solar radiation, cold air temperatures, arid atmosphere, and high wind speeds (e.g. MacDonell et al., 2013a; Réveillet et al., 2020), and ii) shallow snow depths due to very low precipitation amounts (Scaff et al., 2017; Réveillet et al., 2020; Ayala et al., 2017). In previous studies the effect of wind on snow cover pattern distribution has been assessed by Gascoin et al. (2013) and the relative importance of melt versus sublimation has been studied over one catchment by Réveillet et al. (2020), both making use of the physically-based snow model SnowModel (Liston and Elder, 2006b). The study performed by Mengual Henríquez (2017) assessed the snow types in different Chilean regions with SNOWPACK (Bartelt and Lehning, 2002; Lehning et al., 2002b, a). Nevertheless, an accurate assessment of different snow models' sensitivity to parameterization choice or input forcing is currently missing, although it is expected to have a large impact.

In this work, the sensitivity of SnowModel and SNOWPACK, the common snow models previously used in this region, is assessed in function of parameterization choices and forcing uncertainty. First, the models are calibrated similarly to allow later comparisons and a most ideal setup for both models is designed to acquire a precipitation dataset that corrects for the underestimation of precipitation. Second, both models are run with different combinations of parameterizations to assess the uncertainty of parameterizations~~(Sect. 3.2). Second~~. Subsequently, forcing uncertainty propagation in the snow model is considered by running the models with 1000 sets of perturbed forcing~~(Sect. 3.2.1. Considering model sensitivities to calibration and forcing, models are compared and differences are discussed (Sect. 3.3)~~. The combination of sensitivity analysis to model parameterizations and meteorological forcing allows to evaluate and compare the two models.

2 Study area and data

We assess the sensitivity of both models using data from an AWS over the snow season of 2017. First this study area and the meteorological observations are described, followed by the data preprocessing procedure.

2.1 Study area

The study area is located in the La Laguna catchment, in the Chilean Coquimbo region, close to the Argentinian border (Fig. 1a). To assess the sensitivity of the models to the representation of snow physics and meteorological forcing, we use data from the Tapado Automatic Weather Station (AWS), a permanent ~~station~~ meteorological tower since 2009 located close to the terminus of the Tapado Glacier at 30°S, 69°W, 4306 m a.s.l. (Fig. 1c). The site shows a complex topographic setting with average (maximum) wind speeds of 4.2 m s^{-1} (>15 m s^{-1}) in 2017 and little precipitation (<200 mm a^{-1}) that arrives as snow during fewer than 10 events per year. Precipitation events mainly occur during the winter season (>90%) (Rabatel et al., 2011; Réveillet et al., 2020). Therefore the area surrounding the AWS is only covered with snow in austral winter. At this elevation, ~~no vegetation is observed~~vegetation is extremly sparse.

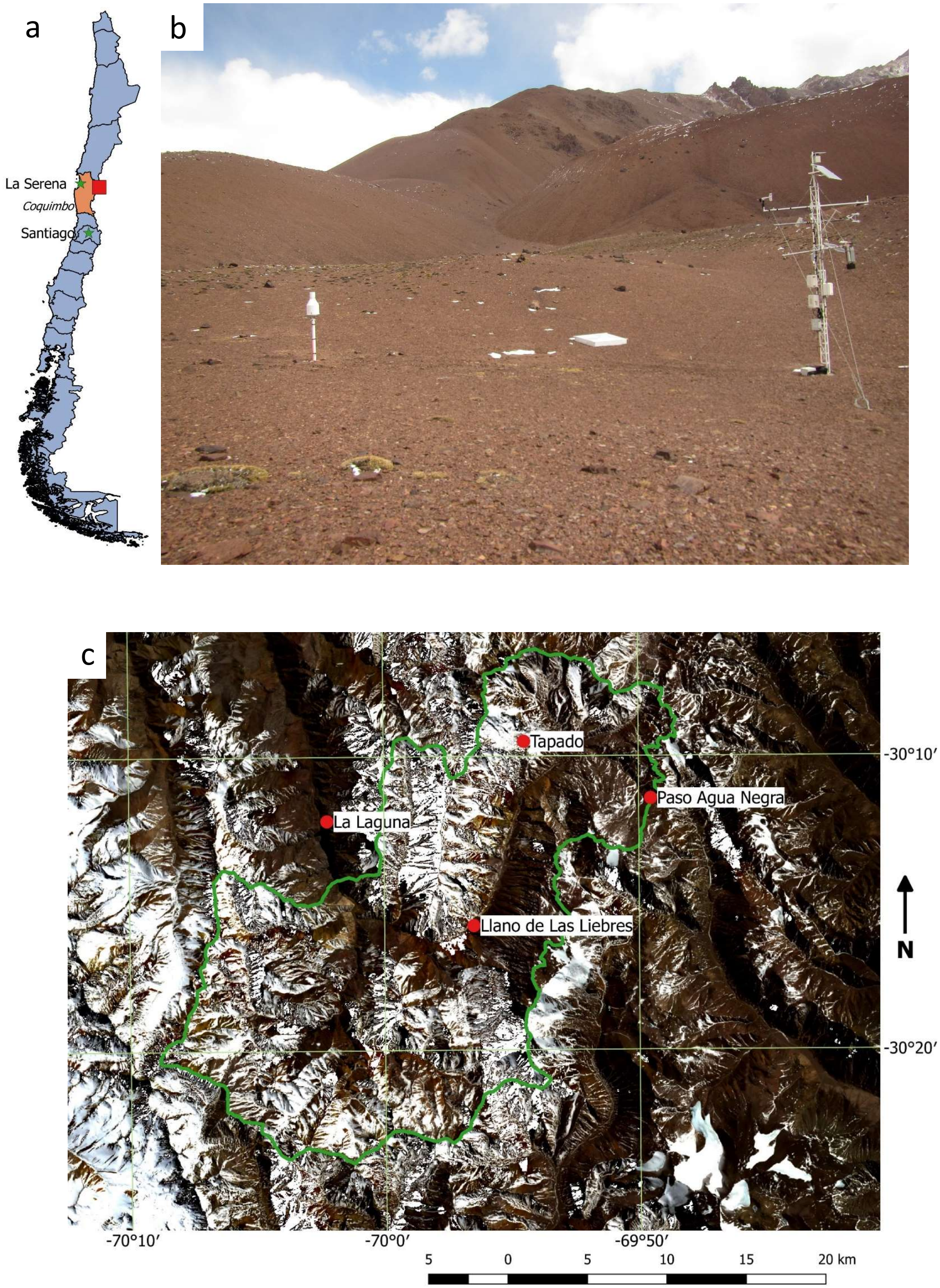

Figure 1. a) Map of Chile with the Coquimbo Region (orange) and the study area location (red box). b) Tapado AWS at the 26th of April 2018 showing the Geonor precipitation gauge (left) which is 10 m from the central mast of the AWS (right). c) Map of the borders of the La Laguna catchment (green), with the AWS locations (red points). Landsat 8 images of 29 August 2017 are used as background and maps and photo are made by A. Voordendag.

2.2 Meteorological observations

The meteorological forcing data consisted of hourly mean values of air temperature (TA), relative humidity (RH), incoming shortwave radiation ($S_{\downarrow}$), incoming longwave radiation ($L_{\downarrow}$), wind speed (WS), wind direction (WD) and air pressure (PA) measured by the AWS (Fig. 2, Table~~S1.1 in the Supplementary Material (SM)~~ 1). Precipitation forcing consisted of hourly data by a Geonor rain gauge (Fig. 1b). This gauge is an unshielted, unheated weighing bucket precipitation gauge filled with anti-freeze liquid and oil to prevent freezing and evaporation respectively. During the snow season, defined as the period with snow on the ground (i.e. between 10 May and 6 November 2017), the station recorded meteorological observations continuously except for the TA and RH, for which gaps have been filled using three nearby AWSs (Fig. 1c, Sect. 2.3).

Hourly snow depth (SD), reflected shortwave radiation ($S_{\uparrow}$) and six-hourly means of snow water equivalent (SWE) were also recorded at the station and used for model calibration and evaluation. SWE was measured with a CS725 sensor by Campbell Scientific which passively detects the change in naturally emitted terrestrial gamma radiation from the ground after it passes through snow cover. It provided two independent SWE observations measuring both potassium and thallium gamma rays (Wright, 2011). The uncertainty given by the manufacturer is $\pm$15 mm from 0 to 300 mm and $\pm$15% from 300 to 600 mm, but differences of up to 82 mm w.e. between gamma ray measurements at $\sim$300 mm w.e. were measured. The manufacturer suggests that the output with the higher count is generally the most reliable, which were the potassium gamma rays measurements (Randall, 2018, personal communication). We display both data sets and estimate an accuracy of $\pm$25 mm for this data set.

2.3 Preprocessing of forcing data

~~When modelling the snow evolution, we covered a period (~~The period between 5 May ~~to~~ and 30 ~~Nov~~ November 2017 ~~) extending outside the snow season to allow a five day spin-up phase to enable the model to reach a physical equilibrium with the applied forcing~~has been covered to model the snow evolution in the austral winter, as this was a season where SWE data was available to validate the models. Since continuous data are required for both snow models, preprocessing was necessary to fill the gaps in the TA and RH data sets (23 June 11:00 ~~and~~ to 31 October 10:00 due to sensor failure) and to correct the wind-induced undercatch in the precipitation data. Therefore, TA and RH data was interpolated based on lapse rates from nearby AWSs (Agua Negra (4774 m a.s.l.), Llano de las Liebres (3565 m a.s.l.) and La Laguna (3209 m a.s.l.; Fig. 1c).

For TA, a daily ~~moist adiabatic lapse rate~~ temperature lapse rate (Blandford et al., 2008) was calculated using TA measured at La Laguna and Paso Agua Negra AWSs (1565 m elevation difference) between 2014 and 2017. We fitted a sinusoidal trend over these lapse rates for the ~~4-year~~ four-year period and found daily lapse rates with a maximum of -6.9°C km^{-1} in winter and a minimum of -8.0°C km^{-1} in summer. These daily lapse rates were subsequently applied to TA observations of Llano de las Liebres AWS which is the only AWS that covers the entire period of missing data in 2017. For RH a similar approach was applied using the lapse rate of the daily dew point temperature between the Paso Agua Negra and La Laguna AWSs and applying it to data measured at the Llano de las Liebres AWS. Dew point temperature was converted to RH following Liston

Table 1. Available observations, sensor height from the ground and the manufacturers and type of the corresponding sensor at Tapado AWS.

Measurement	Unit	Height (m)	Brand/type	Uncertainty given by manufacturer
Accumulated precipitation	mm	1.5	Geonor/T-200B 1000mm	0.1% Full Scale
Air pressure (Pa)	hPa	3.5	Vaisala/PTB110	±1.0hPa
Air temperature (TA)	°C	3.5	Vaisala/HMP45C	±0.3°C at 0°C
Incoming LW radiation ($L_{\downarrow}$)	W m^{-2}	3.5	Kipp and Zonen/CNR4	10% (95% confidence level)
Incoming SW radiation ($S_{\downarrow}$)	W m^{-2}	3.5	Kipp and Zonen/CNR4	5% (95% confidence level)
Outgoing LW radiation ($L_{\uparrow}$)	W m^{-2}	3.5	Kipp and Zonen/CNR4	10% (95% confidence level)
Reflected SW radiation ($S_{\uparrow}$)	W m^{-2}	3.5	Kipp and Zonen/CNR4	5% (95% confidence level)
Relative Humidity (RH)	%	3.5	Vaisala/HMP45C	±2% RH (0 to 90% RH) ±3% RH (90% to 100% RH) and ±0.05% RH/°C
Wind speed (WS)	m s^{-1}	5.4	RM Young/5103	±0.3 m/s
Wind direction (WD)	°	5.4	RM Young/5103	±3°
Snow depth (SD)	m	3.5	Campbell/SR50A	±1 cm
Water equivalent (SWE, thallium, Tl)	mm	3.5	Campbell/CS725	±15 mm from 0 to 300 mm ±15% from 300 to 600 mm
Water equivalent (SWE, potassium, K)	mm	3.5	Campbell/CS725	±15 mm from 0 to 300 mm ±15% from 300 to 600 mm

and Elder (2006a). Evaluation of this lapse rate interpolation, based on 1638 overlapping observations at Tapado, shows an uncertainty (i.e. RMSE) of 2.8°C and 9.97% for TA and RH respectively.

Since the precipitation observations were ~~strongly~~ directly influenced by wind, an undercatch in the precipitation gauge is likely ~~(e.g., MacDonald and Pomeroy, 2007; Smith, 2007; Wolff et al., 2015). Therefore the Geonor precipitation observations~~

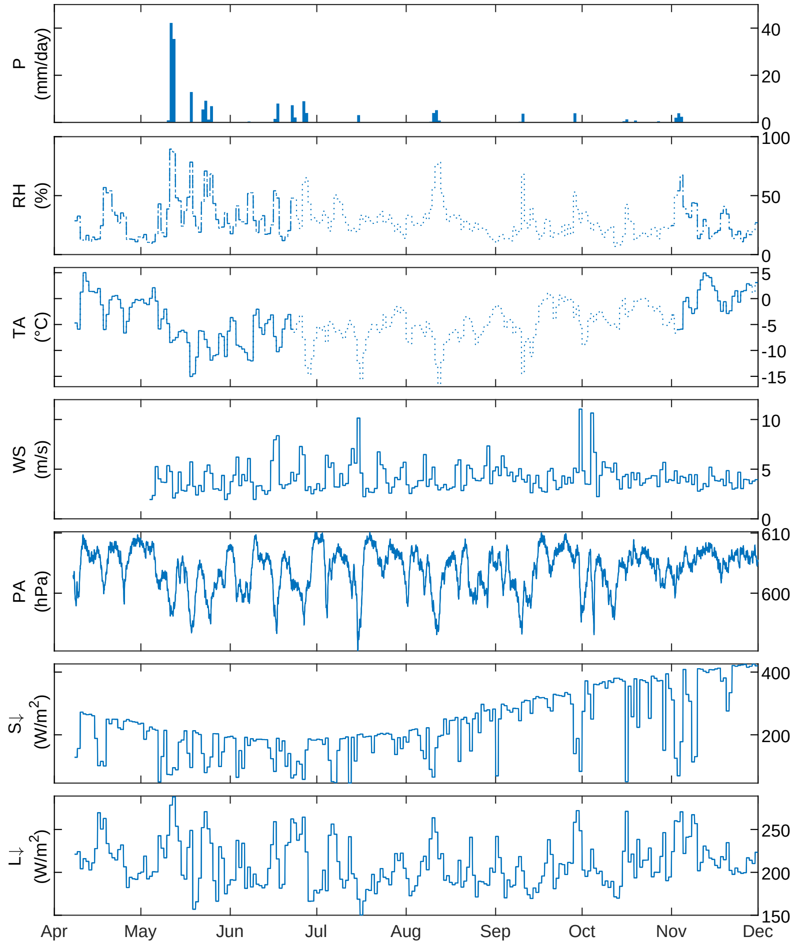

Figure 2. Meteorological observations at Tapado with with daily precipitation (P), average daily relative humidity (RH), temperature (TA), wind speed (WS), air pressure (PA), incoming shortwave radiation ($S_{\downarrow}$) and incoming longwave radiation ($L_{\downarrow}$) from April to December 2017. Dotted lines indicate the TA and RH interpolations.

~~were corrected by applying the catch efficiency factor CE using MacDonald and Pomeroy (2007, Eq. (3)) based on WS observations adjusted from the wind sensor height (5.4 m) to gauge height using a logarithmic wind profile (e.g., Lehning et al., 2002a) and a roughness length of 0.001 m:~~

$$\underline{CE_{Geonor} = 1.01 \exp(-0.09\, WS)}$$

~~This corrected precipitation is hereinafter referred to as reference precipitation (P_{ref}). Precipitation uncertainty was quantified as the difference between the uncorrected, observed precipitation and the precipitation reconstructed from the SWE observations, which was computed using the cumulative positive SWE (potassium) changes during precipitation events (detected by the Geonor T-200B). This cumulative SWE approach reduced the inclusion of deposition of drifting snow resulting which would result in an overestimation of SWE. Based on thisapproach, the precipitation ranges between 190 mm a^{-1} (from precipitation gauge) and 470 mm a^{-1} (from SWE reconstruction). The difference between these values is likely caused by the location offset between the precipitation gauge and the SWE and SD sensor (Fig. 1b), as the AWS is placed in a concave area that collects more snow than the Geonor precipitation gauge~~(e.g. MacDonald and Pomeroy, 2007; Smith, 2007; Wolff et al., 2015). As there are different possibilities to correct for this, the assimilation and correction of precipitation data is explained in Sect. 3.2.2.

3 Methods

3.1 Model descriptions

3.1.1 SNOWPACK

SNOWPACK was developed by the Swiss Federal Institute (SLF) for Snow and Avalanche Research (Bartelt and Lehning, 2002; Lehning et al., 2002a, b). It is a one-dimensional model, but can be implemented in the spatially distributed, three-dimensional snow cover and earth surface model *Alpine3D* (Lehning et al., 2006). SNOWPACK includes a *MeteoIO* preprocessing library for meteorological data (Bavay and Egger, 2014) which was not used, as we implemented a homogeneous preprocessing approach for both models (see Sect. 2.3). SNOWPACK is a physically-based model which has the ability to simulate snow physical properties (e.g. snowpack temperature, layer thickness, snow microstructure and density) and snow processes (e.g. refreezing, sublimation, melt, evaporation) for multiple layers, which are merged if layers become too thin. Sublimation and evaporation are calculated for the top element of the snowpack and melt is simulated using a water transport bucket scheme. In this bucket scheme, all the liquid water exceeding a threshold water content is transported downward in the snowpack or soil (Wever et al., 2014). An extensive description of the model can be found in Bartelt and Lehning (2002); Lehning et al. (2002a, b).

3.1.2 SnowModel

SnowModel is a spatially distributed snowpack evolution modelling system composed of four submodels *MicroMet*, *EnBal*, *SnowPack* and *SnowTran3D* (Liston and Elder, 2006b). MicroMet is preprocessing library for meteorological data interpolation, which was not used in this study as we focused on one location only while we implemented a homogeneous preprocessing approach for both models (see Sect. 2.3). EnBal calculates standard surface energy balance exchanges (Liston and Hall, 1995). SnowModel's SnowPack subroutine is a single or multi-layer (max. six layers) snowpack evolution and runoff model that describes snowpack changes in response to precipitation and melt fluxes defined by MicroMet and EnBal (Liston and Hall, 1995; Liston and Elder, 2006b). In SnowModel, the melted snow is redistributed through the new snow depth up to a maximum density threshold of 550 kg m^{-3}. Any additional melt water is added to the runoff. In this study the model was run with the maximum of snow layers (i.e. six layers) to be comparable with the multiple amount of layers in SNOWPACK. Finally, the three-dimensional model to simulate snow erosion and deposition SnowTran3D (Liston and Sturm, 1998) is not activated in this study; this choice is discussed in Sect. 5.2.

3.2 Model ~~calibration~~setup and sensitivity analysis

To assess the sensitivity of both models to parameterization choice and input uncertainty, we applied a ~~three~~ four step approach. First, we calibrated both models similarly to allow later comparisons (Sect. 3.2.1). Second, we designed the most ideal setup for both models to acquire a precipitation dataset that corrects for the underestimation of precipitation. Third, we varied the parameterization settings of each model to determine the effect of parameterization choice (Sect. 3.2.3). ~~Third~~Last, we implemented forcing ~~perturbations~~ biases (Sect. 3.2.1) to evaluate ~~model sensitivities~~ the sensitivity of the models to the meteorological forcing uncertainties. The combination of sensitivity analysis to model ~~calibration~~ parameterizations and meteorological forcing allowed to evaluate and compare the two models (Sect. 3.3).

3.2.1 Parameter values used in both models

Initially, both models were calibrated using similar parameters to facilitate intercomparison. These parameters were derived from observations or previous studies (Table 2). For example, the soil albedo was set to 0.15, as this is the average observed albedo at the moment when there is no snow. The observed daily albedo is defined as the daily sum of the average hourly ~~reflected shortwave ($S_\uparrow$)~~ divided by the daily sum of average hourly ~~incoming shortwave radiation ($S_\downarrow$)~~(Fig. 4e,f). In the absence of roughness length measurements, the roughness length of the bare soil is set to 0.020 m, corresponding to the default roughness length of pebbles and rocks in SnowModel. As surface ground temperature measurements are not available, we set it to -1°C in both models. -1°C is the default value in SnowModel and ensures that the first snow does not immediately melt.

~~Atmospheric stability and roughness length (z_0) are key parameters in semi-arid regions where sublimation is an important process. As SnowModel only allows atmospheric stability corrections based on the Richardson number, we opted for this method and similar roughness lengths in both models to assure intercomparability. The snow roughness length was fixed to 0.001 m based on an earlier sensitivity study (Réveillet et al., 2020) and unpublished eddy covariance measurements (MacDonell et al., 201~~ ~~Additionally, the sensitivity of both models to roughness length is further analysed in Sect. S3 of the SM, as it was not used to select the final model configurations.~~

3.2.2 Idealised setup

Preliminary results showed simulated SWE and SD to be more than two times lower than the observed SWE. This is caused by an underestimation of the precipitation measurements, as the AWS is placed in a concave area that collects more snow than the Geonor precipitation gauge. This is in correspondence with research by Grünewald and Lehning (2014) on the spatial variability of SD measurements. Therefore, to determine the optimal precipitation input for the models, an idealised setup is

Table 2. SNOWPACK and SnowModel parameter characteristics. The possible snow albedo parameterizations and fresh snow density models are described in Sect. ~~S4~~S1.

	SNOWPACK	SnowModel
Soil albedo	0.15 (calibrated)	0.15 (calibrated)
Max/min snow albedo	None	0.6/0.9 (calibrated)
Atmospheric stability correction model	Richardson number	Richardson number (default)
Roughness length (soil)	0.02 m	0.02 m (default)
Roughness length (snow)	0.001~~m (calibrated)~~ m and 0.01 m	0.001 m ~~(calibrated)~~ and 0.01 m
Surface ground temperature	-1°C	-1°C (default)
Thermal conductivity	Default	Multilayer subroutine
Wind erosion/snow transport by wind	Off	Off
Maximum number of snow layers	Unlimited	6 layers
Fresh snow density parameterizations	5 options	1 default options and 2 from SNOWPACK
Albedo parameterizations	6 options, 4 used	2 options
Simulated ablation processes	Sublimation, runoff, evaporation	Sublimation, runoff
Water transport in snowpack	Bucket scheme (default)	Default

designed, making use of all the data that the models allow as forcing. Two approaches are designed to adjust the precipitation data set. First, it is chosen to assimilate a precipitation data set, which both models perform in different ways. SNOWPACK assimilates the data if SD is given as input. Reflected SW radiation is also given as input, to prevent inaccurate parameterizations of the albedo. The precipitation data set is assimilated with the five possible fresh snow density parameterizations in SNOWPACK. SnowModel allows the possiblity to assimilate the precipitation when SWE is given, but is not able to cope with reflected SW radiation as input. Therefore, six ensembles are made out of two albedo and three fresh snow density parameterizations to find an assimilated precipitation data set. Second, the precipitation is reconstructed from the SWE observations, which was computed using the cumulative positive SWE (potassium) changes during precipitation events (detected by the Geonor T-200B). This cumulative SWE approach reduced the inclusion of deposition of drifting snow resulting which would result in an overestimation of SWE. Hereinafter this precipitation dataset is called PSWE.

Atmospheric stability and snow roughness length (z_0) are key parameters in semi-arid regions where sublimation is an important process. As SnowModel only allows atmospheric stability corrections based on the Richardson number, we opted for this method and similar roughness lengths in both models to assure intercomparability. The z_0 was set to both 0.001 m and 0.01 m. 0.001 m is based on an earlier sensitivity study (Réveillet et al., 2020) and unpublished eddy covariance measurements (MacDonell et al., 2013a); 0.01 m is based on literature (e.g Brock et al., 2006; Cuffey and Paterson, 2010). The first and second approach of the idealised setup are both tested with z_0 of 0.001 m and 0.01 mm, thus the idealised setup in total consists of four cases of each five (SNOWPACK) or six (SnowModel) simulations. The simulated SWE and SD are compared to the observed

SWE and SD and the assimilated precipitation data sets are shown. The P that leads to a best correspondence to the observed SWE and SD is used in the further study.

3.2.3 Sensitivity analysis of variable parameterizations

To assess the impact of the parameterizations on the snowpack simulation, an ensemble approach based on different combinations of albedo and snow density parameterizations ~~was used (e.g. Essery et al., 2013; Lafaysse et al., 2017)~~and z_0 was used (e.g. Essery et al., 2013; Lafaysse et al., 2017, and Sect. 3.2.2). The choice to limit the sensitivity test to these ~~two~~ three parameters is discussed in Sect. 5.2.

For SNOWPACK, ~~20~~ 40 runs were performed over the 2017 season based on four different albedo~~and~~ , five fresh snow density parameterizations and two different z_0 values. Each of the albedo parameterizations is based on empirical relations derived from continuous observations at Weissfluhjoch (Lehning et al., 2002a) or on grain size (Schmucki et al., 2014), while the fresh snow parameterizations are empirical formulas depending on the TA, RH, WS and surface temperature. More details are found in the Sect. ~~S4~~ S1 (Supplementary Material (SM)) and the mentioned references.

For SnowModel, an ensemble of ~~six~~ 12 simulations was run, considering two albedo~~and~~ , three snow density parameterizations and two different z_0. The albedo parameterizations range between 0.6 and 0.9 depending i) on TA solely (more details in Liston and Hall (1995), Liston and Elder (2006b) and in Sect. ~~S4~~ S1 (SM)) or ii) on TA and time ~~(Strack et al., 2004, and Sect. S4)~~(Strack et al., 20 SnowModel's default fresh snow density parameterization depends on the wet bulb temperature, but we included two fresh snow density parameterizations from SNOWPACK depending on TA, RH, WS and surface temperature to test the model more extensively. In these additional parameterizations, we preserved the SnowModel defaults for minimum (50 kg m^{-3}) and maximum fresh snow density (158.5 kg m^{-3}).

Each of the ensemble simulations was forced by the ~~unperturbed measurements (P_{ref},~~ observations (TA, RH, PA, WS, WD, $S_{\downarrow}$, $L_{\downarrow}$) as described in Sect. 2.3 and the P data acquired after the idealised setup (Sect. 3.2.2). The simulations are evaluated by comparing the model output of SD, SWE and albedo with the corresponding observations. Based on this evaluation the simulation with the lowest $RMSE$ and highest R^2 between the observed and modelled albedo is chosen as the reference for the forcing sensitivity analysis discussed in Sect. 3.2.1.

3.3 ~~Forcing uncertainty estimation~~

3.2.1 Forcing uncertainty estimation

To assess the model sensitivity to meteorological measurement uncertainties, a bias has been applied to the meteorological forcing presented in Sect. 2.3 ~~is perturbed~~ to generate an ensemble of 1000 forcing files. Raleigh et al. (2015) have shown that the model outputs are more sensitive to forcing biases than random errors. Therefore, all input variables except P were modified by adding hourly ~~perturbations~~ biases with a normal distribution $N(\mu = 0, \sigma^2)$ with σ the uncertainty range taken from Raleigh et al. (2015) and reported in Table 3. ~~The precipitation perturbations consisted of adding random variations $N(\mu = 0, \sigma^2)$ to the reference precipitation ($P_{ref,tot}$: 278 mm w.e. a^{-1}) based on the mean $P_{mean,t}$ between the reconstructed~~

Table 3. Forcing data for the snow models with the corresponding uncertainty σ used in the ~~MC~~ ensemble simulation. The ~~uncertainties~~ ranges of PA, ~~WS, WD~~TA, $S_{\downarrow}$, ~~and~~ $L_{\downarrow}$, RH and WS are ~~the uncertainties~~ ranges as used by Raleigh et al. (2015). The WD range is according to the uncertainty given by the ~~manufacturers and for TA~~ manufacturer and ~~RH,~~ the ~~uncertainties as~~ P range is described in Sect. ~~2.3 are used~~3.2.1.

~~Observation~~ Forcing	~~Uncertainty σ in MC simulation~~Distribution	Range	Unit
Accumulated precipitation (P)	~~Between 190 and 470~~ Uniform	[-100,+100]	mm a^{-1}~~, Fig. S2.1~~
Air pressure (PA)	~~±1.0hPa~~Normal	[-100,+100]	Pa
Air temperature (TA)	~~±2.8~~Normal	[-3.0,+3.0]	°C
Incoming longwave radiation ($L_{\downarrow}$)	~~10% (95% confidence level)~~Normal	[-25,+25]	W m^{-2}
Incoming shortwave radiation ($S_{\downarrow}$)	~~5% (95% confidence level)~~Normal	[-100,+100]	W m^{-2}
Relative humidity (RH)	~~±9.97%RH~~Normal	[-0.25,+0.25]	%
Wind speed (WS)	~~±0.3 m /s~~Normal	[-3,+3]	m s^{-1}
Wind direction (WD)	~~±~~Normal	[-3,+3]	°

~~and measured hourly precipitation rate at every time step),~~ The biases has been kept within their corresponding range (Table 3) by assuming that the 99.7% of the bias, thus 3σ~~is considered~~ , is within this range. This positive component of the range is divided by three and multiplied with a normally distributed random number ~~, used as sealing factor and kept constant throughout the time series. Based on this sealing factor, the reference precipitation for each of the~~ and added to the observed forcing. We have chosen 1000 ~~simulations ($P_{sim,t}$) at time t is modified using:~~

$$P_{sim,t} = P_{ref,t} + \sigma P_{mean,t}$$

~~The precipitation distribution results are shown in Fig. S2.1a where the total precipitation~~ runs as a compromise between computational effort and a reliable confidence interval.
The distribution of the precipitation uncertainty is chosen to be uniform, as the observed precipitation was low (i.e. 180.7 mm w.e. at the end of the season~~is normally distributed around the reference precipitation (Fig. S2.1b) .~~
) and the results in Sect. 4.1 between the assimilated precipitation (SnowModel) and PSWE cover approx. 200 mm w.e..

Subsequently, based on the disturbed input data, 2000 snow model simulations are performed: 1000 with meteorological ~~perturbations~~ biases and 1000 with combined meteorological/precipitation ~~perturbations~~biases. This setup was chosen to enable the differentiation between meteorological and precipitation uncertainties, which would be difficult in a combined approach where precipitation uncertainty would dominate.

3.3 Model evaluation

Model evaluation consists of comparing the model output of SD, SWE and albedo with the corresponding observations. For the idealised case this consists of evaluating the $RMSE$ and R^2 between the modelled and the observed SD to acquire precipitation data that approaches the correct mass balance. For the parameterization uncertainty, this consists of evaluating the $RMSE$ and

R^2 between the modelled and the observed albedo, to select the best reference for each model (i.e. ~~twenty~~ 40 for SNOWPACK and ~~six~~ 12 for SnowModel). ~~It~~ In this case, it is chosen to only compare between modelled and observed albedo, as this ensures the best possible net shortwave radiation term in the energy balance equation. The forcing uncertainty is evaluated by comparing the differences of end of snow season. Last, the differences in ablation processes of the parameterizations are shown~~along with the energy fluxes of each reference simulation~~.

4 Results

4.1 Idealised simulations

The assimilated precipitation datasets markedly differ between SNOWPACK and SnowModel (blue lines, Fig. 3e,f). For clarity Fig. 3 shows only the results of the idealised simulations for the z_0 value of 1 cm; the results for both z_0 is 1 cm and 1 mm is displayed in Sect. S2. For SNOWPACK, eight out of ten runs with z_0=1 cm crashed, thus only two simulations are shown.

Assimilation of SD in SNOWPACK results in SWE that approximates the PSWE (Fig. 3), leading to assimilated precipitation amounts of 2.55 to 3.02 times the observed precipitation and a good correspondence with the observed SD (i.e. $RMSE$ between 9.2 and 11.5 cm and R^2 between 0.90 and 0.93 calculated with the observed and simulated SD, Fig. 3a).

Assimilation of SWE in SnowModel only adjusts the precipitation between 22 and 27 June and between 7 and 12 August. The amount of precipitation is not adjusted at the beginning of the season and thus, the assimilated data by SnowModel still leads to an underestimation of the SD and SWE (Fig. 3b,d). The missing adjustment of the SWE is probably caused by SnowModel taking a maximum of 99 SWE observations and the observations do not exactly allign with the precipitation events, which leads to correction factors of one (i.e. no change) to the precipitation data. The assimilated precipitation is approx. 1.6 times the observed precipitation and the correspondence of the observed with the SD is lower than with SNOWPACK (i.e. $RMSE$ between 7.1 and 17.1 cm and R^2 between 0.19 and 0.90 calculated with the observed and simulated SD, Fig. 3b).

Both models overestimate the SWE between mid-July and September when PSWE was given as input (red lines in Fig. 3c,d). This is likely caused by an overestimation of the PSWE at the end of June. Only small amounts of precipitation are observed at the precipitation gauge, but the observed SWE distinctly increases probably due to snow drift, as strong winds were also observed. A similar thing occurs at the end of September. The models markedly increase the amount of precipitation (between 97 and 137 mm w.e.) in the assimilation runs, but only small amounts of precipitation and strong winds were observed. This overestimation and the need for SWE as validation data are indications that the PSWE is not a valid precipitation dataset for our simulations , but it is also unfeasible to select one of the assimilated precipitation sets by SNOWPACK as the amount of precipitation is markedly increased at the end of September and we want to use SD as validation data.

Therefore, three different precipitation corrections depending on WS (Smith, 2007; MacDonald and Pomeroy, 2007) or on TA and WS (Wolff et al., 2015) were applied to the observed precipitation (See Sect. S3). Eq. (12) from Wolff et al. (2015) with WS corrected to gauge height using a logarithmic wind profile (e.g. Lehning et al., 2002a) and a z_0 of 0.01 m is used as precipitation data in the further study, as this correction approaches the PSWE and shows an increase in precipitation of 2.35 times the observed precipitation at the end of the season.

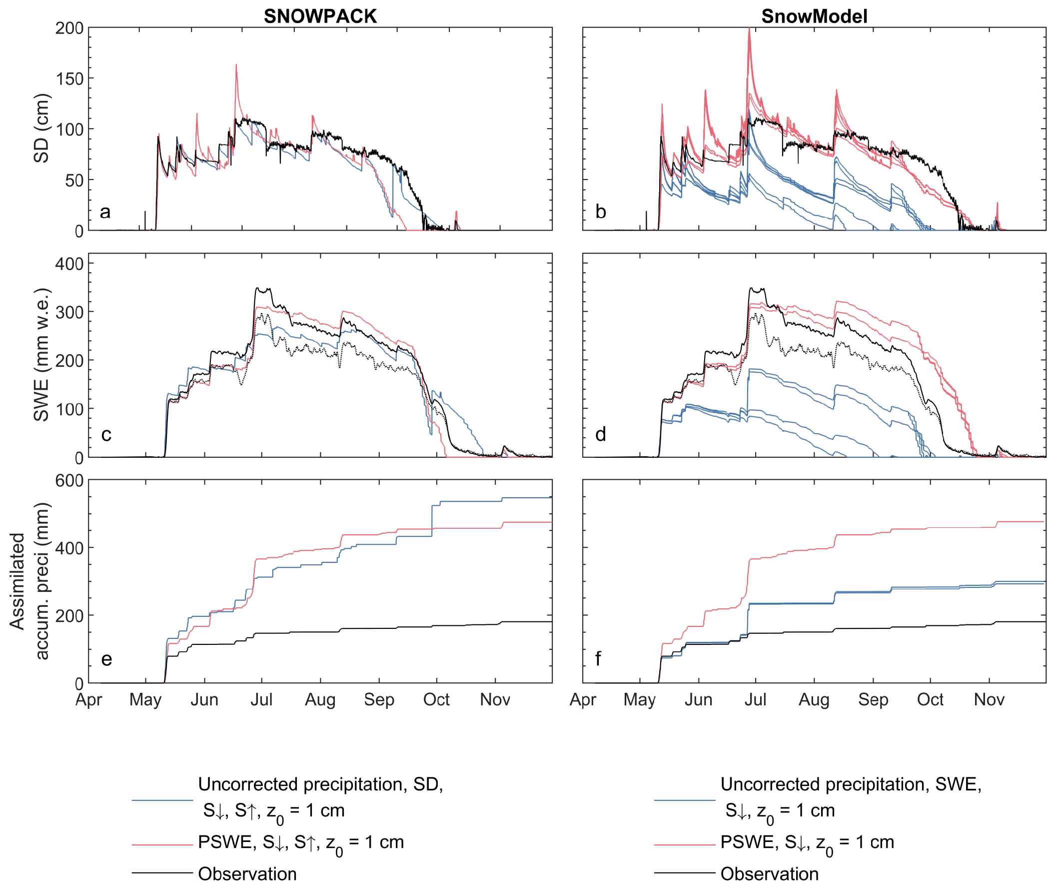

Figure 3. a-b) SD, c-d) SWE and e-f) the cumulative assimilated precipitation for SNOWPACK (a,c,e) and SnowModel (b,d,f) and observations (black). The different forcing parameters are given in the legenda. The dotted line of SWE indicates the less reliable (lower) SWE measurement from thallium rays (See Sect. 2.3). The simulations for z_0 = 1 mm are found in Sect. S2.

4.2 Sensitivity analysis of parameterizations

The ~~SW~~ SD and SWE evaluation of the parameterization sensitivity analysis ~~shows~~ show that the simulated SD and SWE are ~~lower than~~ in good correspondence with the observations for both models (Fig. 4)~~. This can be explained by the precipitation underestimation recorded by the precipitation gauge despite correcting for wind~~, but also overestimate the SWE at the beginning of the season. The correction of the precipitation with the equation from Wolff et al. (2015) overestimates the precipitation in this period, and also leads to an overestimation in the simulations.

For SNOWPACK, the spread of the simulated SD from the ~~20 different calibrations~~ 40 different parameterizations (20 simulations for z_0 = 1 mm and 20 for z_0 = 1 cm) is the largest at the ~~beginning (i.e. May, June) and at the~~ end of the snow season (i.e. October) (Fig. 4a). The date of snow free surface ~~ranges between~~ is the earliest at 8 October and ~~19 November~~ exceeds the simulated period (i.e. ~~41 days of differences~~after 30 November), depending on parameterization choice, and covers the observed date of snow removal (i.e. 16 October). The different SNOWPACK parameterizations (equations in Sect. S1) show a mean SD difference of ~~10~~ 32 cm (which corresponds to ~~9~~28.9% of the total SD) between the minimum and maximum simulated SD (Fig. 4a), with a maximum of ~~84~~ 127 cm observed at ~~12 May, the first day of snowfall~~27 June. For the SWE, this corresponds to a mean difference of ~~23.9~~ 98.3 mm w.e. (i.e ~~6.9~~28.2% of the total SWE) (Fig. 4c). The large modelled SD spread in May and June can be explained by the different density parameterization choices as it is not apparent in the SWE simulations (Fig. 4g). The rapid decrease (3-8 cm d^{-1}) of snow depth until July, caused by compaction of the snowpack, is simulated by the majority of fresh snow density parameterizations, while only one fresh density parameterization models a more moderate compaction (Fig. 4a,g). From July onward, the measured snow depth reduces 10 centimetres per 25 days, which is only simulated by the fresh density parameterization that simulated moderate compaction before July. Snow density measurements were unavailable in 2017 and the observed snow density in Fig. 4g is calculated with SWE/SD. The observed snow density is only shown until the end of August, as the calculation led to unrealistic decreasing snow densities after August, likely caused by higher readings at the SD sensor than at the SWE sensor, as those sensors were placed on different sides of the meteorological tower or to a bias of the SWE sensor in the ablation season as explained by Smith et al. (2017).

The albedo evaluation (Fig. 4e) and corresponding statistics (Sect. ~~S5~~S4) highlight one parameterization choice that outperforms all other parameterizations (i.e. $RMSE$ of 0.09 (-) and R^2 of ~~0.85 compared to the~~ 0.86 calculated with the observed and simulated albedo) in terms of snow compaction after snowfall events, end of snow season, and albedo evolution (Fig. 4e). Therefore ~~it~~ this simulation with a z_0 of 1 cm is selected as the reference simulation (represented in bold lines in Fig. 4) for the forcing uncertainty simulations.

For SnowModel, the largest SD spread of the ~~six ensembles~~ 12 ensembles (six for every z_0, equations in Sect. S1) occurs at the end of the simulated snow season (i.e. August~~and September~~, September and October) with complete snow removal between ~~19 and 27 September~~ 21 October and 12 November (i.e. ~~8~~ 22 days) (Fig. 4b). The mean SD difference between the parameterizations is ~~11~~ 20 cm (i.e. ~~10~~18% of the total SD), with a maximum of ~~108~~ 152 cm at 12 May, the first snow fall, while for SWE the mean difference is ~~18~~ 57 mm (i.e. ~~5.2~~19.2% of the total SWE) with a maximum of ~~187 mm at 27 and 28 June~~ 317 mm w.e. at 12 August (Fig. 4d).

~~A closer analysis of Fig. 4f shows that SnowModel's output clusters into two groups, where the grouping is determined by the albedo parameterization with limited influence of fresh snow density parameterizations. The differences between the two clusters increases as the difference in albedo between the parameterizations increases at the end of May.~~ Quantitative analysis (Sect. ~~S5~~S4) shows best performance scores for the time-evolution albedo approach in combination with the reference snow density parameterization and a z_0 of 1 cm ($RMSE$ of ~~0.123~~ 0.150 (-) and R^2 of ~~0.776~~0.600). Therefore, these are used as the reference simulation (bold line in Fig. 4).

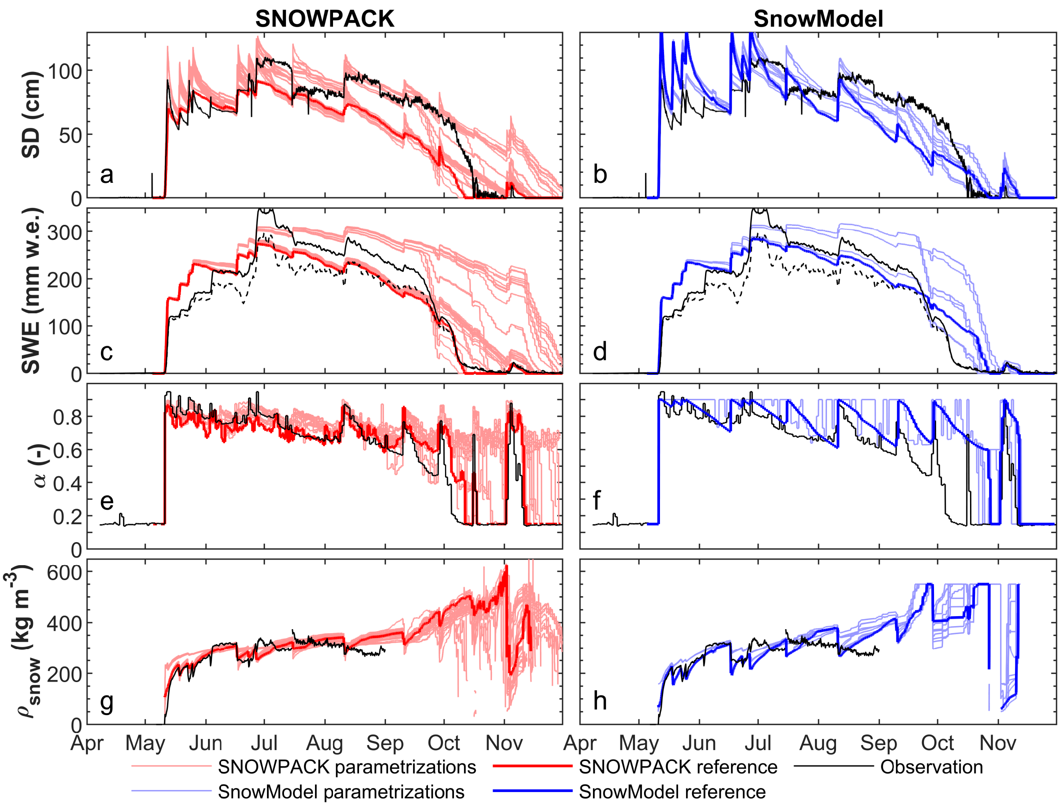

Figure 4. a-b) SD, c-d) SWE~~and~~ , e-f) albedo and g-h) snow density simulations (coloured) of the ensemble approaches for SNOWPACK (red) and SnowModel (blue) and observations (black). The bold coloured lines show the reference simulations chosen as the most optimal ~~calibration~~ parameterizaton set according to the measured albedo.~~The dotted line of SWE indicates the less reliable (lower) SWE measurement from thallium rays (See Sect. 2.3).~~

Comparison of the SNOWPACK and SnowModel output shows similar SD variations attributed to snow density parameterizations that simulate low density snowfall with ~~strong subsequent compaction~~ notable subsequent compaction (Fig. 3g,h). In reality, this happens at Tapado until June, followed by a different regime with denser fresh snow and less compaction. The biggest difference between the models, however, is the result of the albedo parameterizations. Where SnowModel relies on two ~~simplistic~~ albedo models based on TA and albedo ranges, SNOWPACK relies on empirical relations ~~which are~~ calibrated with measurements in Switserland and not adapted to the arid Tapado climate. Nevertheless, the albedo of the reference run of SNOWPACK performs well in a semi-arid area. Last, the simulations by SnowModel all approximate the end of snow season within a period of 22 days, whereas the simulations at the end of season noticeably diverge for SNOWPACK. This is likely caused by TA above the freezing point at the end of October and a fast melt simulated for all ensembles by SnowModel.

4.3 Sensitivity analysis of forcing data

4.3.1 Without precipitation uncertainties

The ~~forcing perturbations~~ biased forcing without precipitation uncertainty show ~~that SNOWPACK is more sensitive than~~ a similar sensitivity for SNOWPACK and SnowModel (Figures 5a,c) with mean SD/SWE differences of ~~7.9~~ 52 cm/~~17.5~~ 163 mm w.e. for SNOWPACK and ~~5.7~~ 47 cm/~~7.7~~ 172 mm w.e. for SnowModel. ~~Also temporally, SNOWPACK shows larger sensitivities with larger spread uncertainty throughout the season, whereas for SnowModel the spread only emerges after mid-June~~The simulations with SnowModel show more uncertainty in the melting period (e.g. in October), but otherwise, the simulations mainly overlap. The forcing uncertainty results in complete snow removal simulations ranging from ~~25 Sep - 15 Oct~~ 27 August to 28 November (i.e. ~~19~~ 93 days) for SNOWPACK and ~~17-20 September~~ 30 August to 25 November (i.e. ~~3~~ 87 days) for SnowModel. The ~~larger difference between both models in terms of complete snow removal is probably caused by precipitation input variations in SNOWPACK as result of TA, RH and~~ ~~$S_{\downarrow}$~~ ~~constraints on snowfall in SNOWPACK. Since these requirements are not always met when the forcing is perturbed, it may result in precipitation variations in the input~~reference simulations of both models are located in the middle of the spread of simulations, which is coherent with the normal distribution of the biases applied to the forcing. The biggest differences between the models are found in the way SD has been simulated. The reference run and the 1000 simulations with biased forcing show marked settling rates througout the season with SnowModel, whereas the settling is more moderate for SNOWPACK. This depends on the chosen snow density parameterization and is discussed further in Sect. 5.

4.3.2 With precipitation uncertainties

The forcing perturbations with precipitation uncertainty shows that precipitation uncertainty has a ~~much larger~~ large impact on SD and SWE ensemble spread (Fig. 5b,d). Averaged over the season this results in SD/SWE differences of ~~18.4~~ 75 cm/~~52.5~~ 257 mm w.e. and ~~24.9~~ 70 cm/~~63.7~~ 262 mm w.e. for SNOWPACK and SnowModel, respectively. ~~Despite~~ Along with the similar average spread over the entire season observed for both models, the range of the simulated day of free snow surface is ~~higher for SNOWPACK (i.e. 24 September and~~ also similar; for SNOWPACK this date is between 29 August and later than 30 ~~October) than for SnowModel~~ November and for SnowModel this is between 20 August and 29 November (i.e. ~~between 16 September and 6 October). Indeed, the spread at the end of the ablation season is diminishing for SnowModel while it is increasing for SNOWPACK~~101 days). Again, the main differences are found in the settling of the snowpack (See Sect. 5).

4.4 Consequences of the model choice and ~~calibration~~ parameterizations on sublimation

Ablation rates (Fig. 6) show that sublimation is the dominant ~~driver for~~ mode of mass loss in both models until September (i.e. cold period), and followed by melt from September to the end of the season (i.e. end of November, and called the melting period). Note that for SNOWPACK, the first day of ~~free snow~~ snow-free surface of the reference run is ~~9~~ 11 October and for SnowModel ~~20 September~~27 October.

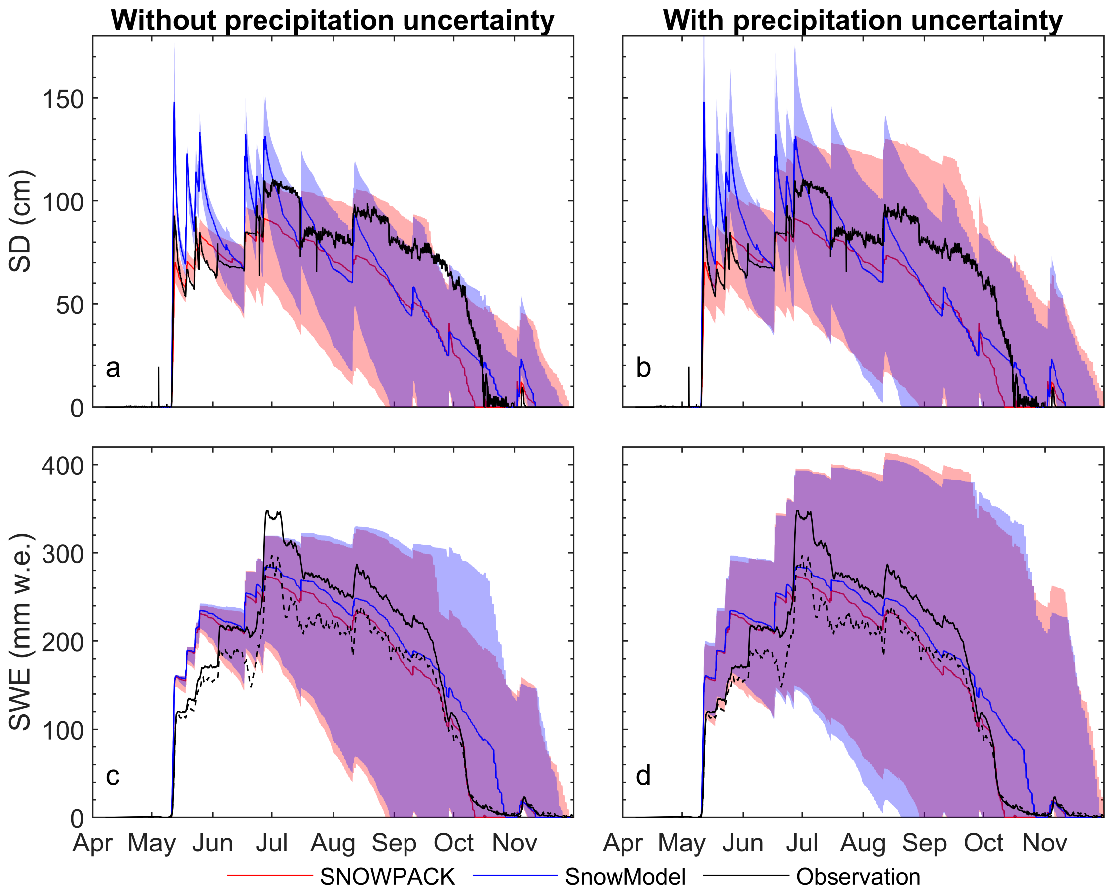

Figure 5. Observed (black) and simulated (colour) SD and SWE by SnowPack and SnowModel forced by the 1000 ensembles of meteorological data. The reference run (see Sect. 3.2.3) of both models is the bold coloured curve. The shaded area corresponds to the 1000 runs of a)-b) snow depth and SWE c)-d) of SNOWPACK (red) and SnowModel (blue) for the ~~MC~~ run with biased forcing. a)-c) are without precipitation uncertainties and b)-d) are with precipitation uncertainties.~~The dotted line of SWE indicates the less reliable (lower) SWE measurement from thallium rays (See Sect. 2.3)~~

For SNOWPACK, the spread of the averaged sublimation rates corresponding to the ensemble runs from the first day of snow to ~~15 October ranges between 1.48 and 1.61~~ 20 September ranges between 1.41 and 2.96 mm w.e. d^{-1} (Fig. 6a). During the cold period, when no melt occurs, ~~all parameterizations result in similar sublimation amounts~~ the sublimation amounts mainly depend on the z_0, with sublimation rates ranging between ~~1.16 and 1.32~~ 1.40 and 3.18 mm w.e. d^{-1}. At the end of the season, the total sublimation ranges between ~~142 and 179~~ 153 and 364 mm w.e. (corresponding to ~~51.3 to 64.6~~36.2 to 86.0% of the total ablation). During the melting period, the ensemble runs show a large spread of melt rates ranging between ~~2.15 to 9.05~~ 0.97 to 17.7 mm w.e. d^{-1}. The total amount of runoff is between ~~60 and 105~~ 28.9 and 236 mm w.e. for SNOWPACK and this model also simulates evaporation, which contributes ~~for 7.3 to~~ between 2.5 and 10.2% of total ablation (Fig. 6a).

For SnowModel, sublimation differences between the parameterizations are ~~larger~~ similar (Fig. 6b) with average sublimation rates from the first day of snow to 20 September ranging between ~~1.26 to 2.04~~ 1.27 to 2.79 mm w.e. d^{-1}~~, and it is worth noting that significant differences are also observed during the cold season (Fig. 6a)~~. At the end of the winter season the sublimation totals range between ~~119 and 176~~ 154 and 342 mm w.e. (which corresponds to ~~42.7 to 63.5~~36.4 to 80.7% of the total ablation). ~~Significant differences are also calculated during the melting period, when the simulated melt rate leads to a total melt difference of 58~~ The runoff is between 81.7 and 269 mm w.e.~~at the end of the season.~~. A closer analysis of Fig. 6b shows that SnowModel's output clusters into four groups, where the grouping is determined by the albedo parameterization and z_0 with limited influence of fresh snow density parameterizations. The two lower clusters are linked to the Z_0 value of 1 mm. The differences between clusters for different z_0 values increase as the difference in albedo between the parameterizations increase at the end of June.

~~The spread of the ensemble runs for both sublimation and melt simulated by SnowModel can be attributed to the choice of the albedo parameterization.~~ While the ensemble ~~calibration~~ parameterization simulations do not lead to significant differences in the modelled end date of the snow season (i.e. difference of ~~eight~~ 22 days), the ~~parameterization strongly impacts~~ albedo parameterization and z_0 value directly impact the proportion of sublimation versus melt to the total ablation (Fig. 6b,d). During the cold period, simulations considering the lowest albedo and z_0 of 1 cm (the reference simulation), lead to a higher sublimation rate (Fig. 6b). Indeed, a lower albedo increases the energy absorbed by the snowpack, and as the temperature is below the freezing point, this energy leads to an increase in the sublimation. A higher z_0 enhances this process even more as this leads to a more negative latent heat flux. Second, the increase of net ~~SW~~ shortwave radiation also affects the physical properties of the snowpack resulting in an increase of compaction (Fig. 4b,h). The snow density of the snowpack is therefore higher~~(Fig. S6.1 in SM)~~, which directly affects the thermal conductivity of the upper snow layers (Yen, 1981). ~~Changing the surface temperature~~ Surface temperature variation is directly linked to the latent heat flux and therefore to ~~the~~ sublimation, explaining the different sublimation ratios simulated depending on the albedo parameterizations and z_0 values.

~~Contrary~~ In contrast to SnowModel, the albedo parameterization ~~of~~ in SNOWPACK does not affect the sublimation but ~~strongly~~ noticeably influences the melting ~~speed~~ rate (Fig.6c), which can be attributed to the more complex characteristics of this model. SNOWPACK allows refreezing and evaporation of melting snow within the snowpack, which can lead to a longer melt season, whereas calculated evaporation leads to a lower amount of runoff from melt. Also, SNOWPACK considers a more complex representation of snow physics, such as the grain size and the snow surface area (SSA), which directly impacts the albedo and can help to explain the wide diversity of melt simulations.

~~When comparing the ablation and energy fluxes (Fig. 6 and **??**), it is clear that the sublimation rates of both reference runs are very similar (difference of 27.6 mm w.e. of the cumulative sublimation at 1 September), whereas subtle differences are related to differences in turbulent and latent heat fluxes (average cold period Q_L is -28.3 W m^{-2} for SNOWPACK vs. -34.5 W m^{-2} for SnowModel).~~

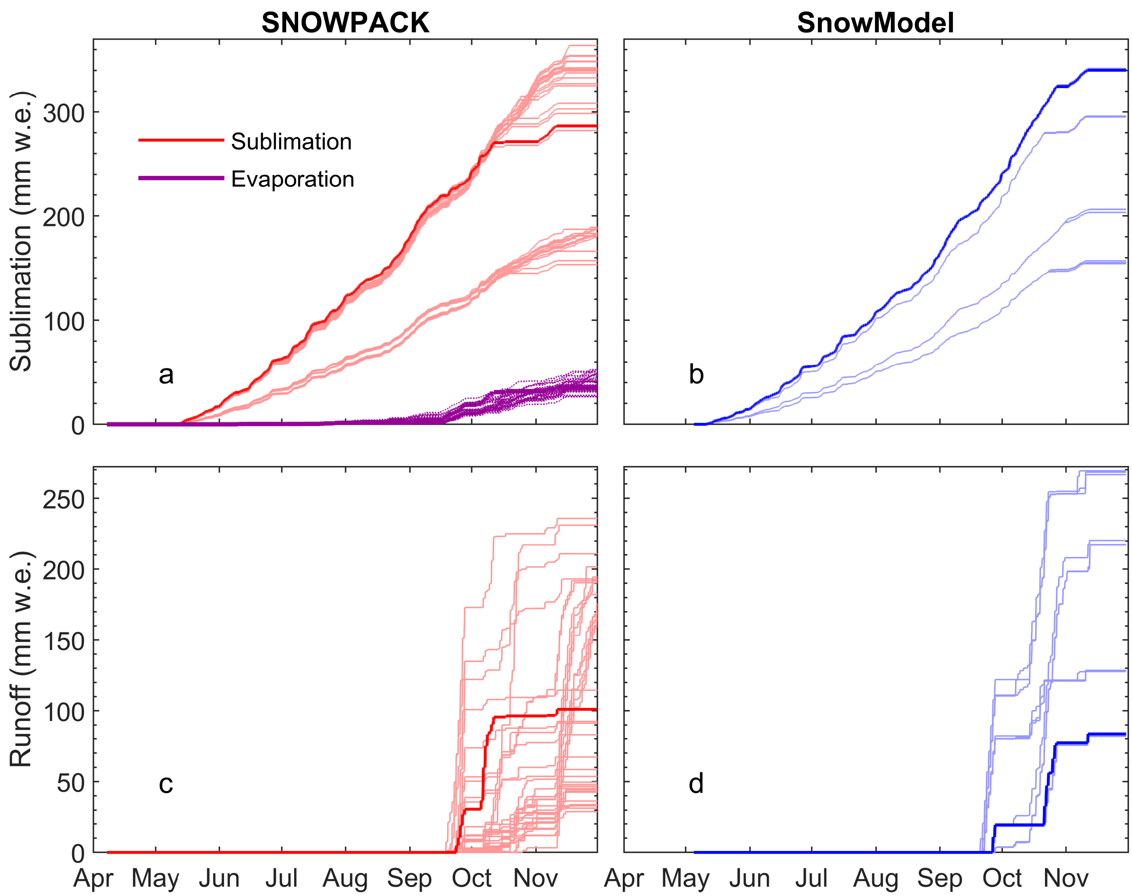

Figure 6. Cumulative sublimation (a,b) and runoff from melt (c,d), simulated by SNOWPACK (red) and SnowModel (blue). For SNOWPACK the cumulative evaporation from melt is shown (purple lines in a). Results for all the ~~ensemble calibrations~~ ensembles of parameterizations are shown and the bold lines correspond to the reference simulations of each model.

~~At the end of the season, the cumulative melt is comparable between the reference simulations of the two models (90.8 mm w.e. for SNOWPACK and 86.5 mm w.e. for SnowModel, Fig. 6c,d), although with large differences in melt distribution. For SnowModel, the entire snowpack melts within a few days (i.e. 11.2 mm w.e. d^{-1} between 12 and 19 September) while melt is later and slower in SNOWPACK (7.3 mm w.e. d^{-1} between 18 and 27 September). This corresponds to higher net SW radiation and lower albedo simulations in SnowModel during the melting period (Fig. **??**), which increases the energy available for melt and subsequent runoff.~~

~~Evaporation and refreezing of meltwater also explain part of the differences in ablation results between the models. SNOWPACK simulates 0.30 mm w.e. d^{-1} of evaporation of the liquid water contained in the snowpack (10% of total ablation corresponding to a total of 28 mm) and also allows for refreezing due to the low temperature during the night. These processes, not considered~~

~~by SnowModel, reduce the runoff simulated by SNOWPACK, whereas SnowModel often encounters the snowpack threshold density to activate direct run-off in the bucket scheme.~~

~~Stacked daily averages of the modelled heat fluxes at Tapado AWS simulated with the reference runs of a) SNOWPACK and b) SnowModel. SW_{net} is net shortwave radiation, LW_{net} is net longwave radiation, Q_S is the sensible heat flux, Q_L is the latent heat flux, Q_G is the heat flux from the ground (only simulated by SNOWPACK), and Q_M is the energy consumed by melt. The fluxes are only plotted at the times where the models simulate snow.~~

5 Discussion

5.1 Model sensitivity and comparison

Our results show the importance of model ~~parameterization and model structure~~ parameterizations and model forcing over the snow model choice, despite the limited model options chosen for the ensemble approach, and the large differences in the two model complexities chosen in this study. This conclusion, found here in an arid environment, is in ~~total~~ agreement with the studies performed in ~~Alpine~~ other alpine areas (Etchevers et al., 2004; Günther et al., 2019). ~~Günther et al. (2019) highlighted for instance the importance of the model structure as being higher than the choice of the parameter values.~~

In this study, the albedo parameterization appears the ~~most important parameter~~ one of the most important parameters to be properly assessed (Fig. 4, 6), as also reported by the studies performed in Alpine regions (Etchevers et al., 2004; Zolles et al., 2019). This can be surprising at first glance as in the semi-arid Andes the ablation is mainly driven by the sublimation and the albedo ~~calibration~~ parameterization is generally crucial to accurately simulate the melt. However, according to the results presented here, the two models agree with the larger sensitivity to the albedo parameterization.

The biggest differences between the models are found as the snow settles and therefore depends on the snow density parameterization. The challenge in this study was that the snow settling showed two distinct regimes. From May to mid-June, high compaction rates were found, whereas the compaction afterwards was more moderate. SNOWPACK is able to model both moderate and high compaction, depending on the parameterization chosen, but the mode of compaction remains the same over the season. SnowModel simulates high compaction rates for all parameterizations, which is correct for the start of the season but an overestimation after mid-June. These compaction rates implicate changes in the thermal conductivity of the snowpack and thus changes in the melting. The different snow density parameterizations in SNOWPACK are still developed and improved (e.g. Keenan et al., 2021), but an improvement of snow density parameterizations in semi-arid regions shows a demand for snow density measurements, as observations are biased (Smith et al., 2017).

~~This is likely related to the absence of the consideration of the~~ The uncertainty associated to the turbulent fluxes parameterization in the ensemble approach is only considered by implementing two different z_0 values, while it can have major implications in surface energy balance modelling (e.g. Dadic et al., 2013; Conway and Cullen, 2013; Litt et al., 2017; Réveillet et al., 2020). The representation of turbulent fluxes in snow models is commonly based on the bulk method, and the Richardson number is ~~usually used~~often used, together with the Monin–Obukhov similarity theory, to evaluate the atmosphere stability (e.g. Liston and Hall, 1995; Vionnet et al., 2012). ~~The roughness value is therefore used to parameterize the turbulent fluxes~~Here, only the

Richardson number is used as both models offered this option. While the stability function cannot be compared between the two models chosen in this study, the sensitivity of SNOWPACK to the six possibilities available in the current version is low (i.e. max SD difference of few centimeters, results not shown). In a study over Brewster Glacier in New-Zealand, Conway and Cullen (2013) pointed out the importance of the stability functions to properly simulate the heat fluxes with low wind speed and large temperature gradients and also that the modelled latent heat fluxes were unaffected by the choice of exchange coefficient parameterization. The present study takes place in a dry and windy environment without a large temperature gradient and ~~thus, this explains the low~~ this helps to explain the small differences observed related to the ~~chosen atmospheric stability function~~different atmospheric stability functions. The turbulent fluxes parameterization is sensitive to the ~~roughness~~ z_0 value and observations, such as Eddy Covariance measurements, are essential to accurately parameterize the turbulent fluxes (e.g. Conway and Cullen, 2013; Litt et al., 2017; Réveillet et al., 2018).

Due to the absence of such measurements, ~~and also due to the strong~~ the variability of this value over ~~the time (e.g. MacDonell et al., 2013~~ ~~a sensitivity analysis was performed (see Sect. S3) . Here we varied the value within a random range and found a larger range of uncertainty for SNOWPACK than for SnowModel . This is likely due to the higher complexity in physical representation of SNOWPACK, but it does not allow us to conclude if this uncertainty is higher than the albedo parameterization~~time (e.g. MacDonell et al., 2013b; Pellicciotti et al., 2005; Nicholson et al., 2016) and due to other values in literature at other locations showing a wide variety (two orders of magnitude) of the snow roughness length (Gromke et al., 2011; Poggi, 1977; Bintanja and Broeke, 19 it was decided to use two different values for z_0 (1 mm and 1 cm). Similar sensitivity ranges for SNOWPACK and SnowModel were found (e.g. Fig. 4), along with similar sublimation rates, but this directly depended on the value for z_0. For both models, a z_0 of 1 cm led to better simulations (Fig. 3, 4), but as applied more often, this can also be seen as optimizing parameter (e.g. Stigter et al., 2018). In future work, the z_0 can be verified with eddy covariance measurements.

The impact of parameterization choice differs for the two models as the uncertainty is directly related to the difference in snow physical representation and the characteristics of the models. Indeed, the range of the ensemble approach simulated by SNOWPACK is higher than that simulated by SnowModel directly related to both higher ~~calibration choice~~ number of parameterization possibilities for SNOWPACK and more complex physical representation of the processes.

Likewise, results presented here show that the main sensitivity remains in the forcing uncertainty, in agreement with previous studies ~~(Magnusson et al., 2015; Günther et al., 2019; Raleigh et al., 2015)~~(Magnusson et al., 2015; Günther et al., 2019; Raleigh et al., 20 For instance, ~~(Magnusson et al., 2015)~~ Magnusson et al. (2015) found that the models of different complexity (temperature-index models vs. physical models) show similar ability to reproduce daily observed snowpack runoff, and concluded that the forcing uncertainties is the greatest factor affecting model performance, rather than model parameterizations. However, as mentioned by Raleigh et al. (2015), simulated SD and SWE are critically sensitive to the relative magnitude of errors in forcing. ~~They~~ Raleigh et al. (2015) and Schlögl et al. (2016) also mentioned that precipitation bias (or correction of the undercatch of the precipitation gauge) was the most important factor, in agreement with the findings of our study.

Finally, while Rutter et al. (2009) pointed out that no universal 'best' model exists and model performance ~~strongly~~ directly depends on the study site~~, here~~. Here we show, that in a semi-arid environment, the conclusions are similar for Alpine environments in that model structure and parameterization choices represent an important step to improve model performance.

5.2 Limitations of the study and further ~~works~~work

The sensitivity study of the two models to the forcing is done by ~~perturbing~~ adding a bias to the meteorological variables with ~~random noise and~~ ranges derived from literature. It was also possible to add random noise to the data, but this does not necessarily preserve the physical consistency ~~. Some studies (e.g. Charrois et al., 2016) maintained this consistency by applying~~ and would lead to low sensitivity of the models (results not shown), as random noise counterbalances each other, which has also been investigated by Raleigh et al. (2015). It would also have been possible to apply a random perturbation (e.g. Charrois et al., 2016) using a first-order auto-regressive model (Deodatis and Shinozuka, 1988). However, the forcing ~~perturbation~~ bias does not affect the conclusion of the relative comparison of the two models which only requires the exact same forcing as input to be relevant. For the same reason, the choice of method applied for the forcing correction (i.e. for precipitation) and reconstruction (i.e. for the TA and RH) would not affect the conclusions of the model comparison. However, due to the ~~forcing~~ precipitation uncertainty related to measurements errors, and also because the sensor locations may not be representative of the area, ~~only a qualitative comparison between the simulated and observed SWE and SD measurement is done in this work. Günther et al. (2019)~~ different ways to correct the precipitation data were proposed. Günther et al. (2019) and Grünewald and Lehning (2014) already outlined that the snow cover is spatially heterogeneous even at very small scales due to topographic and microclimatic effects on accumulation, redistribution, and ablation processes, introducing an uncertainty in validation data. We also show that in any case, due to i) the question of the sensor location representativity of the area, ii) the precipitation undercatch because of the wind, and, iii) the ~~strong~~ high sensitivity of models to precipitation uncertainty, this study highlights the complexity and necessity of accurately measuring precipitation. Additionally, possible corrections also depend on the availability of observations, but this study was restricted to not using SWE and SD as forcing, as these parameters were needed as validation data. Therefore, a precipitation correction has been chosen that overestimates the start of the season, but also, for example, does not take a rise of SD and SWE mid-June into account leading to good correspondence of simulated and observed SWE from the start of July (e.g. Fig. 4).

The ensemble approach ~~of the calibration~~ with different parameterizations is built considering limited parameterization options, contrary to other studies where a large number of physical options are considered (e.g. Essery et al., 2013; Lafaysse et al., 2017; Zolles et al., 2019). In our case, choosing snow models with different physical complexities limits the number of calibration possibilities, as ~~calibration of~~ the parameterization of the same variables are chosen for the comparison. Thus, some parameterizations, such as the choice of atmospheric stability correction only available in SNOWPACK, were excluded and this model is calibrated following the same options found in SnowModel (Sect. 3.2.1).

Testing different albedo parameterizations is chosen as i) different options are possible in both models and ii) previous studies concluded that the largest absolute uncertainties originate from the shortwave radiation and the albedo parameterizations (e.g. Zolles et al., 2019). The sensitivity test to different fresh snow density parameterizations was also chosen as previous studies identified this ~~calibration~~ parameterization as a significant ~~uncertainties~~ uncertainty in model calibration (e.g. Essery et al., 2013). Finally, energy balance models are known to be sensitive to z_0, especially in cold and dry regions where sublimation is the main ablation process (e.g. Réveillet et al., 2020). However, due to this important sensitivity and the absence of measurements to properly calibrate this value, ~~the sensitivity test by changing the value within a random range is performed independently (see Sect. S3)~~two values for z_0 were implemented, but this still might underestimate the possible range of z_0 values.

Otherwise, despite the choice of limiting the ~~calibration~~ parameterization options, SNOWPACK's sensitivity to model ~~calibration~~ parameterizations is evaluated based on ~~20~~ 40 simulations, whereas SnowModel's evaluation is based on ~~six~~ 12 simulations only. However, the difference of the number of ~~simulation~~ simulations does not impact the conclusion, as the width of the spread of different ~~calibrations~~ parameterizations was not quantitatively assessed.

Among the possible ~~calibrations~~settings of the model, the snow transport option has not been activated, while the option is available in both models. However, due to the strong wind speed characteristics of the study area (Gascoin et al., 2013, and Fig. 2), snow transport is expected to be considerable. Yet, snow transport estimation remains out of the scope of this study, focused on energy balance comparisons, mainly to assess differences in sublimation rates. Also, in a study performed in the Pascua-Lama catchment, a region to the north of Tapado AWS, Gascoin et al. (2013) highlighted that the inclusion of Snow-Tran3D does not change the fact that the model is unable to capture the small-scale snow depth spatial variability (as captured by in-situ snow depth sensors). ~~Last~~Finally, snow transport in SNOWPACK can only be simulated ~~with a SD assimilation, which~~ in the 3D domain with SD as forcing, which then could not be used as validation data. However, due to the importance and complexity in modelling snow transport, properly assessing its impact could be assessed in future work.

6 ~~Conclusion~~Conclusions

Snow models are key to quantify runoff and provide accurate water availability projections. The aim of this study is to compare two snow models, SNOWPACK and SnowModel, and evaluate their sensitivity relative to ~~calibration~~ parameterization and forcing. For that purpose, the two models are run over the 2017 snow season, at local point, and forced with i) the most ideal set of input parameters, ii) an ensemble of different physical ~~calibrations and ii~~parameterizations and iii) an ensemble of ~~perturbed~~ biased forcing.

The most ideal set of input parameters consisted of observed forcing and the validation parameters (SD and $S_{\downarrow}$ for SNOWPACK; SWE for SnowModel) given as input. Hence, the models were able to assimilate the forced precipitation to correct for undercatch in the precipitation gauge. SNOWPACK is able to approach the observation very well (i.e. min. $RMSE$ of 9.2 cm, max. R^2 of 0.93 calculated with the observed and simulated SD), but SnowModel only adjusts the precipitation at two precipitation events, still leading to undercatch. The final correction of the precipitation data was done with an equation based on TA and WS, as it was unwanted to adjust the precipitation with SNOWPACK's assimilated data, as this assimilated data is built from data that is needed as validation data. The ~~calibration ensemble simulations were built~~ parameterization simulations were done considering different parameterizations of the albedo and the fresh snow density and different values for z_0. Results indicated a significant difference related mainly to the ~~calibration~~ parameterization choice of the albedo and z_0. However, the impact of ~~this choice~~ the albedo affects the two models differently. For SnowModel, the albedo parameterization has a significant impact on the simulated sublimation during the cold period ~~(i.e. difference of ∼40 mm w.e.)~~ while SNOWPACK simulates similar sublimation rates for all the possible ~~calibrations. The albedo parameterization of SNOWPACK has stronger~~ parameterizations. The choice of albedo parameterization in SNOWPACK has direct consequences on melt at the end of the season~~leading to a difference of 45 mm w.e. of the total melt, and a difference of the end of the season of 41 days, related to the calibration chosen. This~~ . The model differences are mainly related to the model characteristics (~~i.e.~~e.g. the consideration of the water evaporation and refreezing into the snowpack), and the more complex representation of the snow physics in SNOWPACK. However, the models are both sensitive to the chosen z_0, leading to sublimation rates ranging rom 36% up to 86%. In addition, results presented in this study highlight a larger uncertainty depending on the model ~~calibration~~ parameterization (despite the limited number of options chosen) than between the two models (even though the significant differences in their physical complexity).

~~The sensitivity of both modelsto the forcing data is highly influenced by the precipitation uncertainties. As for ensemble calibration study, SNOWPACK shows larger uncertainty in modelling the end of the season (i.e. difference of 36 days) compared to SnowModel (i.e. 20 days), likely related to more complex physical processes. SnowModel shows larger uncertainty during the start of the melting period, but , a small difference on the simulated date of the end of the season. This is attributed to the fast melt of the entire snowpack (i.e. few days) simulated by this model.~~

Otherwise, for both models~~, results show uncertainty of similar magnitude between the calibration and forcing. However, this result strongly depends on the bias chosen for the forcing perturbation as well as the number of calibrations selected, which, here, does not enable us to draw a conclusion about the larger source of uncertainty or the most appropriate model to model snow depth evolution in the semi-arid Andes.~~ , results show high levels of uncertainty related to forcings which is direcly related to the bias chosen, but the spread of the uncertainty for both models is approximately the same. SNOWPACK and SnowModel are highly influenced by precipitation uncertainties. Both models show similar levels of uncertainty, in modelling the end of the season.

.

This study is performed over one winter season providing conclusions of the model sensitivity, specific for this winter. In further studies, simulations could be performed over a larger time period, and at distinct places to reinforced and discuss the conclusions presented here. In addition, the model choice could be extended to other models, and in particular snow models with similar physical complexity. Such work would provide additional information of the ~~calibration~~ parameterization sensitivity by allowing a comparison based on a larger choice of possible ~~calibrations~~parameterizations.

Data availability. Part of the data used in this paper (AWS data) can be accessed at https://www.ceazamet.cl. SnowModel can be accessed by contacting the administrator, Glen E. Liston. SNOWPACK is an Open Source and can be accessed at https://models.slf.ch/p/snowpack/. For any other access to the data presented in this study, please contact the authors.

Author contributions. AV conducted data preparation, ran the numerical experiments and produced the figures. AV and MR designed the modelling strategy. MR and SM designed the study. All authors contributed to the results analysis and to the preparation of the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. ~~The authors~~ We thank Michael Lehning and two anonymous referees as well as the editor, Brice Noël, for their constructive comments, which helped to sharpen the scope of this study. We also thank Glen E. Liston for providing the code of Snow-Model ~~. We are also~~ and are grateful to CEAZAmet and the CEAZA glaciology group for maintaining the Tapado AWS and data centre.

Marion Réveillet and Shelley MacDonell were supported by CONICYT-Programa Regional R16A10003, and the Coquimbo regional government FIC-R(2015) BIP 30403127-0. Marion Réveillet was also supported by the ANR program: ANR-16-CE01-0006 EBONI.

References

Andreas, E. L., Jordan, R. E., and Makshtas, A. P.: Parameterizing turbulent exchange over sea ice: the ice station weddell results, Boundary-Layer Meteorology, 114, 439–460, https://doi.org/10.1007/s10546-004-1414-7, 2005.

Ayala, A., Pellicciotti, F., Peleg, N., and Burlando, P.: Melt and surface sublimation across a glacier in a dry environment: distributed energy-balance modelling of Juncal Norte Glacier, Chile, Journal of Glaciology, 63, 803–822, https://doi.org/10.1017/jog.2017.46, 2017.

Bartelt, P. and Lehning, M.: A physical SNOWPACK model for the Swiss avalanche warning: Part I: Numerical model, Cold Regions Science and Technology, 35, 123–145, https://doi.org/10.1016/s0165-232x(02)00074-5, 2002.

Bavay, M. and Egger, T.: MeteoIO 2.4. 2: a preprocessing library for meteorological data, Geosci, Model Dev. Discuss, 7, 3595–3645, 2014.

Bintanja, R. and Broeke, M. R. V. D.: Momentum and scalar transfer coefficients over aerodynamically smooth antarctic surfaces, Boundary-Layer Meteorology, 74, 89–111, https://doi.org/10.1007/bf00715712, 1995.

Blandford, T. R., Humes, K. S., Harshburger, B. J., Moore, B. C., Walden, V. P., and Ye, H.: Seasonal and Synoptic Variations in Near-Surface Air Temperature Lapse Rates in a Mountainous Basin, Journal of Applied Meteorology and Climatology, 47, 249–261, https://doi.org/10.1175/2007jamc1565.1, 2008.

Braithwaite, R. J. and Olesen, O. B.: Calculation of Glacier Ablation from Air Temperature, West Greenland, in: Glaciology and Quaternary Geology, pp. 219–233, Springer Netherlands, https://doi.org/10.1007/978-94-015-7823-3_15, 1989.

Brock, B. W., Willis, I. C., and Sharp, M. J.: Measurement and parameterization of albedo variations at Haut Glacier d'Arolla, Switzerland, Journal of Glaciology, 46, 675–688, https://doi.org/10.3189/172756500781832675, 2000.

Brock, B. W., Willis, I. C., and Sharp, M. J.: Measurement and parameterization of aerodynamic roughness length variations at Haut Glacier d'Arolla, Switzerland, Journal of Glaciology, 52, 281–297, https://doi.org/10.3189/172756506781828746, 2006.

Charrois, L., Cosme, E., Dumont, M., Lafaysse, M., Morin, S., Libois, Q., and Picard, G.: On the assimilation of optical reflectances and snow depth observations into a detailed snowpack model, The Cryosphere, 10, 1021–1038, https://doi.org/10.5194/tc-10-1021-2016, 2016.

Conway, J. P. and Cullen, N. J.: Constraining turbulent heat flux parameterization over a temperate maritime glacier in New-Zealand, Annals of Glaciology, 54, 41–51, https://doi.org/10.3189/2013aog63a604, 2013.

Cuffey, K. M. and Paterson, W. S. B.: The Physics of Glaciers, Elsevier LTD, Oxford, https://www.ebook.de/de/product/10550595/kurt_m_cuffey_w_s_b_paterson_the_physics_of_glaciers.html, 2010.

Dadic, R., Mott, R., Lehning, M., Carenzo, M., Anderson, B., and Mackintosh, A.: Sensitivity of turbulent fluxes to wind speed over snow surfaces in different climatic settings, Advances in Water Resources, 55, 178–189, https://doi.org/10.1016/j.advwatres.2012.06.010, 2013.

Deodatis, G. and Shinozuka, M.: Auto-Regressive Model for Nonstationary Stochastic Processes, Journal of Engineering Mechanics, 114, 1995–2012, https://doi.org/10.1061/(asce)0733-9399(1988)114:11(1995), 1988.

Douville, H., Royer, J.-F., and Mahfouf, J.-F.: A new snow parameterization for the Météo-France climate model. Part I: validation in stand-alone experiments, Climate Dynamics, 12, 21–35, https://doi.org/10.1007/s003820050092, 1995.

Dutra, E., Balsamo, G., Viterbo, P., Miranda, P. M. A., Beljaars, A., Schär, C., and Elder, K.: An Improved Snow Scheme for the ECMWF Land Surface Model: Description and Offline Validation, Journal of Hydrometeorology, 11, 899–916, https://doi.org/10.1175/2010jhm1249.1, 2010.

Essery, R.: A factorial snowpack model (FSM 1.0), Geoscientific Model Development, 8, 3867–3876, https://doi.org/10.5194/gmd-8-3867-2015, 2015.

Essery, R., Morin, S., Lejeune, Y., and Ménard, C. B.: A comparison of 1701 snow models using observations from an alpine site, Advances in Water Resources, 55, 131–148, https://doi.org/10.1016/j.advwatres.2012.07.013, 2013.

Etchevers, P., Martin, E., Brown, R., Fierz, C., Lejeune, Y., Bazile, E., Boone, A., Dai, Y.-J., Essery, R., Fernandez, A., Gusev, Y., Jordan, R., Koren, V., Kowalczyk, E., Nasonova, N. O., Pyles, R. D., Schlosser, A., Shmakin, A. B., Smirnova, T. G., Strasser, U., Verseghy, D., Yamazaki, T., and Yang, Z.-L.: Validation of the energy budget of an alpine snowpack simulated by several snow models (Snow MIP project), Annals of Glaciology, 38, 150–158, https://doi.org/10.3189/172756404781814825, 2004.

Favier, V., Falvey, M., Rabatel, A., Praderio, E., and López, D.: Interpreting discrepancies between discharge and precipitation in high-altitude area of Chile's Norte Chico region (26-32°S), Water Resources Research, 45, https://doi.org/10.1029/2008wr006802, 2009.

Gascoin, S., Lhermitte, S., Kinnard, C., Bortels, K., and Liston, G. E.: Wind effects on snow cover in Pascua-Lama, Dry Andes of Chile, Advances in Water Resources, 55, 25–39, https://doi.org/10.1016/j.advwatres.2012.11.013, 2013.

Gromke, C., Manes, C., Walter, B., Lehning, M., and Guala, M.: Aerodynamic Roughness Length of Fresh Snow, Boundary-Layer Meteorology, 141, 21–34, https://doi.org/10.1007/s10546-011-9623-3, 2011.

Grünewald, T. and Lehning, M.: Are flat-field snow depth measurements representative? A comparison of selected index sites with areal snow depth measurements at the small catchment scale, Hydrological Processes, 29, 1717–1728, https://doi.org/10.1002/hyp.10295, 2014.

Günther, D., Marke, T., Essery, R., and Strasser, U.: Uncertainties in Snowpack Simulations—Assessing the Impact of Model Structure, Parameter Choice, and Forcing Data Error on Point-Scale Energy Balance Snow Model Performance, Water Resources Research, 55, 2779–2800, https://doi.org/10.1029/2018wr023403, 2019.

Hock, R.: Temperature index melt modelling in mountain areas, Journal of Hydrology, 282, 104–115, https://doi.org/10.1016/s0022-1694(03)00257-9, 2003.

Keenan, E., Wever, N., Dattler, M., Lenaerts, J. T. M., Medley, B., Munneke, P. K., and Reijmer, C.: Physics-based SNOWPACK model improves representation of near-surface Antarctic snow and firn density, The Cryosphere, 15, 1065–1085, https://doi.org/10.5194/tc-15-1065-2021, 2021.

Krinner, G., Derksen, C., Essery, R., Flanner, M., Hagemann, S., Clark, M., Hall, A., Rott, H., Brutel-Vuilmet, C., Kim, H., Ménard, C. B., Mudryk, L., Thackeray, C., Wang, L., Arduini, G., Balsamo, G., Bartlett, P., Boike, J., Boone, A., Chéruy, F., Colin, J., Cuntz, M., Dai, Y., Decharme, B., Derry, J., Ducharne, A., Dutra, E., Fang, X., Fierz, C., Ghattas, J., Gusev, Y., Haverd, V., Kontu, A., Lafaysse, M., Law, R., Lawrence, D., Li, W., Marke, T., Marks, D., Ménégoz, M., Nasonova, O., Nitta, T., Niwano, M., Pomeroy, J., Raleigh, M. S., Schaedler, G., Semenov, V., Smirnova, T. G., Stacke, T., Strasser, U., Svenson, S., Turkov, D., Wang, T., Wever, N., Yuan, H., Zhou, W., and Zhu, D.: ESM-SnowMIP: assessing snow models and quantifying snow-related climate feedbacks, Geoscientific Model Development, 11, 5027–5049, https://doi.org/10.5194/gmd-11-5027-2018, 2018.

Lafaysse, M., Cluzet, B., Dumont, M., Lejeune, Y., Vionnet, V., and Morin, S.: A multiphysical ensemble system of numerical snow modelling, The Cryosphere, 11, 1173–1198, https://doi.org/10.5194/tc-11-1173-2017, 2017.

Lehning, M., Bartelt, P., Brown, B., and Fierz, C.: A physical SNOWPACK model for the Swiss avalanche warning: Part III: Meteorological forcing, thin layer formation and evaluation, Cold Regions Science and Technology, 35, 169–184, https://doi.org/10.1016/s0165-232x(02)00072-1, 2002a.

Lehning, M., Bartelt, P., Brown, B., Fierz, C., and Satyawali, P.: A physical SNOWPACK model for the Swiss avalanche warning: Part II: Snow microstructure, Cold Regions Science and Technology, 35, 147–167, https://doi.org/10.1016/s0165-232x(02)00073-3, 2002b.

Lehning, M., Völksch, I., Gustafsson, D., Nguyen, T. A., Stähli, M., and Zappa, M.: ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology, Hydrological Processes, 20, 2111–2128, https://doi.org/10.1002/hyp.6204, 2006.

Liston, G. E. and Elder, K.: A Meteorological Distribution System for High-Resolution Terrestrial Modeling (MicroMet), Journal of Hydrometeorology, 7, 217–234, https://doi.org/10.1175/jhm486.1, 2006a.

Liston, G. E. and Elder, K.: A Distributed Snow-Evolution Modeling System (SnowModel), Journal of Hydrometeorology, 7, 1259–1276, https://doi.org/10.1175/jhm548.1, 2006b.

Liston, G. E. and Hall, D. K.: An energy-balance model of lake-ice evolution, Journal of Glaciology, 41, 373–382, https://doi.org/10.3189/s0022143000016245, 1995.

Liston, G. E. and Sturm, M.: A snow-transport model for complex terrain, Journal of Glaciology, 44, 498–516, https://doi.org/10.3189/s0022143000002021, 1998.

Litt, M., Sicart, J.-E., Six, D., Wagnon, P., and Helgason, W. D.: Surface-layer turbulence, energy balance and links to atmospheric circulations over a mountain glacier in the French Alps, The Cryosphere, 11, 971–987, https://doi.org/10.5194/tc-11-971-2017, 2017.

MacDonald, J. and Pomeroy, J.: Gauge undercatch of two common snowfall gauges in a prairie environment, in: Proceedings of the 64th Eastern Snow Conference, vol. 29, pp. 119–126, 2007.

MacDonell, S., Kinnard, C., Mölg, T., Nicholson, L., and Abermann, J.: Meteorological drivers of ablation processes on a cold glacier in the semi-arid Andes of Chile, The Cryosphere, 7, 1513–1526, https://doi.org/10.5194/tc-7-1513-2013, 2013a.

MacDonell, S., Nicholson, L., and Kinnard, C.: Parameterisation of incoming longwave radiation over glacier surfaces in the semi-arid Andes of Chile, Theoretical and Applied Climatology, 111, 513–528, https://doi.org/10.1007/s00704-012-0675-1, 2013b.

Magnusson, J., Wever, N., Essery, R., Helbig, N., Winstral, A., and Jonas, T.: Evaluating snow models with varying process representations for hydrological applications, Water Resources Research, 51, 2707–2723, https://doi.org/10.1002/2014wr016498, 2015.

Menard, C. B., Essery, R., Krinner, G., Arduini, G., Bartlett, P., Boone, A., Brutel-Vuilmet, C., Burke, E., Cuntz, M., Dai, Y., Decharme, B., Dutra, E., Fang, X., Fierz, C., Gusev, Y., Hagemann, S., Haverd, V., Kim, H., Lafaysse, M., Marke, T., Nasonova, O., Nitta, T., Niwano, M., Pomeroy, J., Schädler, G., Semenov, V. A., Smirnova, T., Strasser, U., Swenson, S., Turkov, D., Wever, N., and Yuan, H.: Scientific and Human Errors in a Snow Model Intercomparison, Bulletin of the American Meteorological Society, 102, E61–E79, https://doi.org/10.1175/bams-d-19-0329.1, 2021.

Mengual Henríquez, S. A.: Caracterización de la nieve de distintas localidades de Chile mediante el uso del modelo SNOWPACK, Master's thesis, Universidad de Chile, 2017.

Nicholson, L. I., Pętlicki, M., Partan, B., and MacDonell, S.: 3-D surface properties of glacier penitentes over an ablation season, measured using a Microsoft Xbox Kinect, The Cryosphere, 10, 1897–1913, https://doi.org/10.5194/tc-10-1897-2016, 2016.

Oke, T. R.: Boundary layer climates, Routledge, 2002.

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, Journal of Glaciology, 51, 573–587, https://doi.org/10.3189/172756505781829124, 2005.

Poggi, A.: Heat Balance in the Ablation Area of the Ampere Glacier (Kerguelen Islands), Journal of Applied Meteorology (1962-1982), 16, 48–55, https://doi.org/10.1175/1520-0450(1977)016<0048:hbitaa>2.0.co;2, http://www.jstor.org/stable/26177592, 1977.

Rabatel, A., Castebrunet, H., Favier, V., Nicholson, L., and Kinnard, C.: Glacier changes in the Pascua-Lama region, Chilean Andes (29° S): recent mass balance and 50 yr surface area variations, The Cryosphere, 5, 1029–1041, https://doi.org/10.5194/tc-5-1029-2011, 2011.

Raleigh, M. S., Lundquist, J. D., and Clark, M. P.: Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework, Hydrology and Earth System Sciences, 19, 3153–3179, https://doi.org/10.5194/hess-19-3153-2015, 2015.

Randall, K. L.: Campbell Scientific Web Request (case:83574), Personal communication, 2018.

Réveillet, M., Six, D., Vincent, C., Rabatel, A., Dumont, M., Lafaysse, M., Morin, S., Vionnet, V., and Litt, M.: Relative performance of empirical and physical models in assessing the seasonal and annual glacier surface mass balance of Saint-Sorlin Glacier (French Alps), The Cryosphere, 12, 1367–1386, https://doi.org/10.5194/tc-12-1367-2018, 2018.

Réveillet, M., MacDonell, S., Gascoin, S., Kinnard, C., Lhermitte, S., and Schaffer, N.: Impact of forcing on sublimation simulations for a high mountain catchment in the semiarid Andes, The Cryosphere, 14, 147–163, https://doi.org/10.5194/tc-14-147-2020, 2020.

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., Barr, A., Bartlett, P., Boone, A., Deng, H., Douville, H., Dutra, E., Elder, K., Ellis, C., Feng, X., Gelfan, A., Goodbody, A., Gusev, Y., Gustafsson, D., Hellström, R., Hirabayashi, Y., Hirota, T., Jonas, T., Koren, V., Kuragina, A., Lettenmaier, D., Li, W.-P., Luce, C., Martin, E., Nasonova, O., Pumpanen, J., Pyles, R. D., Samuelsson, P., Sandells, M., Schädler, G., Shmakin, A., Smirnova, T. G., Stähli, M., Stöckli, R., Strasser, U., Su, H., Suzuki, K., Takata, K., Tanaka, K., Thompson, E., Vesala, T., Viterbo, P., Wiltshire, A., Xia, K., Xue, Y., and Yamazaki, T.: Evaluation of forest snow processes models (SnowMIP2), Journal of Geophysical Research, 114, https://doi.org/10.1029/2008jd011063, 2009.

Scaff, L., Rutllant, J. A., Rahn, D., Gascoin, S., and Rondanelli, R.: Meteorological Interpretation of Orographic Precipitation Gradients along an Andes West Slope Basin at 30°S (Elqui Valley, Chile), Journal of Hydrometeorology, 18, 713–727, https://doi.org/10.1175/jhm-d-16-0073.1, 2017.

Schlögl, S., Marty, C., Bavay, M., and Lehning, M.: Sensitivity of Alpine3D modeled snow cover to modifications in DEM resolution, station coverage and meteorological input quantities, Environmental Modelling & Software, 83, 387–396, https://doi.org/10.1016/j.envsoft.2016.02.017, 2016.

Schmucki, E., Marty, C., Fierz, C., and Lehning, M.: Evaluation of modelled snow depth and snow water equivalent at three contrasting sites in Switzerland using SNOWPACK simulations driven by different meteorological data input, Cold Regions Science and Technology, 99, 27–37, https://doi.org/10.1016/j.coldregions.2013.12.004, 2014.

Smith, C. D.: Correcting the wind bias in snowfall measurements made with a Geonor T-200B precipitation gauge and alter wind shield, in: 87th American Meteorological Society Annual Meeting, San Antonio, TX, 2007.

Smith, C. D., Kontu, A., Laffin, R., and Pomeroy, J. W.: An assessment of two automated snow water equivalent instruments during the WMO Solid Precipitation Intercomparison Experiment, The Cryosphere, 11, 101–116, https://doi.org/10.5194/tc-11-101-2017, 2017.

Stigter, E. E., Litt, M., Steiner, J. F., Bonekamp, P. N. J., Shea, J. M., Bierkens, M. F. P., and Immerzeel, W. W.: The Importance of Snow Sublimation on a Himalayan Glacier, Frontiers in Earth Science, 6, https://doi.org/10.3389/feart.2018.00108, 2018.

Strack, J. E., Liston, G. E., and Pielke Sr, R. A.: Modeling snow depth for improved simulation of snow-vegetation-atmosphere interactions, Journal of Hydrometeorology, 5, 723–734, 2004.

Strasser, U. and Marke, T.: ESCIMO.spread - a spreadsheet-based point snow surface energy balance model to calculate hourly snow water equivalent and melt rates for historical and changing climate conditions, Geoscientific Model Development, 3, 643–652, https://doi.org/10.5194/gmd-3-643-2010, 2010.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Moigne, P. L., Martin, E., and Willemet, J.-M.: The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2, Geoscientific Model Development, 5, 773–791, https://doi.org/10.5194/gmd-5-773-2012, 2012.

Weiser, U., Olefs, M., Schöner, W., Weyss, G., and Hynek, B.: Correction of broadband snow albedo measurements affected by unknown slope and sensor tilts, The Cryosphere, 10, 775–790, https://doi.org/10.5194/tc-10-775-2016, 2016.

Wever, N., Fierz, C., Mitterer, C., Hirashima, H., and Lehning, M.: Solving Richards Equation for snow improves snowpack meltwater runoff estimations in detailed multi-layer snowpack model, The Cryosphere, 8, 257–274, https://doi.org/10.5194/tc-8-257-2014, 2014.

Wolff, M. A., Isaksen, K., Petersen-Øverleir, A., Ødemark, K., Reitan, T., and Brækkan, R.: Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: results of a Norwegian field study, Hydrology and Earth System Sciences, 19, 951–967, https://doi.org/10.5194/hess-19-951-2015, 2015.

Wright, M.: Performance Analysis of CS725 Snow Water Equivalent Sensor, Edmonton, AB: Campbell Scientific Corp, 2011.

Yen, Y.-C.: Review of thermal properties of snow, ice, and sea ice, vol. 81, US Army, Corps of Engineers, Cold Regions Research and Engineering Laboratory, 1981.

Zolles, T., Maussion, F., Galos, S. P., Gurgiser, W., and Nicholson, L.: Robust uncertainty assessment of the spatio-temporal transferability of glacier mass and energy balance models, The Cryosphere, 13, 469–489, https://doi.org/10.5194/tc-13-469-2019, 2019.