Multi-criteria evaluation of snowpack simulations in complex alpine terrain with two spatialization approaches

~~Distributed *vs.* semi-distributed simulations of snowpack dynamics in alpine areas: case study in the upper Arve catchment, French Alps, 1989–2015~~

Jesús Revuelto[1,2], Grégoire Lecourt[1], Matthieu. Lafaysse[1], Isabella Zin[2], Luc Charrois[1], Vincent Vionnet[1], Marie Dumont[1], Antoine Rabatel[2], Delphine Six[2], Thomas Condom[2], Samuel Morin[1], Alessandra Viani[2,3], Pascal. Sirguey[4]

[1] Météo-France - CNRS, CNRM, UMR 3589, CEN, Grenoble, France
[2] Université Grenoble Alpes, CNRS, IRD, Institut des Géosciences de l'Environnement (IGE, UMR 5001), Grenoble, France
[3] University of Brescia, Department of Civil Engineering, Architecture, Land, Environment and Mathematics (DICATAM), Brescia, Italy
[4] National School of Surveying, University of Otago, Dunedin, New Zealand

Abstract: We evaluated distributed and semi-distributed modeling approaches to simulating the spatial and temporal evolution of snow and ice over an extended mountain catchment, using the Crocus snowpack model. The distributed approach simulated the snowpack dynamics on a 250-m grid, enabling inclusion of terrain shadowing effects. The semi-distributed approach simulated the snowpack dynamics for discrete topographic classes characterized by elevation range, aspect, and slope. This provided a categorical simulation that was subsequently spatially re-projected over the 250-m grid used for the distributed simulations. The study area (the upper Arve catchment, western Alps~~,~~ France) is characterized by complex topography, including steep slopes, an extensive glaciated area, and snow cover throughout the year. Simulations were carried out for the period 1989–2015 using the SAFRAN meteorological forcing system. The simulations were compared and evaluated using four observation datasets including point snow depth measurements, seasonal and annual glacier surface mass balance, snow covered area evolution based on optical satellite imagery~~sensors~~, and the annual equilibrium-line altitude of glacier zones, derived from satellite images. The results showed that in both approaches the Crocus snowpack model ~~effectively~~ accurately reproduced the snowpack distribution over the study period. Slightly better results were obtained using the distributed approach. The improvement is statistically significant mainly because it includes~~ed~~ the effects of shadows and terrain characteristics (local values of aspect, slope and elevation for each grid cell). However, the minor improvement observed with a much higher computational time does not justify the recommendation of this approach for all

applications; as long as distributed simulations are not combined with new data assimilation techniques and higher-resolution meteorological inputs.

Key words: snowpack simulation, distributed, semi-distributed, mountain areas, glacierized catchments

1. Introduction

The dynamics of the accumulation and melting of snow and ice in mountain areas has major effects on the timing and level of discharge from rivers in downstream areas. One-sixth of the Earth's population depends directly on the water supply from snow and ice melt in mountain areas (Barnett *et al.*, 2005). Thus, significant research effort has been applied to the study of snow and ice dynamics in these regions (Egli and Jonas, 2009; Lehning *et al.*, 2011; López-Moreno *et al.*, 2013; McCreight *et al.*, 2012), with particular focus on mountain hydrology (DeBeer and Pomeroy, 2009; López-Moreno and García-Ruiz, 2004; Oreiller *et al.*, 2014; Viviroli *et al.*, 2007). The snowpack dynamics and its spatial extent also control many mountain processes, including soil erosion (Meusburger *et al.*, 2014), plant survival (Wipf *et al.*, 2009), and the glacier surface mass balance (López-Moreno *et al.*, 2016; Réveillet *et al.*, 2017; Sold *et al.*, 2013).

Some of the most dangerous natural hazards in mountain areas are also directly related to the distribution of the snowpack and ice, and their evolution over time. This is the case for snow avalanches (Schweizer *et al.*, 2008), and floods in mountain rivers and downstream areas (Gaál *et al.*, 2015). To enable anticipation of the occurrence of snow-related hazards and to reduce the threat to populations and infrastructure (Berghuijs *et al.*, 2016; Tacnet *et al.*, 2014); various models have been developed to reproduce and forecast the evolution of the snowpack on a daily or sub-daily basis.

Detailed snowpack models (Bartelt and Lehning, 2002; Vionnet *et al.*, 2012) are increasingly coupled with hydrological models to forecast river discharges, and this depends on reliable simulation of snow and ice melting (Avanzi *et al.*, 2016; Braun *et al.*, 1994; Lehning *et al.*, 2006). The more accurate the information on snowpack dynamics, the better will be the discharge forecasts based on hydrological models. However, the spatio-temporal distribution of the snowpack is highly variable in mountain areas (López-Moreno *et al.*, 2011, 2013; Scipión *et al.*, 2013; Seidel *et al.*, 2016), and the runoff from mountain catchments depends on many interrelated processes that are highly variable in space and time, including infiltration, surface runoff, groundwater recharge, freezing of soil, and the snowpack distribution (Seyfried and Wilcox, 1995). For example, in areas where snow persists throughout the year the snowpack dynamics has a major impact on groundwater storage (Hood and Hayashi, 2015). Finally, snowpack models are also combined with other models and techniques to forecast avalanche hazards (Bartelt and Lehning, 2002; Durand *et al.*, 1999).

Reproducing snowpack dynamics in heterogeneous mountain areas remains challenging. Some snowpack processes, including wind-induced redistribution and small scale topographic control on the snow distribution (Mott et al., 2010; Revuelto *et al.*, 2016a; Schirmer *et al.*, 2011; Trujillo *et al.*, 2007; Vionnet et al., 2014~~3~~) have not yet been fully integrated into numerical snowpack models which can be used operationally. Moreover, the additive nature of snowpack dynamics involves discrepancies between observed and simulated snowpacks, which can accumulate over the simulation period (e.g., Raleigh *et al.*, 2015).

The various approaches available for running snowpack simulations range from punctual simulations (snowpack dynamics simulated for a particular location having specific characteristics) to semi-distributed and distributed approaches that simulate snow dynamics over broad areas.

The semi-distributed approach, based on an unstructured grid design, involves simulating the snowpack evolution over areas defined using discrete values for topographic variables including altitude, aspect, and slope (Fiddes and Gruber, 2012, 2014); . The French numerical chain S2M (SAFRAN-SURFEX-MEPRA; Lafaysse *et al.*, 2013), simulates the snowpack evolution using a semi-distributed approach. In this chain the SURFEX/ISBA-Crocus snowpack model (Vionnet *et al.*, 2012; hereafter referred to as Crocus) is applied over a semi-distributed discretization of the French mountain ranges to diagnose the avalanche hazard for various topographic classes. Semi-distributed hydrological simulations are also widely used, which involves discretizing catchments into hydrologic response units (HRU), with the flow contribution from the HRUs being routed and compounded into an overall catchment discharge (Nester *et al.*, 2012; Pomeroy *et al.*, 2012). This simulation method is also applied to river discharge forecasting in mountain areas, with the output of semi-distributed snowpack simulations used as inputs to the hydrological models (Braun *et al.*, 1994).

The other modeling approach to simulating snowpack dynamics over extended areas is distributed simulations. This method involves simulation of the temporal evolution of environmental variables (e.g., snowpack or other hydrological variables) over a gridded representation of the terrain. In this approach the terrain is not discretized in classes; rather, it explicitly considers the characteristics (e.g. elevation, slope, aspect) for each pixel when simulating its snowpack evolution. Both approaches (distributed and semi-distributed) have advantages and disadvantages, particularly the lower computing resource requirements of semi-distributed simulations, and the more accurate terrain representation of distributed simulations. Some snowpack processes cannot be accurately reproduced using the semi-distributed approach, including wind-induced snow redistribution, small scale topographic control of precipitation, and terrain shadowing effects (Grünewald *et al.*, 2010; Revuelto *et al.*, 2014; Vionnet *et al.*, 2014). However, evaluating the performance of these simulation approaches depends on the intended use of the simulations (Carpenter and Georgakakos, 2006; Orth *et al.*, 2015). Similarly, the results obtained will depend on the spatial scale and the quality of the meteorological forcing model, and whether it is distributed or semi-distributed (Queno *et al.*; 2016; Vionnet *et al.*, 2016).

Many studies have compared the performance of hydrological models based on distributed and semi-distributed approaches in reproducing streamflow dynamics for alpine watersheds (Grusson *et al.*, 2015; Kling and Nachtnebel, 2009; Li *et al.*, 2015), but none have directly analyzed and compared representation of the spatio-temporal evolution of the snowpack using these simulation approaches. This is significant because direct implementation of the most promising advances in simulation ~~requires the use~~~~are~~is mainly considered for~~ of~~ distributed simulations. This is the case for assimilation of satellite data (Charrois *et al.*, 2016; Dumont *et al.*, 2012a; Thirel *et al.*, 2013); the inclusion of small scale processes in simulations, including snow redistribution by wind (Schirmer *et al.*, 2011; Vionnet *et al.*, 2014); and gravitational or topographic controls on snow movements (Bernhardt and Schulz, 2010; Christen *et al.*, 2010; Revuelto *et al.*, 2016a). Semi-distributed simulations may also allow the implementation of satellite data assimilation techniques (Mary *et al*, 2013) but would require specific routines for aggregating observations and they would reduce potential benefits of high resolution satellite observations. Similarly, blowing snow can be simulated in the semi-distributed approach (MacDonald *et al.*, 2009, Vionnet *et al.*, 2018). Vionnet *et al.*, (2018) show that strong assumptions on the topography are necessary to transport snow mass from one aspect to another (virtual ridge between opposite aspect classes for any elevation band). In MacDonald *et al.*, (2009), the model parametrization requires a discretization of the study site based on a strong knowledge of the area from previous works (McCartney *et al.*, 2006, Pomeroy *et al.*, 1999, 2006). Thus, the transferability of these results to large domains for which detailed information on the landscape features is not available is questionable.

~~Thus, comparison of distributed and semi-distributed simulations is needed to evaluate potential improvements, based on similar simulation setups (including the same study period and area, meteorological forcing, and simulation initialization). The newest meteorological models provide high spatial resolution information on the evolution of atmospheric variables (Seity et al., 2010); this is an improvement that distributed snowpack simulations can fully incorporate.~~

Recent studies have assessed the impact of high-resolution atmospheric forcing from the Numerical Weather Prediction system AROME (Seity *et al.*, 2010) on distributed snowpack simulations with Crocus. Queno *et al.*, (2016) and Vionnet *et al.*, (2016) compared simulations at a 2.5 km spatial resolution forced by AROME forecasts or by SAFRAN reanalysis (Durand *et al.*, 2009a). These works demonstrated that the geographical patterns simulated by the AROME-Crocus model chain are realistic and more detailed than the SAFRAN-Crocus model chain over large areas (the Pyrenees and French Alps). Nevertheless these studies also exhibit some significant biases in meteorological and snow variables with the AROME-Crocus chain which do not assimilate any meteorological observation, in particular precipitation. As a result, Queno *et al.*, (2016) and Vionnet *et al.*, (2016) exhibit a better skill of snowpack simulations when SAFRAN is used as forcing. They conclude that the potential of the high spatial resolution atmospheric forcing from the NWP system will be more beneficial in snowpack simulations with the development of a high-resolution distributed analysis combining observations and AROME forecast and the development of downscaling methods to fill the gap between their kilometric resolution and the resolution required to capture the variability of slopes and aspects in alpine environments.

Moreover the impact of topographic effects on snowpack simulations (implemented in the snowpack model) has not yet been assessed in detail. At present, the implementation of terrain shadowing effects on Crocus snowpack model (achieved in distributed simulations) has not been analyzed in complex alpine terrain. ~~This way, it~~It is therefore necessary to compare distributed and semi-distributed snowpack simulations with a spatial resolution that enables a detailed representation of alpine terrain. In this regard, ~~T~~this study provides a comprehensive evaluation of semi-distributed and distributed snowpack simulations for a mountain catchment, using the Crocus snowpack model (Brun *et al.*, 1992; Vionnet *et al.*, 2012)~~ over a long time period~~. ~~With the purpose of by one side using the longest time period with a suitable meteorological forcing in mountain areas and by the other side taking benefit from data assimilation of meteorological variables along the snow season, the SAFRAN re-analysis (Durand *et al.*, 2009a, 2009b) was selected as the meteorological forcing.~~ The SAFRAN re-analysis (Durand *et al.*, 2009a, 2009b) was selected as the meteorological forcing since it is available over a long period of time and assimilates meteorological observations over mountain areas.

The final products of both simulations are 250 m gridded snowpack ~~distribution maps~~datasets. This spatial resolution was selected because it renders~~ a sufficient~~ slopes sufficiently well to ~~representation to~~ describe small valleys with significant shadowing effects. It will also allow to explore~~ing~~ snow mechanical stability in future avalanche hazard forecasting applications. Indeed ~~(note the impact of~~ broader resolutions imply a too strong smoothing of terrain to represent~~ation)~~ slopes steep enough for avalanche release. The 250 m grid cell size of the simulations also enables a direct comparison with optical satellite products at the same spatial resolution. ~~Additionally, using SAFRAN as meteorological forcing allows to This study provided a comprehensive evaluation of semi-distributed and distributed snowpack simulations for a mountain catchment, using the Crocus snowpack model (Brun *et al.*, 1992; Vionnet *et al.*, 2012).~~

We firstly assessed the ability of the model to simulate the snowpack evolution at a local scale for specific stations having continuous snow observation data. For these stations, the punctual simulations accounted for local topographic characteristics. These punctual simulations enabled initial analysis of the capacity of the model to subsequently evaluate the distributed and semi-distributed approaches to simulating the snowpack dynamics over a broader area, using the same meteorological forcing. The simulation results obtained using the distributed and semi-distributed approaches were compared with observations for the snow covered area based on MODIS satellite sensors, the glacier surface mass balance (winter, summer, and annual), and the glacier equilibrium-line altitude derived from satellite images (Landsat, SPOT, and ASTER). This enabled assessment of the use of distributed simulations for analysis of snow and ice dynamics. The simulations were based on data for the upper Arve catchment (French Alps) for the 26 years from 1989 to 2015.

~~This way, t~~The SAFRAN-Crocus simulations shown in this work enable a complete evaluation of model performance in highly heterogeneous mountain terrain over a study period that captures all possible climatologies within the study area. Moreover the effect of using distributed simulations is compared with results obtained with semi-distributed simulations which nowadays are operationally exploited, showing the interest or not of changing the modelling app~~ora~~roach. Such evaluation also aims at~~pretends to~~ showing the *pros* and *con~~tra~~s* of simulating snowpack evolution over large mountain areas with different techniques that offer substantial flexibility on terms of their set-up and i~~o~~n terms of the computational requirements.

2. Study area

The upper Arve catchment is located in the western Alps, France, between the northeast slopes of the Mont Blanc massif and the southwest slopes of the Aiguilles Rouges massif. The catchment extends from the headwaters of the Arve River to the town of Chamonix (Fig. 1), and includes major tributaries carrying melt water from three glaciated areas (*Arveyron de la Mer de Glace, Arveyron d'Argentière, and Bisme du Tour*) to the main river. The upper Arve catchment covers 205 km^2 and has a high degree of topographic heterogeneity, with steep slopes in some areas, and gentle slopes on large glaciated areas and at the lower elevation zones of the valley, which is a typical U-shaped glacial valley. Elevation ranges from 1020 to 4225 m.a.s.l., with 65% of the surface area above 2000 m.a.s.l. Glaciers cover 33% of the area (Gardent *et al*., 2014), and 22% is covered by forests, mainly in the lower elevation areas. The water discharge regime is strongly dependent on the snow melt dynamics during spring and early summer, with the major contribution of melt water from glacierized areas occurring during late summer and autumn; this is termed a nivo-glacial regime of river discharge (Viani *et al*., submitted). The Mont Blanc and Aiguilles Rouges massifs are also highly spatially heterogeneous, having various slopes and aspects over a wide range of elevations in glaciated and non-glaciated areas; this affects the spatio-temporal evolution of snow and ice.

The area is one subject to severe flood hazards. This is a consequence of the steepness of the terrain, which results in a rapid hydrological response to precipitation, the typically rapid meteorological changes that occur in this mountain area (mainly associated with convective episodes during spring and summer), and the high population densities and infrastructure in the bottom of the valley.

3. Methods

3.1. Simulation setup

We used the Crocus snowpack model to simulate the temporal evolution of snow and ice in the upper Arve catchment. Crocus is a multilayer model that simulates snowpack evolution based on the energy and mass exchanges between the various snow layers within the snowpack, and between the snowpack and its interface with the atmosphere and the soil (i.e. the top and bottom of the snow column). The maximum number of layers in Crocus is set to 50. Crocus is implemented in the externalized surface model SURFEX (Vionnet *et al.*, 2012). Within SURFEX (Masson *et al.*, 2013), Crocus is coupled to the multilayer land surface model ISBA-DIF (Interaction between Soil, Biosphere and Atmosphere; diffusion version; Decharme *et al.*, 2011).

The meteorological forcing required to drive the temporal evolution of the simulations was obtained from the SAFRAN meteorological analysis system (Durand *et al.*, 1993). This provides the atmospheric variables needed to run ISBA-Crocus, including air temperature, specific humidity, long wave radiation, direct and diffuse short wave radiation, wind speed, and precipitation phase and rate. SAFRAN was specifically developed to provide meteorological forcing for mountain areas at a suitable elevational resolution. The SAFRAN analysis combines observational data obtained from automatic weather stations with manual observations with the guess from the global numerical weather prediction system ARPEGE (Courtier and Thépaut, 1994). We used SAFRAN re-analysis, which benefitted from meteorological observations not available in real time (Durand *et al.,* 2009a, 2009b). This analysis system can provide outputs for punctual simulations, or semi-distributed outputs. In the first case the analysis is performed directly for the elevations of the stations involved, while in the second case the analysis is performed for 300-m elevation bands. In both cases the spatial extent of the analysis is approximately 1000 km². These regions (known as “massifs”) were defined by Durand *et al*. (1993) who took climatic homogeneity into account. In this study the SAFRAN analysis was only used for that part of the Mont Blanc “massif” which covers the entire study catchment. Particularly, this “massif” has an extension of 580 km^2, so when taking into account the extension of the study area, it is covered a 36% of the “massif” extension- SAFRAN and SURFEX/ISBA-Crocus (hereafter SAFRAN-Crocus) are used in avalanche hazard forecasting in France, using the S2M chain (Lafaysse *et al.*, 2013); this takes account of the altitude, aspect, and slope classes (semi-distributed simulation).

3.2. Punctual, semi-distributed, and distributed approaches

The temporal evolution of snow and ice was simulated using punctual, semi-distributed, and distributed approaches, based on the same meteorological forcing. Despite ~~here in~~ ~~after~~ these three approaches being~~are~~ described individually and ~~also~~ their results ~~are~~ presented in different sub-sections, ~~we~~the reader must bear in mind that these are based o~~i~~n the same simulation setup.

Punctual simulation

Punctual snowpack simulations were performed for the five Météo-France stations within the study area, based on the elevation, slope, and aspect for each station. Punctual simulations included a topographic mask from a 50-m digital elevation model (DEM) to account for any terrain shadowing effect on simulation of the incoming shortwave radiation (provided by the SAFRAN meteorological model).

Semi-distributed simulation

Snow and ice semi-distributed simulations were carried out based on the topographic classes of the SAFRAN model (300-m elevation bands from 900 m.a.s.l. to 4100 m.a.s.l) for eight aspect classes (north, northeast, east, southeast, south, southwest, west, and northwest) and two slope values (20° and 40°). For each elevation band a simulation over flat terrain (no aspect classification) was also carried out. These topographic classes are the same as those used for avalanche forecasting (Lafaysse *et al.*, 2013). To consider snow and ice evolution on glacierized and non-glacierized areas, two distinct simulations were run for all terrain classes, one involving a given thickness of ice to initialize the simulation, and another initialized using bare ground (see section 3.3).

In a final stage the snowpack semi-distributed simulations (-which have an unstructured grid design) were assigned or re-projected onto the pixels of the study area DEM (the same DEM used for distributed simulations; 250x250 m grid size). The pixels were categorized according to the semi-distributed terrain classes: slopes from 0 to 10° were considered flat, those from 11 to 30° were assigned to the 20° slope class, and those > 30.1° were assigned to the 40° class. From this categorization of the DEM the snowpack simulation outputs were assigned to each terrain class for all time steps. Thereby, for each time step a snow and ice distribution map was generated that spatially distributed the semi-distributed snowpack simulation obtained for the various terrain classes. This enabled comparison of the two approaches based on the same observation dataset.

Distributed simulation

The distributed snowpack simulations were performed in a DEM having a 250 x 250 m grid spacing and covering the 205 km^2 of the study area. As SAFRAN reanalysis provides semi-distributed outputs, the meteorological forcing at hourly time steps was spatially distributed over the 250-m grid DEM using specific routines that accounted for the elevation and aspect ~~–~~of each grid cell.~~,~~ For each cell of the 250-m grid, the spatialization of meteorological variables from the 300-m elevation bands of SAFRAN is based on a linear interpolation between the two closest elevation bands. Only one SAFRAN aspect class is considered for each pixel (nearest-neighbour technique for the aspect) ~~-~~(Vionnet *et al.*, 2016). Therefore, the meteorological input data are similar for all simulations: only minor differences occur~~red~~ because elevation~~s~~ differences (< 300 m) may impact meteorological forcing variables.

The distributed Crocus simulations included the elevation, aspect, slope, soil, and land cover characteristics for each pixel (the last two obtained from ECOCLIMAP-II/Europe; Faroux *et al.*, 2013) to simulate the evolution of the snowpack (snow and ice). A routine to account for the topographic shadowing effect of short wave radiation (Revuelto *et al.*, 2016a) was included in the distributed simulations. The inclusion of particular pixel features and topographic shadowing is the main difference between the semi-distributed and distributed methods. Figure 2 shows a schematic representation of distributed and semi-distributed simulation approaches.

3.3. Simulation initialization

Snowpack simulations were run for the period 1989–2015. However, the ISBA ground state (including temperature and soil humidity) must be initialized to accurately reproduce the evolution of the snowpack. A spin-up simulation for the 1988–89 snow year (1 August 1988 to 31 July 1989) was repeated iteratively 10 times, to ensure a realistic ground state when launching simulations.

Similarly, to adequately replicate the snow and ice evolution over glacierized areas a glacier initialization was performed. Thus, for the simulations a sufficiently thick ice layer (several tens of meters) was incorporated beneath the snow layers to ensure glacier presence during each season in the glacierized areas. As Crocus is a multilayer snowpack model that simulates the energy and mass interchanges between the various snowpack layers, it also enables simulation of the glacier surface mass balance (Dumont *et al.*, 2012a; Gerbaux *et al.*, 2005; Lejeune *et al.*, 2013). Glacierized areas were initialized at the beginning of each snow season (1 August) using a 40-m ice thickness (if the total ice thickness was less than this value), which ensured that it was present for the entire snow season (from 1 August of one year to 31 July of the next year). Thus, the six deepest Crocus layers were initialized with a density value of 917 kg/m^3 and a temperature of 273.16 K (the Crocus default density and temperature values for ice, and representative of temperate glaciers). The thickness of these layers progressively transitioned from a shallow thickness for the upper layer (0.01 m) to thicker layers in the deepest part of the ice (with a 5-fold difference factor between one layer and the one above); this resulted in a total ice thickness of 39.06 m. The ice initialization was also performed during the spin-up of soil to reproduce the ground state over glaciarized areas. The extent of glacierized areas was based on the most recent data on their surface area, inventoried in 2012 (Rabatel *et al.*, 2013). Although other historic surface inventories of glacierized areas within the upper Arve catchment were available (1986 and 2003; Gardent *et al.*, 2014), the most recent inventory was used for simplicity because the change in the glacierized surface area between the inventoried dates represents less than a 1% of the total study surface area.

3.4 Evaluation strategy

The availability of direct snow and ice observations for mountain areas is limited. Broadly, when the time between observations is short, the spatial extent is limited and oppositely, when large areas are observed, the temporal frequency is low. Consequently, evaluation of the performance of a model in reproducing the snowpack evolution is difficult because of a lack of information. Although we did not evaluate a hydrological model in this study, the “observation scale” defined by Blöschl and Sivapalan (1995) aided assessment of the representativeness of the available observations. The observation scale is defined by: i) the spatial/temporal extent (coverage) of a dataset; ii) the spacing (space and time resolution) between samples; and iii) the integration volume (time) of a sample (also known as support). These three criteria can rarely be optimized simultaneously. Hanzer *et al.* (2016) introduced a representation to depict the suitability of an observation dataset to evaluate model performance. To evaluate the simulations in this study we used four datasets based on: *in situ* snow depth from Météo-France stations; the snow covered area (SCA) from MODIS images; the punctual glacier surface mass balance (SMB); and the glacier equilibrium-line altitude (ELA) from Landsat/SPOT/ASTER. Based on the radar charts presented by Hanzer *et al.* (2016), shown in their Figure 5, the information available for our study matches four of the datasets exploited in their study (Snow Depth, MODIS, Landsat and Glacier mass balance). These four datasets cover almost the full radar chart space ("optimal" validation dataset), thus providing almost a full evaluation of the simulation performance.

The analyses presented below enabled us to draw conclusions about the impact of the methods used on the various spatio-temporal scales considered, also enabling an overall evaluation of the simulation platform.

The four datasets ~~used in evaluation of the simulations are described below~~allowed performing a multi-criteria evaluation of the simulations with all observations available within the study area. However ~~N~~not all simulations (punctual, semi-distributed, and distributed) were evaluated using all four observation datasets. The punctual snow depth simulations only provided a preliminary evaluation of the simulation setup in terms of reproducing the temporal snowpack evolution, so only punctual snow depth observations were used in the evaluation of this simulation approach. The three other datasets (SCA, and glacier SMB and ELA) were used in evaluating the semi-distributed and distributed simulations, as these datasets had the appropriate spatial and temporal extents needed to assess the performance of these two approaches.

Punctual snow depth observations

The Météo-France observation network has 5 stations in the study area (Fig. 1), located at different elevations. Some of these stations acquired data during all snow seasons throughout the entire study period, including at Nivose Aiguilles Rouges (2365 m.a.s.l.), Chamonix (1025 m.a.s.l.), and Le Tour (1470 m.a.s.l.). Other stations were installed later, and provided observational data since the 1994–95 snow season (Lognan station; 1970 m.a.s.l) and since the 2003–04 snow season (La Flegere station; 1850 m.a.s.l.). At these stations the temporal evolution of the snow depth was observed at daily or sub-daily time intervals, and these data were used to evaluate SAFRAN-Crocus in non-glacierized areas during winter and spring (periods with snow presence).

Snow cover area based on the MODIS sensor

i) Evolution of the snow covered area

Many studies have demonstrated the usefulness of MODIS images for snow cover mapping in mountain areas (Gascoin *et al.*, 2015; Klein and Barnett, 2003; Parajka and Blöschl, 2008). The MODIS mission database provides long temporal coverage (the mission was launched in 2000, and obtains daily images), so enabled a comparison between the simulated and observed snow cover evolution for 14 snow seasons (out of the 26) simulated on an almost daily basis (comparisons were limited by cloud cover in the study area). Sub-pixel snow monitoring of the snow cover at 250-m spatial resolution was performed using MODImLab software (which is based on Dumont *et al.*, 2012b; Sirguey *et al.*, 2009~~)~~. Multispectral fusion between MOD02HKM (500 m; bands 3–7) and MOD02QKM (bands 1 and 2) (Sirguey *et al.*, 2008), enabled this software to generate images at 250 × 250 m spatial resolution to derive various snow–ice products. We used the unmixing_wholesnow (UWS, Sirguey 2016) product, as it has been shown to outperform other snow–ice products for assessing evolution of the SCA (spectral unmixing technique in Charrois *et al.*, 2013). We also considered the cloudiness product in MODImLab to determine the proportion of the catchment affected by cloud cover. Generation of the UWS and cloudiness products in MODImLab software was based on the same DEM used for the snowpack simulations. This ensured a direct match between of observation and simulation pixels. To avoid errors related to cloud presence in the study area, only days having cloud cover representing < 20% of the total surface area were considered in the analysis.

~~The~~ Different UWS thresholds for considering a pixel to be snow covered w~~as~~ere ~~se~~tested. The UWS values used in the sensibility test were 0.25, 0.35 and 0.45~~ to 0.35 (i.e., fractional snow cover > 35%; Charrois *et al.*, 2013; Dedieu *et al.*, 2016)~~. Similarly, ~~T~~three snow depth threshold values (0.10, 0.15, and 0.20 m (Gascoin *et al.*, 2015; Quéno *et al.*, 2016) were examined to consider a pixel as snow covered in the simulations on the sensibility test. Since snowpack simulations on forested areas (sub-canopy snowpack simulations not implemented in the simulations) and also satellite observations could have important deviations from real snowpack evolution; ~~T~~the SCA evolution in forested areas was not evaluated, and these areas were masked in the analysis.

The temporal evolution of the snow covered area (SCA) within the study area predicted by each simulation approach (semi-distributed and distributed) was analyzed in terms of the root mean squared error (RMSE), the mean absolute error (MAE), and R^2 for comparisons between simulations and observations. Despite the study period being ~~is~~ long in terms of snow observations, it only -spans over a 14 year time period. Thus, to assess whether the results obtained with distributed and semi-distributed simulations are significantly different, the uncertainty of the scores was quantified by a bootstrap approach.~~a t-student test has been applied.~~ From the annual SCA database, a ~~100~~ bootstrap sample of 100 members was obtained by random sampling with replacement of the different years of observations. This bootstrapping was ~~exploited~~used to compute the standard deviation of each score, considered as a random variable. -Thus, the scores samples of the semi-distributed and distributed simulations can be compared by a t-student test.

The temporal evolution of the SCA for specific snow seasons was also analyzed to assess the difference between observations and simulations in different time periods. Error metrics obtained on these particular snow seasons were compared to average values from the bootstrapped sample. The SCA evolution in forested areas was not evaluated, and these areas were masked in the analysis.

ii) Evaluation of spatial similarity

The spatial similarity between the observed and simulated SCA was evaluated for each simulation approach based on two similarity metrics: the Jaccard index (J), and the average symmetric surface distance (ASSD). As the grid cells coincided because the simulations and observations were based on the same DEM, we were able to obtain binary maps of snow presence from the simulated and observed maps, using the thresholds established.

The Jaccard index is the ratio of the intersection between the observed (O) and the simulated (S) SCA and the union of O and S (Equation 1). The index values range from 0 to 1, with a value of 1 representing a perfect match between the observed and simulated SCA.

$$J = \frac{|O \cap S|}{|O \cup S|} \qquad (1)$$

The ASSD is complementary to J, as it evaluates the distance between the boundaries of the observed and simulated SCA. ASSD is based in the modified directed Hausdroff distance between boundaries (Dubuisson and Jain, 1994; see Quéno *et al.*, 2016 and Sirguey *et al.*, 2009 for more details). The ASSD unit is meters, and the smaller the distance the better the match between surface boundaries. The Jaccard index and ASSD were calculated for the 2001–02 to the 2014–15 snow seasons. To assess the performance of the two SCA simulation approaches for specific periods, the 2006–07 and 2007–08 snow seasons (both of which were characterized by low average levels of snow accumulation) and the 2011–12 and 2012–13 snow seasons (characterized by high levels of snow accumulation) were analyzed for both the accumulation period (January, February, and March; JFM) and the melt period (May, June, and July; MJJ).

Glacier surface mass balance

Glaciers located in the Mer de Glace and Argentière sub-catchments have been monitored, in a sufficient number of measurement locations for our analysis, since 1995 by the French Service National d'Observation GLACIOCLIM. During this period field data were obtained twice per year, during the maximum (end April–May) and minimum (around October) snow accumulation periods. These data enabled calculation of the SMB for summer (SSMB; annual difference between the maximum and minimum acquisitions), winter (WSMB; annual difference between the minimum of the previous year and the maximum acquisitions), and annually (ASMB; year to year differences in the minimum acquisitions) at each individual point of the network (Fig. 3). The observation procedure involved use of glaciological methods (Cuffey and Paterson, 2010) to retrieve the surface mass balance for the various time periods (SSMB, WSMB, and ASMB). Stakes (markers over the glaciers) are set up in both accumulation and ablation areas throughout the glaciers, and so reflect the evolution of the various zones of the glaciers. The spatial distribution of the stakes is shown in Figure 3. For further information on the methods for SMB data collection, see Réveillet *et al.* (2017).

The observations of SMB for the various time periods at more than 65 locations encompassing different glaciers enabled assessment of the snow and ice evolution over glacierized areas, as these measurements included snow and ice ablation (SSMB) and snow accumulation (WSMB) periods. Thus, the simulated SMB for the same observation periods and locations were computed based on Crocus results. With this information~~, a linear regression and~~ RMSE, MAE and R^2 coefficient were computed for each sub-basin for the three periods, and these were used to measure the performance of the modeling approaches. Similarly to the SCA evaluation,– the significance of the differences between both simulation approaches is assessed with a bootstrap method based on the resampling of the 20 available years, in the observations.

Finally~~T~~ the simulated (distributed and semi-distributed) and observed temporal evolutions of the SMBs were compared based on the SAFRAN elevation bands (the average and standard deviation for all points within each band were calculated). To assess any elevational dependence of the SMB, the seasonal evolution of the observed and simulated SSMB, WSMB, and ASMB were compared for two snow seasons having opposite characteristics (high and low levels of snow accumulation) for the Mer de Glace glacier, which had a large gradient for assessing elevational dependence.

Glacier equilibrium-line altitude

The glacier equilibrium-line altitude (ELA) is the annual maximum elevation of the snow–ice transition over glacierized areas. Since 1984 the temporal evolution of the ELA for the five largest glaciers in the study area has been monitored using various satellite sensors (Rabatel *et al.*, 2013, 2016). Data on the inter-annual evolution of the ELA for the Tour, Argentière, and Mer de Glace glaciers (and its main tributaries, the Leschaux and Talèfre glaciers) was available for the entire study period

Images from Landsat 4TM, 5TM, 7 ETM+, SPOT 1–5, and ASTER were used to obtain the ELA for the study period. The spatial resolution of these images ranges from 2.5 to 30 m. The method of snow line delineation using multispectral images combining green, near-infrared, and short-wave infrared bands has been fully described by Rabatel *et al.* (2012). The satellite acquisition date depends on various factors including the availability of satellite images for the study area and cloud presence, but images obtained during the period of minimum snow accumulation (late August to early October) were used to obtain the ELA. Thus, the simulated ELA was obtained for the same dates as the satellite acquisitions. Because of the difference in the spatial resolution of the simulation (250 m) and satellite observations (≤ 30m), the average and standard deviations of the ELA were compared.

4. Results

4.1. Punctual snow depth

The observed and simulated snow depth evolution for the 2007–08 and 2012–13 snow seasons (low and high average snow accumulation years, respectively) for the five stations are shown in Figure 4. The snow depth evolution shows the capacity of the SAFRAN-Crocus model chain to reproduce the temporal evolution at locations having differing topographic characteristics.

It is important to note that the results shown in Figure 4 indicate the capacity of the simulations to reproduce snow depth dynamics at specific points having well known topographic characteristics. Punctual simulations include the impact of surrounding topography on incident solar radiation (terrain shadowing masks). Additionally, the meteorological forcing was taken at the station elevation (SAFRAN forcing not yet discretized on elevation bands). Nevertheless, the spatial scale of the meteorological forcing was that of the Mont Blanc SAFRAN massif. Therefore the spatial variability of solid/liquid precipitation within the massif is not taken into account.

Some snow accumulation events were underestimated or overestimated in the SAFRAN-Crocus simulation, evident in discrepancies between the simulated and observed snow depths, including for the Le Tour (overestimation) and La Flégère (underestimation) stations for the 2007–08 snow season. Despite these discrepancies resulting from meteorological forcing, the simulated evolution of the snow depth shows a good~~correct~~ temporal timing.~~appeared reliable, in particular during melt periods.~~

Table 1 shows the RMSE and bias errors between observations and simulations at the five stations. There was a high level of variability between the errors for the various stations, mainly because all local effects were not included in the simulations. It is noteworthy that the number of observations available and the time periods (which could have marked differences on total seasonal snow accumulation) affected the significance of the RMSE and bias for the various stations (Table 1). The RMSE values ranged from 20.8 to 66.6 cm and the bias ranged from –19.1 to 49.4 cm. These values are small relative to the total snowpack thickness (snow depth observations were commonly > 200 cm, and in some cases exceeded 300 cm). However, for the Aiguilles Rouges station the RMSE and bias estimates were higher than for the other stations. This may be because this station is exposed to major wind-induced snow transport episodes that were not accounted for in the simulation. In addition to these events, this station is also affected by forecasting errors related to the meteorological forcing, such as the large underestimation for the first snowfall in 2007–08.

4.2. Snow Cover Area evaluation

Figure 5 shows an example of the SCA obtained using the UWS product for 24 July 2008, and the corresponding simulated snow depth determined using the distributed approach. This date was selected because it was a cloud-free day with high elevation areas covered by snow.

Table 2 shows the SCA simulation results estimated based on 0.1, 0.15 and 0.2 m snow depth thresholds compared with the ~~observed~~ various UWS ~~(0.35~~ thresholds tested~~)s~~, for the 2008–09 and 2009–10 snow seasons (average snow accumulations) and for both spatialization approaches. This table shows that the evaluation metrics are only slightly sensitive to the choice of these thresholds and that for every threshold and metrics the ranking of the two approaches remains the same. In light of the sensibility test~~-se~~ results, we selected a 0.15 m snow depth ~~simulation~~ threshold for the simulations and 0.35 SCA threshold for MODImLab UWS product for classifying a pixel as snow-covered ~~deciding whether a pixel was snow covered~~.

i) Evolution of the snow covered area

The results of simulation of the SCA in the study area for 10 of the 14 snow seasons (for ease of visualization) based on MODIS data are shown in Figure 6. This figure shows that both approaches were able to reproduce the SCA evolution based on MODIS images. ~~During winter and early spring, when large areas of the catchment are covered with snow, there was a high degree of consistency between the observations, and simulations based on each approach. In contrast, during summer and early autumn, when snow is only present at high elevations and on preferential accumulation areas, there was less consistency between observations and simulations, particularly for the semi-distributed simulations.~~

Figure 7 shows the SCA evolution for four non-consecutive snow seasons, two having low levels of snow accumulation (2006–07 and 2007–08 seasons) and two having high levels of snow accumulation (2011–12 and 2012–13 seasons). In winter the simulation slightly overestimated the SCA compared with observations, but during summer and autumn the simulations underestimated the SCA. However, the distributed simulations most closely reproduced the observed SCA ~~(Table 3)~~. In all four seasons the semi-distributed simulations generated larger underestimates of the SCA during summer and early autumn.

Using the terrain aspect classification for semi-distributed simulations it is possible to evaluate the impact of terrain shadowing effects. From the eight orientation classes we identified two main groups: those having a northern aspect (N, NW, NE) and those having a southern aspect (S, SE, SW). Figure 8 shows the observed and simulated SCA evolution for high and low snow accumulation seasons in relation to these two terrain classes. The variability in the SCA was well captured for both aspects by both the semi-distributed and distributed simulations. ~~Moreover, ~~The SCA temporal evolution shows that overall the simulation underestimated the SCA, during late spring and summer in northern aspects. For southern aspects, simulation of the SCA evolution was poorer during winter.

Error estimates for the SCA simulated for the whole study site and for ~~in relation to the~~ north and south aspects (Tables 3, 4 and 5) were lower for the distributed simulations compared with the satellite observations. RMSE and MAE standard deviations obtained from the bootstrapping (Table 3) are lower than the difference between scores for both approaches. The p-values for these two error metrics are lower than 0.01 and thus the null hypothesis is rejected with a 99% confidence interval interval and the skills of distributed and semi-distributed simulations are not statistically equivalent Conversely the R2 standard deviation of the SCA is high compared to the difference between the scores of both approaches. As a result, the high p-value indicated (in this case above 0.05) that the null hypothesis should be accepted and that these scores are not statistically different between both approaches. R2, MAE and RMSE average values for high (2006-2008 snow seasons, Table 4) and low (2011-2013 snow seasons, Table 5) levels of snow accumulation also show the better capacity of distributed simulations to reproduce SCA evolution. T~~t~~he t-student test has demonstrated that RMSE and MAE results for both approaches are not statistically equivalent and that for all aspects~~sorientations~~ and periods, the distributed simulations presents lower errors~~; ~~. ~~w~~We can conclude that this latter approach significantly better reproduce the SCA evolution.

The differences in the error metrics (RMSE and MAE) between distributed and semi-distributed simulations are significant for both, north and south aspects but higher for north aspect. However, it must be highlighted that for the whole catchment and for any aspect, the null hypothesis can be accepted based on the R2 value between distributed and semi-distributed approaches. This means that the added value of the distributed approach is not visible on this criterion. ~~Moreover, the SCA temporal evolution shown in Figure 8 shows that overall the simulation underestimated the SCA, during late spring and summer in northern aspects. For southern aspects, simulation of the SCA evolution was poorer during winter.~~

ii) Evaluation of the spatial similarity

The spatial similarity between the observed and simulated SCA is exemplified in the temporal evolution of the Jaccard index and ASSD. Table 6 shows the average values for J and ASSD for the entire study period and for the 2006–07 and 2007–08 snow seasons (low levels of snow accumulation) and the 2011–12 and 2012–13 snow seasons (high levels of snow accumulation).

The higher scores found during seasons having high levels of snow accumulation were expected because of the larger areas covered by snow. Figure 9 shows the temporal evolution of the Jaccard index and ASSD for high and low level snow accumulation seasons. ~~Although the~~ D~~d~~ifference between the distributed and semi-distributed simulations was almost unappreciable~~low~~ for most dates ~~,~~and only ~~-~~during late melting (August-September) ~~the Jaccard index values for the~~ distributed simulations were slightly better~~higher~~ (higher J index and lower ASSD). This shows that except for some particular time periods differences in the spatial similarity with the observed SCA with both simulation approaches are minor., ~~showing a greater capacity for simulating the SCA (Table 6). Similarly, ASSD values were lower for distributed simulations, which showed reduced distances between the limits of snow free and snow covered areas. The differences between the two approaches are also evident in the average values shown in Table 6.~~

~~The performance of the simulations appeared to differ between periods of maximum and minimum snow accumulation (Fig. 9).~~ Table 7 shows the average Jaccard and ASSD index values obtained for the JFM and MJJ periods for the four snow seasons analyzed in detail (high and low level snow accumulation seasons). Again, a minor improvement ~~The better performance of distributed simulations was a result of better reproduction of the SCA evolution, and their ability to~~ on captureing ~~better~~the spatial patterns in heterogeneous mountain terrain is obtained with distributed snowpack simulations. Not surprisingly, the values in Table 7 also show higher scores for both simulations during winter and early spring, when the SCA was high.

4.3. Glacier surface mass balance

Analysis of the glacier surface mass balance enabled assessment of the effectiveness of simulations of the seasonal and annual evolution of snow and ice on glacier surfaces. Figures 10 and 11 show the simulated and observed temporal evolution of the surface mass balance for the 300-m elevation bands. These show good agreement between observations and simulations with respect to year-to-year SMB variability. During winter the snow accumulation at high elevations was underestimated. For elevations above 2700 m.a.s.l. a higher positive glacier SMB was observed, and the difference between the observed and simulated SMB increased at higher elevations. During summer, when solid precipitation has no or marginal influence in low elevation areas and little influence at higher elevations, the observed and simulated SMB values were similar for elevations above 2100 m.a.s.l. for the Mer de Glace glacier, and above 2400 m.a.s.l. for the Argentière glacier. Nevertheless, in high elevation areas the SSMB deviation was also underestimated on the simulations. ~~This was probably because of the lower level of snow accumulation simulated during winter (using SAFRAN model) which induces an earlier complete melting of snow in the simulation in low elevations. This is presumably because of more rapid melting of ice insulated from solar radiation by the snow layers above, and because of the impact of variations in wind speed or long wave radiation on the simulation.~~

Combination of the simulated WSMB and SSMB produced an ASMB that underestimated snow accumulation at high elevations (> 3000 m.a.s.l.) and melting at low elevations (2400 m.a.s.l. for the Argentière glacier, and < 2400 m.a.s.l. for the Mer de Glace glacier). Thus, the glacier ASMB included summer and winter variations, which in some cases negated each other. The contrasting performance of the simulations in reproducing the SMB between high and low elevations is clearly illustrated in Figure 12. This shows the altitudinal dependence of the SMB for two snow seasons, one having a low level of snow accumulation and the other a high level. The simulated SSMB, WSMB, and ASMB values for both approaches underestimated the observed values at both low (higher negative loss of water equivalents observed) and high (lower positive loss of water equivalents observed) elevation areas. Nevertheless, the SMB simulations at intermediate elevations correctly reproduce the observed values, and the temporal evolution of the SMB for the 20 years (Figs 10 and 11) was well reproduced by the simulations.

The performance of simulations in reproducing glacier SMB must take account of the areal extent at differing elevations. Elevations > 3000 m.a.s.l. represent 37% and 52% of the surface areas of the Argentière and Mer de Glace glaciers, respectively. The Argentière glacier has < 10% of its surface area below 2400 m.a.s.l., and the Mer de Glace glacier has < 7% below 2100 m.a.s.l. These relative extents of glacierized surface area show that for large areas of the glaciers the SMB was accurately reproduced by the simulations. However, for large glacierized areas there were marked differences between the observations and simulations; although the year-to-year evolution was accurately reproduced, this demonstrates the need to improve simulation methods.

In general, the distributed simulation values for the SMB were slightly closer to the observed SMB values than were those from the semi-distributed simulations.

Table 8 shows RMSE, MAE and R2 means and standard deviations obtained from the 100-member bootstrap sample. For most of the error metrics, the standard deviations are lower than score differences and the p-values are low enough to reject null hypothesis. This way results obtained with both simulation approaches are statistically different. In winter, the SMB simulations show similar results. For both glaciers, lower RMSE and MAE are obtained for distributed simulations and better R2 for semi-distributed simulations. Oppositely during summer all scores show better results for the distributed approach. The annual SMB also exhibit better results for distributed simulations. ~~shows that the RMSE values were lower for the distributed simulations and the R^2 values were higher for most periods in both glacierized areas. However, the WSMB simulations obtained using the semi-distributed approach were slightly better at reproducing the SMB.~~

4.4. Glacier Equilibrium Line Altitude

The temporal evolution of the ELA for the five largest glaciers in the study area is shown in Figure 13. ~~Overall, and d~~Despite differences in the spatial resolutions of simulations and observations of ELA, ~~the ability of the~~ simulations were able to capture the ~~temporal~~ evolution of the ELA changes during the 26 years of the study~~ was satisfactory, with lower variations found for distributed simulations for most seasons~~For most of the years and glaciers simulated, ELA values derived from the distributed approach were closer to those observed. However, for certain years, more precise results were obtained with semi-distributed simulations.

Table 9 shows the average absolute differences between observations and simulations and the linear adjustments for the five glaciers. These results show a systematic positive bias on the simulated ELA which is consistent with the summer underestimation revealed by the previous tests.

5. Discussion

5.1. Overview of SAFRAN-Crocus performance

The observation dataset used in this study enabled ~~multilevel~~ multi-criteria spatio-temporal validation of the performance of snowpack simulations at the scale of a large alpine catchment. The analysis of the results of semi-distributed and distributed simulations provided a holistic evaluation of the snow and ice dynamics in the study area. Overall, the SAFRAN-Crocus simulations have shown a good capability on reproducing the temporal evolution and spatial variability of snow and ice during the study period.

The simulations were evaluated using snow depth data from five Météo-France stations. Their ability to reproduce a bulk variable such as snow depth suggests that the main simulation processes were satisfactory, especially those related to the various components of the energy and mass balance. These findings are consistent with previous evaluations of the SAFRAN-Crocus system (Durand *et al.*, 2009a; Lafaysse *et al.*, 2013).

Crocus simulates the energy and mass exchanges with soil and atmosphere and also within the snowpack layers, but it does not simulate small scale topographic effects on snow depth distribution (Revuelto et al., 2016a). Given the fact that the final objective of this study is to compare two simulation approaches and that~~,~~ one of them~~,~~ would not allow an appropriate parametrization of topographic control on snow distribution (semi-distributed approach), we have not considered novel approaches for distributing snow based on terrain parameters (Cristera et al., 2017, Helbig et al., 2015), which may also require a higher spatial resolution for accounting topographic effect on snow distribution (Deems et al., 2006, Trujillo et al., 2007). Hence, we decided to simulate snowpack evolution ~~in~~with a spatial scale in which satellite observations were available over a long time period with a suitable temporal resolution, wh~~at~~ich lead to select 250m spatial resolution simulations (same as MODImLab products). Moreover this spatial resolution provides an –appropriate representation of slopes for future applications forecasting snow avalanches with expert systems (MEPRA, Lafaysse *et al.*, 2013).

Distributed information on the snowpack evolution from the MODIS sensor enabled evaluation of the simulation results on a suitable temporal scale. Although many MODIS images were discarded because of cloud cover, they demonstrated the capacity of SAFRAN-Crocus to simulate the spatial distribution of the SCA over time for large areas having high spatial heterogeneity. The 14-year time period spanned is longer than in all previous similar evaluations, and at a higher spatial resolution (Quéno *et al.*, 2016). Evaluation of the spatial similarity between simulations and observations (Jaccard index and ASSD) showed that the SCA spatial pattern was well reproduced. The simulated SCA for winter was in close agreement with observations, as most of the study area was covered by snow. In contrast, during summer the performance of simulations declined, as evidenced by the increase in ASSD and the decrease in the Jaccard index. As small scale topographic effects that control snow accumulation on preferential accumulation areas were not included in the simulations, deviations from observations would have increased for certain periods, particularly the late melt period. These processes, which are mainly driven by small topographic features, can be long-lasting during the late melt period (Revuelto *et al.*, 2016b; Sturm and Wagner, 2010). This was particularity evident in comparisons of the scores for the 2006–07 and 2007–08 periods with those for the 2011–12 and 2012–13 periods (Table 3). The differences in response may have originated from the higher weight of glacier melt processes in years with shallow snow depth. For these years, the good capability of the model on reproducing snow melting is lumped because the snow distribution is not appropriately simulated.

The availability of observations of the glacier SMB over a long time period provided an opportunity to evaluate the performance of the simulations in capturing the snow and ice temporal evolution over a wide range of elevations over glacierized areas. Contrasting simulation performances were found in the various elevation bands, and changed with the time period involved (summer, winter, or annual scales). The performances in simulating the SMB for the Argentière and Mer de Glace glaciers differed at high and low elevations. Although the observed SMB was always higher than the simulated one for elevations exceeding 2700 m, the opposite was observed for areas having elevations below 2100–2400 m. As the temporal variability of solid precipitation generally explains the temporal variability of the WSMB (Réveillet *et al.*, 2017), it is important to consider differences between simulated and observed solid precipitation, and how these could affect underestimation of the SMB in simulations. Studies in the same study area and nearby glaciers suggest that at high elevations the SAFRAN reanalysis may underestimate solid precipitation at ratios ranging from 1:1.2 at 2000 m.a.s.l. and 1:2.0 at 3200 m.a.s.l., with an average of 1:1.5 at the glacier scale (

Gerbaux *et al.*, 2005; Réveillet *et al.*, 2017; Viani *et al.*, submitted). This mainly results from the lack of precipitation observations at high elevations available for assimilation into the SAFRAN reanalysis; consequently divergences increase with elevation. Despite this shortcoming, the simulations captured the inter-annual fluctuation of the WSMB for all elevation bands. During summer the SMB could be explained by temperature variability in the two glaciers (Réveillet *et al.*, 2017), thus simulations results are closer to observations, particularly at higher elevations. In summer, most precipitation is liquid, and so has little impact on the energy balance of the glaciers (Hock, 2005); this may explain the improvement in summer simulations for most elevations.

It has recently been shown that Crocus is able to accurately simulate snow albedo (Réveillet *et al.*, in prep), which is important because of its influence on the surface mass balance (Essery and Etchevers, 2004; Essery *et al.*, 1999). However, it has been demonstrated that Crocus results are directly affected by uncertainties in the estimation of long wave radiation and wind (Réveillet *et al.*, in prep). Such effects may be significant for elevations where the snow completely melts during summer and do not insulate ice from the atmosphere during late melt season; this includes the low elevation areas of glaciers, where high SSMB errors were found. At the annual time scale, glacier differences between the observed and simulated SMB at high elevations during winter and at low elevations during summer were reduced because the SMB underestimates for winter (note these were negative/positive at high/low elevations) were compensated for by more accurate simulations during summer, and vice versa. Regardless of these errors, SAFRAN-Crocus was able to replicate the interannual evolution of the SMB. Additionally, there was a good match between observations and simulations for the 2100–2400 to 3000 m.a.s.l. elevation bands for the Mer de Glace and Argentière glaciers, respectively; these elevation bands encompassed large proportions of the glaciers (approximately 40 and 53%, respectively).

For the entire study period the SAFRAN-Crocus simulations effectively reproduced the observed inter-annual evolution of the study area glacier ELA. However, some differences were evident, particularly on steeper glaciers, because the high spatial heterogeneity was not well captured by the simulations. For mid-latitude mountain glaciers, the annual evolution of the ELA can be considered to be a good proxy for the glacier surface mass balance (Braithwaite, 1984; Rabatel *et al.*, 2005). Thus, observations of the glacier SMB, together with the ELA, provide for a complete evaluation of glacier temporal evolution.

5.2~~3~~. Distributed *vs.* semi-distributed approaches

In this study we performed distributed and semi-distributed snowpack simulations using the same model and evaluation setup (including ice initialization, meteorological forcing, projection on the same grid, observation databases). Thus, both approaches were affected by the same methodological limitations. The simulation results were consistent with the observed SCA evolution using both approaches. However, better results were obtained from the distributed simulations~~,~~ ~~especially~~ during late summer. Similarly, ~~This was because Tthe energy balance was more accurately simulated in the distributed approach, as it accounted for terrain shadowing effects on incoming solar radiation. The distributed simulations also accounted for the specific characteristics of each pixel rather than categorization based on topographic classes.~~

~~The distributedBoth approaches also produced more accuratesimilar SCA simulations of the SCA for the various time periods, except particularly during the late melt period, when deviations from observed values were higher with semi-distributed simulations. Similarly, spatial similarity evaluation (Jaccard index and ASSD) also showed that the distributed approach was slightly superior at reproducing the SCA distribution. The semi-distributed approach better simulated the temporal evolution of the SCA for areas having a southern aspect, because of terrain shadowing effects in areas having a northern aspect are not appropriately considered. Oppositely, tit iswas also observed~~ an improvement on simulation results ~~with~~was observed with ~~-~~distributed simulations~~-he simulation in northern aspects obtained with the distributed approach is superior because theseis approach is are able to include terrain shadowing on the simulations~~. This ~~was~~is because the energy balance ~~was~~is more accurately simulated in the distributed approach, as it account~~ed~~s for terrain shadowing effects on incoming solar radiation. Thereby, for ~~aspects~~areas (deep valleys) ~~(northern aspect)~~ and/or time periods (~~mainly~~ long time periods since last snowfalls on late summer) for ~~in~~ which the differences on simulation of the incoming solar radiation has a determinant weight, differences on snowpack simulation between both approaches are marked.~~.~~

Based on the glacier SMB scores and their temporal evolution, we concluded that the best simulation approach depends on the season involved. Thus, the WSMB evaluation showed that similar results were obtained using the two methods. In contrast, the distributed approach was better at simulating the SSMB. The similar performances of the semi-distributed and distributed simulations during winter, but the better results for the distributed simulations for summer resulted in the distributed approach providing greater accuracy at the annual scale. The rather better results obtained for both glaciers analyzed for a long time period (ASMB) using the distributed simulations suggests that this approach is likely to provide more reliable results over longer periods.

The distributed simulation of the ELA generally showed closest agreement with observations, but for certain years the semi-distributed simulations most accurately reproduced the observed values. Thus, it is not possible to conclude that one approach to reproducing the ELA was superior. This uncertainty may be related to the coarse pixel size, which did not enable the high spatial heterogeneity of the terrain to be captured. The annual ELA covers a small area of the glaciers (it represents the snow line limit between snow-free and snow-covered areas), and thus the effect of spatial heterogeneity is likely to be significant.

Overall, the distributed simulations were slightly better at reproducing observational data. Thus, distributed simulations, which better represent the spatial heterogeneity of mountain areas, in general produce more accurate snowpack simulations~~, and are the recommended modeling approach~~. However, depending on the purpose of the simulations and the accuracy required, other factors must be considered. For instance, semi-distributed simulations have lower computing resource requirements; in this study, the distributed approach had computing requirements that were a factor of 100 greater. The accuracy of semi-distributed simulations in reproducing the snowpack evolution over large areas makes them useful in many applications.

A good example of an application in which the computational requirements have a determinant weight are ensemble simulations for projections in several climate scenarios (e.g. Verfaille et al, 2017).

5.~~2~~3. Limitations ~~of the evaluations performedand simulations~~

Although the observation dataset enabled comprehensive evaluation of the simulations, it had limitations. First, the discrepancy in spatial scale between the SAFRAN meteorological analysis and the snow depth observations, and the low number of stations, limited the interpretation of results in terms of the simulated snow depth. Differences in the temporal evolution of snow depth between observation and simulations were in part associated with the unresolved sub-massif spatial variability in the level of precipitation, as previously described (Durand *et al.*, 2009a; Lafaysse *et al.*, 2013; Vionnet *et al.*, 2016). *In situ* observations are also subject to local effects associated with the topographic control at each site, including exposure to dominant winds, which markedly affects the snow depth dynamics. Such effects remain difficult to capture in snowpack modeling (Dadic *et al.*, 2010a; Liston *et al.*, 2007; Revuelto *et al.*, 2016a; Schirmer *et al.*, 2011; Vionnet *et al.*, 2014), and were not included in the modeling involved in our study. Similarly other processes such a~~u~~s lateral heat flux exchanges amongst grid-cells are not implemented in Crocus snowpack model and thus could impact the final result of simulations (Harder and Pomeroy 2017).

Discrepancies originating from the snow–rain limit can also influence the snow depth. Stations at high elevation (Aiguilles Rouges: 2365 m.a.s.l.) are typically not affected by this phenomenon during winter, as the 0°C isotherm is located at lower elevations. In contrast, low elevation stations (Le Tour: 1470 m.a.s.l.; Chamonix: 1025 m.a.s.l.) are potentially affected by differences between the simulated and observed snow–rain limit, even during winter. In mid-latitude regions including the Alps, elevational shifts in the 0°C isotherm cover a significant variation throughout the year, including the elevations where each of the stations in this study is located.

~~data~~ Data on the spatial extent of SCA derived from MODIS images enabled distributed evaluation of the simulations. However, its usefulness in analysis of the performance of spatial simulations is limited, as it does not provide information on other snowpack variables, and imposes restrictions on the spatial resolution. Satellite observations also involve uncertainty, depending on the routines applied for generating the final product and the thresholds used to decide whether a pixel area as covered by snow. After a sensibility test ~~W~~we adopted a 0.35 UWS threshold for considering a pixelas snow covered in satellite imagery (Charrois *et al.*, 2013; Dedieu *et al.*, 2016). We also performed an analysis to select the simulated snow depth threshold for considering a pixel to be snow covered. The 0.15 m threshold selected is consistent with values reported in previous studies (Gascoin *et al.*, 2015; Quéno *et al.*, 2016). Despite mountain areas having~~ve~~ a high spatial heterogeneity which also affects snowpack distribution, these thresholds enabled a binary representation of snow presence/absence which finally ensured a consistent SCA evaluation of both simulation approaches.

The results obtained in this study, i.e. slightly but significantly better skill for the distributed approach, are sensitive to the choice of the spatial resolution. Using resolution coarser than 250 m would lead to smaller differences between both spatialization approaches because the pixel elevations would be less accurate and because all the shadows would not be resolved. Conversely, higher resolutions may improve the accuracy of shadowing effects but with computational times which can become unaffordable for large areas applications.

.~~-~~Moreover, ~~I~~in low elevation areas, where ice is exposed to the atmosphere for longer periods during the year (snow does not insulate ice from the atmosphere since it has disappeared), differences in meteorological forcing variables including wind and temperature can have a marked influence on simulation results (Réveillet *et al.*, submitted). Similarly, at low elevations the glaciers are usually covered by debris, as is the case for the Mer de Glace glacier. This was not considered in our simulations, but differences in the behavior of the snow–ice interface in debris-covered areas could be expected to affect the simulation results (Lejeune *et al.*, 2013).

~~In addition to the above issues, satellite products can have errors for specific dates. For a small number of days during the study period the SCA obtained from MODIS images did not describe the real extent of snow cover. For these days the SCA did not match the temporal SCA evolution observed on previous and later dates. Furthermore, days having the maximum cloud cover allowed in our analysis could have ± 20% SCA variability. This induces uncertainty in the observation for certain dates which can be greater than this of the pixel classification as snow covered in the simulations (note the ± 0.05 m snow depths threshold tested). In addition, pixels classified as snow covered in which bare soil may have a non-negligible extension (pixels close to the 0.35 UWS threshold) could introduce discrepancies between observations and simulations, mainly during summer.~~

~~Glacier surface mass balance observations also involve limitations. For instance, infrequent glacier SMB observations for certain temporal windows limited evaluation of the simulated SMB. The spatial sampling involved in the glaciological method can also be a significant source of uncertainty, especially for elevation bands for which there are a limited number of observations. Additionally, the average SMB obtained for the elevation bands can lump the high SMB spatial variability that occurs within a specific band. For most years and all the elevation bands the uncertainty associated with the~~

~~average SMB measurements (± 0.2 m water equivalent; Réveillet *et al.*, 2017) was exceeded by the uncertainty associated with the observations for each band. This could have affected the results presented here, indicating that the standard deviations for the observed SMB values should be retained when analyzing the results of the simulations.~~

~~The simulations underestimated the observed SMB for the lowest elevations having SMB observations, despite the temporal variability being replicated. This may have been related to errors in precipitation and phase, and in this regard differences in the snow–rain limit could be important. Additionally, the impact of local effects is more important at low elevations, as glaciers are more confined in valleys that have very steep slopes and adjacent high mountains. In low elevation areas, where ice is exposed to the atmosphere for longer periods during the year (snow does not insulate ice from the atmosphere since it has disappeared), differences in meteorological forcing variables including wind and temperature can have a marked influence on simulation results (Réveillet *et al.*, submitted). Similarly, at low elevations the glaciers are usually covered by debris, as is the case for the Mer de Glace glacier. This was not considered in our simulations, but differences in the behavior of the snow–ice interface in debris-covered areas could be expected to affect the simulation results (Lejeune *et al.*, 2013).~~

~~Some issues were also evident in evaluation of the ELA. For the smallest glaciers, a reduced number of pixels having the 250 m pixel resolution were considered. As the ELA observations were based on Landsat, SPOT and ASTER satellite images (2.5–30 m resolution) the spatial variability of the simulation made it difficult to identify the glacier margins. The combination of problems in delimitating glaciated areas over smaller ice bodies, and the smooth topography characterizing the simulations compared with real terrain, could cause simulation errors for smaller glaciers.~~

~~5.3. Distributed *vs.* semi-distributed approaches~~

~~In this study we performed distributed and semi-distributed snowpack simulations using the same model and evaluation setup (including ice initialization, meteorological forcing, projection on the same grid, observation databases). Thus, both approaches were affected by the same methodological limitations. The simulation results were consistent with the observed SCA evolution using both approaches. However, better results were obtained from the distributed simulations, especially during late summer. The energy balance was more accurately simulated in the distributed approach, as it accounted for terrain shadowing effects on incoming solar radiation. The distributed~~

~~simulations also accounted for the specific characteristics of each pixel rather than categorization based on topographic classes. The distributed approach also produced more accurate simulations of the SCA for the various time periods, particularly during the late melt period. Similarly, spatial similarity evaluation (Jaccard index and ASSD) also showed that the distributed approach was slightly superior at reproducing the SCA distribution. The semi-distributed approach better simulated the temporal evolution of the SCA for areas having a southern aspect, because of terrain shadowing effects in areas having a northern aspect are not appropriately considered. Oppositely, the simulation in northern aspects obtained with the distributed approach is superior because these are able to include terrain shadowing on the simulations.~~

~~Based on the glacier SMB scores and their temporal evolution, we concluded that the best simulation approach depends on the season involved. Thus, the WSMB evaluation showed that similar results were obtained using the two methods. In contrast, the distributed approach was better at simulating the SSMB. The similar performances of the semi-distributed and distributed simulations during winter, but the better results for the distributed simulations for summer resulted in the distributed approach providing greater accuracy at the annual scale. The better results obtained for both glaciers analyzed for a long time period (ASMB) using the distributed simulations suggests that this approach is likely to provide more reliable results over longer periods.~~

~~The distributed simulation of the ELA generally showed closest agreement with observations, but for certain years the semi-distributed simulations most accurately reproduced the observed values. Thus, it is not possible to conclude that one approach to reproducing the ELA was superior. This uncertainty may be related to the coarse pixel size, which did not enable the high spatial heterogeneity of the terrain to be captured. The annual ELA covers a small area of the glaciers (it represents the snow line limit between snow-free and snow-covered areas), and thus the effect of spatial heterogeneity is likely to be significant.~~

~~Overall, the distributed simulations were better at reproducing observational data. Thus, distributed simulations, which better represent the spatial heterogeneity of mountain areas, in general produce more accurate snowpack simulations, and are the recommended modeling approach. However, depending on the purpose of the simulations and the accuracy required, other factors must be considered. For instance, semi-distributed simulations have lower computing resource requirements; in this study,~~

~~the distributed approach had computing requirements that were a factor of 100 greater. The accuracy of semi-distributed simulations in reproducing the snowpack evolution over large areas makes them useful in many applications.~~

5.4. Future perspectives on distributed snowpack simulations

Simulating the snowpack evolution in mountain areas is challenging. Although advances in meteorological/snowpack models and simulation approaches are improving the reproduction of observational data, inaccuracies remain. Many studies have highlighted the potential to improve snowpack modeling by assimilating observational data (Griessinger *et al.*, 2016; Thirel *et al.*, 2013). Satellite data enables the distribution of the snowpack over large areas to be determined, and the assimilation of such data into snowpack models has been shown to significantly improve the simulation results (Charrois *et al.*, 2016). In distributed snowpack simulations almost direct satellite data can be assimilated, in contrast to the semi-distributed approach which needs of aggregation routines to enable satellite data assimilation losing part of the information in this process. Additionally, meteorological forcing models having high spatial resolution are improving simulations of the spatial pattern of meteorological variables in mountain areas (Schirmer and Jamieson, 2015; Vionnet *et al.*, 2016; Weusthoff *et al.*, 2010). This will improve snowpack simulations (Förster *et al.*, 2014; Quéno *et al.*, 2016), even though it is challenging to combine high resolution numerical weather prediction models with precipitation measurements assimilation in analysis systems. Interest in distributed snowpack simulations will be enhanced when reliable high spatial resolution meteorological forcing data are available, as only this simulation approach can take full advantage of such data.

Other approaches halfway between our distributed and semi-distributed snowpack simulations are also showing promising results. This is the case of unstructured triangular meshes, which allow better capturing horizon- shadows of surrounding topography than the semi-distributed approach used in this work. These methods are able to improving energy balance simulation results while preserving computational costs (Marsh et al., 2012);

Further research is needed on parameterizing small scale snowpack processes for incorporation in modeling, including wind driven snow transport (Dadic *et al.*, 2010b; Winstral *et al.*, 2012), avalanche snow redistribution (Bernhardt and Schulz, 2010), and topographic control on snow distribution (Revuelto *et al.*,. 2016a). Inclusion of these processes, together with the incorporation of reliable meteorological forcing and satellite data, assimilation will improve the accuracy of snowpack simulations over extensive mountain areas.

6. Conclusions

This study provided a detailed assessment of the ability of the SAFRAN-Crocus system to simulate the snow and ice dynamics in complex alpine terrain using distributed and semi-distributed simulation approaches. The study was undertaken in the upper Arve catchment in the western French Alps, with simulations run for the 1989–90 to the 2014–15 snow seasons.

A preliminary evaluation of the simulations was completed based on observations of snow depth derived from five meteorological stations within the study area. This was only performed using punctual snowpack simulations, to provide an initial assessment of model performance over non-glaciated terrain. Despite some discrepancies between observations and simulations, the model reliably reproduced the snow depth, especially during melt periods.

In regard to the spatial scale of snowpack simulations over extended areas, the semi-distributed and distributed simulations were compared using the same observation datasets, including: (i) the temporal evolution of the snow-covered area based on data from the MODIS sensor; (ii) measurements of surface mass balance of glaciers within the upper Arve catchment; and (iii) observational data on the annual evolution of the equilibrium-line altitude for the various glaciers considered.

Both simulation methods accurately reproduced the evolution of the SCA during accumulation events, as they relied on the same meteorological forcing data. For the winter to early spring period, when the study area is almost completely covered by snow, there was little difference between the two approaches. However, for ~~the~~ late melt period the distributed simulations better reproduced the observations.

The simulations for low elevations and elevations > 2700 m.a.s.l. underestimated (negative underestimation in low elevations and positive in high) the observed SMB. Nevertheless, the results of both simulations were in close agreement with observations at mid-elevation areas, and adequately reproduced the observed annual SMB at all elevations. Overall, the distributed simulations yielded limited better results.

Based on comparison with ELA data obtained from various satellites at the end of summer, the SAFRAN-Crocus accurately reproduced the inter-annual variability of the snowpack over glaciated areas. However, differences between observations and simulations were evident, particularly for the smallest glacierized areas, where the spatial resolution of the simulations did not enable the high spatial variability of the topography to be included. ~~In addition, based on the ELA evaluation, the distributed approach was slightly better at reproducing the snowpack dynamics.~~

Overall, the results of this study demonstrated that distributed simulations reproduce slightly better snowpack dynamics in the alpine terrain of our study area. Distributed simulations take into account the specific topographic characteristics of each pixel (local values of aspect, slope and elevation) and more importantly the effects of terrain shadowing by surrounding areas. Accounting for these two effects over long time periods led to statistically significant better results for the distributed approach. However the lower computational requirements of semi-distributed simulations together with the flexibility on the design and application scale of the simulation make this approach also suitable to simulate snowpack evolution.~~Overall, the results of this study demonstrated that distributed simulations were slightly better at reproducing snowpack dynamics in the alpine terrain of our study area. Distributed simulations take account of the specific topographic characteristics of each pixel and also more importantly the effects of terrain shadowing by surrounding areas. Inclusion of these two effects over long time periods led to better results being obtained using the distributed approach. Distributed simulations will facilitate incorporation of the latest snowpack modeling advances, including assimilation of satellite data and the use of higher spatial resolution meteorological forcing models~~.

7. Acknowledgments

This study was funded by Syndicat mixte d'aménagement de l'Arve et de ses abords (SM3A), Communauté de Communes de la Vallée de Chamonix Mont-Blanc and Fondation Terre Solidaire in the framework of the Programme d'Action de Prévention des Inondations (PAPI). We thank Glacioclim (https://glacioclim.osug.fr) for generating the glacier surface mass balance database used in the study. J. Revuelto benefited from a grant within the above-cited PAPI project and is now supported by a Post-doctoral Fellowship of the AXA research found (le Post-Doctorant Jesús Revuelto est bénéficiaire d'une bourse postdoctorale du Fonds AXA pour la Recherchem Ref: CNRM 3.2.01/17). IGE and CNRM/CEN are part of Labex OSUG@2020.

References

Avanzi, F., De Michele, C., Morin, S., Carmagnola, C.M., Ghezzi, A., and Lejeune, Y. (2016). Model complexity and data requirements in snow hydrology: seeking a balance in practical applications. Hydrol. Process. *30,* 2106–2118.

Bartelt, P., and Lehning, M. (2002). A physical SNOWPACK model for the Swiss avalanche warning: Part I: numerical model. Cold Reg. Sci. Technol. *35*, 123–145.

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

Bartelt, P., and Lehning, M. (2002). A physical SNOWPACK model for the Swiss avalanche warning: Part I: numerical model. Cold Reg. Sci. Technol. *35*, 123–145.

Berghuijs, W.R., Woods, R.A., Hutton, C.J., and Sivapalan, M. (2016). Dominant flood generating mechanisms across the United States. Geophys. Res. Lett. *43*, 2016GL068070.

Bernhardt, M., and Schulz, K. (2010). SnowSlide: A simple routine for calculating gravitational snow transport. Geophys. Res. Lett. 37, L11502.

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

Braithwaite, R.J. (1984). Short Notes: Can the Mass Balance of a Glacier be Estimated from its Equilibrium-Line Altitude? J. Glaciol. *30,* 364–368.

Braun, L.N., Brun, E., Durand, Y., Martin, E., and Tourasse, P. (1994). Simulation of discharge using different methods of meteorological data distibution, basin discretization and snow modelling. Nord. Hydrol. 25, 129–144.

Brun, E., David, P., Sudul, M., and Brunot, G. (1992). A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting. J. Glaciol. *38*, 13–22.

Carpenter, T.M., and Georgakakos, K.P. (2006). Intercomparison of lumped versus distributed hydrologic model ensemble simulations on operational forecast scales. J. Hydrol. *329*, 174–185.

Charrois, L., Dumont, M., Sirguey, P., Morin, S., Lafaysse, M., and Karbou, F. (2013). Comparing different MODIS snow products with distributed distributed simulation of the snowpack in the French Alps. Proceedings of the International Snow Science Workshop Grenoble – Chamonix Mont-Blanc - 2013 (Grenoble, France), 937-941

Charrois, L., Cosme, E., Dumont, M., Lafaysse, M., Morin, S., Libois, Q., and Picard, G. (2016). On the assimilation of optical reflectances and snow depth observations into a detailed snowpack model. The Cryosphere *10*, 1021–1038.

Christen, M., Kowalski, J., and Bartelt, P. (2010). RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. Cold Reg. Sci. Technol. 63, 1–14.

Courtier P, Thépaut J-N, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var using an incremental approach. Q. J. R. Meteorol. Soc. 120 1367–1388

Cristea, N. C., Breckheimer, I., Raleigh, M. S., HilleRisLambers, J., and Lundquist, J. D. (2017). An evaluation of terrainâ˘AR˘ based downscaling of fractional snow covered area datasets based on Lidar-derived snow data and orthoimagery. Water Resources Research, 53, 6802–6820,

Cuffey, K.M., and Paterson, W.S.B. (2010). The Physics of Glaciers (Academic Press Inc, Amsterdam (NL)).

Dadic, R., Mott, R., Lehning, M., and Burlando, P. (2010a). Wind influence on snow depth distribution and accumulation over glaciers. J. Geophys. Res. Earth Surf. *115*, F01012.

Dadic, R., Mott, R., Lehning, M., and Burlando, P. (2010b). Parameterization for wind–induced preferential deposition of snow. Hydrol. Process. *24*, 1994–2006.

DeBeer, C.M., and Pomeroy, J.W. (2009). Modelling snow melt and snowcover depletion in a small alpine cirque, Canadian Rocky Mountains. Hydrol. Process. 23, 2584–2599.

Decharme, B., Boone, A., Delire, C., and Noilhan, J. (2011). Local evaluation of the Interaction between Soil Biosphere Atmosphere soil multilayer diffusion scheme using four pedotransfer functions. J. Geophys. Res. Atmospheres 116, D20126.

Dedieu, J.-P., Carlson, B.Z., Bigot, S., Sirguey, P., Vionnet, V., and Choler, P. (2016). On the Importance of High-Resolution Time Series of Optical Imagery for Quantifying the Effects of Snow Cover Duration on Alpine Plant Habitat. Remote Sens. *8*, 481.

Dubuisson, M.P., and Jain, A.K. (1994). A modified Hausdorff Distance for Object Matching. Proc. Int. Conf. Pattern Recognit. Jerus. Isr. 566-568.

Dumont, M., Durand, Y., Arnaud, Y., and Six, D. (2012a). Variational assimilation of albedo in a snowpack model and reconstruction of the spatial mass-balance distribution of an alpine glacier. J. Glaciol. *58*, 151–164.

Dumont, M., Gardelle, J., Sirguey, P., Guillot, A., Six, D., Rabatel, A., and Arnaud, Y. (2012b). Linking glacier annual mass balance and glacier albedo retrieved from MODIS data. The Cryosphere *6*, 1527–1539.

Durand, Y., Brun, E., Mèrindol, L., Guyomarc'h, G., Lesaffre, B., and Martin, E. (1993). A meteorological estimation of relevant parameters for snow models. Ann. Glaciol. *18*, 65–71.

Durand, Y., Giraud, G., Brun, E., Merindol, L., and Martin, E. (1999). A computer-based system simulating snowpack structures as a tool for regional avalanche forecasting. J. Glaciol. *45*, 469–484.

Durand, Y., Laternser, M., Giraud, G., Etchevers, P., Lesaffre, B., and Mérindol, L. (2009a). Reanalysis of 44 Yr of Climate in the French Alps (1958–2002): Methodology, Model Validation, Climatology, and Trends for Air Temperature and Precipitation. J. Appl. Meteorol. Climatol. *48*, 429–449.

Durand, Y., Giraud, G., Laternser, M., Etchevers, P., Mérindol, L., and Lesaffre, B. (2009b). Reanalysis of 47 Years of Climate in the French Alps (1958–2005): Climatology and Trends for Snow Cover. J. Appl. Meteorol. Climatol. 48, 2487–2512.

Egli, L., and Jonas, T. (2009). Hysteretic dynamics of seasonal snow depth distribution in the Swiss Alps. Geophys. Res. Lett. *36*, L02501.

Essery, R., and Etchevers, P. (2004). Parameter sensitivity in simulations of snowmelt. J. Geophys. Res. Atmospheres *109*, D20111.

Essery, R., Martin, E., Douville, H., Fernández, A., and Brun, E. (1999). A comparison of four snow models using observations from an alpine site. Clim. Dyn. *15*, 583–593.

Faroux, S., Kaptué Tchuenté, A.T., Roujean, J.-L., Masson, V., Martin, E., and Le Moigne, P. (2013). ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1-km resolution based on satellite information for use in land surface, meteorological and climate models. Geosci. Model Dev. 6, 563-582.

Fiddes, J., and Gruber, S. (2012). TopoSUB: a tool for efficient large area numerical modelling in complex topography at sub-grid scales. Geosci Model Dev 5, 1245–1257.

Fiddes, J., and Gruber, S. (2014). TopoSCALE v.1.0: downscaling gridded climate data in complex terrain. Geosci Model Dev 7, 387–405.

Förster, K., Meon, G., Marke, T., and Strasser, U. (2014). Effect of meteorological forcing and snow model complexity on hydrological simulations in the Sieber catchment (Harz Mountains, Germany). Hydrol. Earth Syst. Sci. *18*, 4703–4720.

Gaál, L., Szolgay, J., Kohnová, S., Hlavčová, K., Parajka, J., Viglione, A., Merz, R., and Blöschl, G. (2015). Dependence between flood peaks and volumes: a case study on climate and hydrological controls. Hydrol. Sci. J. *60*, 968–984.

Gardent, M., Rabatel, A., Dedieu, J.-P., and Deline, P. (2014). Multitemporal glacier inventory of the French Alps from the late 1960s to the late 2000s. Glob. Planet. Change *120*, 24–37.

Gascoin, S., Hagolle, O., Huc, M., Jarlan, L., Dejoux, J.-F., Szczypta, C., Marti, R., and Sánchez, R. (2015). A snow cover climatology for the Pyrenees from MODIS snow products. Hydrol Earth Syst Sci *19*, 2337–2351.

Gerbaux, M., Genthon, C., Etchevers, P., Vincent, C., and Dedieu, J.P. (2005). Surface mass balance of glaciers in the French Alps: distributed modeling and sensitivity to climate change. J. Glaciol. *51*, 561–572.

Griessinger, N., Seibert, J., Magnusson, J., and Jonas, T. (2016). Assessing the benefit of snow data assimilation for runoff modeling in Alpine catchments. Hydrol Earth Syst Sci *20*, 3895–3905.

Grünewald, T., Schirmer, M., Mott, R., and Lehning, M. (2010). Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment. The Cryosphere *4*, 215–225.

Grusson, Y., Sun, X., Gascoin, S., Sauvage, S., Raghavan, S., Anctil, F., and Sáchez-Pérez, J.-M. (2015). Assessing the capability of the SWAT model to simulate snow,

snow melt and streamflow dynamics over an alpine watershed. J. Hydrol. *531*, *Part 3*, 574–588.

Hall, D.K., and Riggs, G.A. (2007). Accuracy assessment of the MODIS snow product. Hydrol. Process. *21*, 1534–1547.

Hanzer, F., Helfricht, K., Marke, T., and Strasser, U. (2016). Multilevel spatiotemporal validation of snow/ice mass balance and runoff modeling in glacierized catchments. The Cryosphere *10*, 1859–1881.

Harder, P., Pomeroy, J. W. and Helgason, W. (2017) "Local Scale Advection of Sensible and Latent Heat During Snowmelt." Geophysical Research Letters 44.19 (2017): 9769-9777.

Helbig, N., van Herwijnen, A., Magnusson, J., and Jonas, T. (2015). Fractional snow-covered áreas parametrization over complex topography. Hydrology and Earth System Sciences, 19(3), 1339-1351.

Hock, R. (2005). Glacier melt: a review of processes and their modelling. Prog. Phys. Geogr. 29, 362–391.

Hood, J.L., and Hayashi, M. (2015). Characterization of snowmelt flux and groundwater storage in an alpine headwater basin. J. Hydrol. 521, 482–497.

Klein, A.G., and Barnett, A.C. (2003). Validation of daily MODIS snow cover maps of the Upper Rio Grande River Basin for the 2000–2001 snow year. Remote Sens. Environ. *86*, 162–176.

Kling, H., and Nachtnebel, H.P. (2009). A spatio-temporal comparison of water balance modelling in an Alpine catchment. Hydrol. Process. *23*, 997–1009.

Lafaysse, M., Morin, S., Coléou, C., Vernay, M., Serça, D., Besson, F., Willemet, J.M., Giraud, G., and Durand, Y. (2013). Towards a new chain of models for avalanche hazard forecasting in French mountain ranges, including low altitude mountains. Int. Snow Sci. Workshop Grenoble-Chamonix Mont-Blanc.

Lehning, M., Völksch, I., Gustafsson, D., Nguyen, T.A., Stähli, M., and Zappa, M. (2006). ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology. Hydrol. Process. *20*, 2111–2128.

Lehning, M., Grünewald, T., and Schirmer, M. (2011). Mountain snow distribution governed by an altitudinal gradient and terrain roughness. Geophys. Res. Lett. *38*, L19504.

Lejeune, Y., Bertrand, J.-M., Wagnon, P., and Morin, S. (2013). A physically based model of the year-round surface energy and mass balance of debris-covered glaciers. J. Glaciol. *59*, 327–344.

Li, H., Xu, C.-Y., and Beldring, S. (2015). How much can we gain with increasing model complexity with the same model concepts? J. Hydrol. *527*, 858–871.

Liston, G.E., Haehnel, R.B., Sturm, M., Hiemstra, C.A., Berezovskaya, S., and Tabler, R.D. (2007). Simulating complex snow distributions in windy environments using SnowTran-3D. J. Glaciol. *53*, 241–256.

López-Moreno, J.I., and García-Ruiz, J.M. (2004). Influence of snow accumulation and snowmelt on streamflow in the central Spanish Pyrenees / Influence de l'accumulation et de la fonte de la neige sur les écoulements dans les Pyrénées centrales espagnoles. Hydrol. Sci. J. *49*, 787–802.

López-Moreno, J.I., Fassnacht, S.R., Heath, J.T., Musselman, K.N., Revuelto, J., Latron, J., Morán-Tejeda, E., and Jonas, T. (2013). Small scale spatial variability of snow density and depth over complex alpine terrain: Implications for estimating snow water equivalent. Adv. Water Resour. *55*, 40–52.

López-Moreno, J.I., Revuelto, J., Rico, I., Chueca-Cía, J., Julián, A., Serreta, A., Serrano, E., Vicente-Serrano, S.M., Azorin-Molina, C., Alonso-González, E., *et al*. (2016). Thinning of the Monte Perdido Glacier in the Spanish Pyrenees since 1981. The Cryosphere 10, 681–694.

MacDonald, M. K., Pomeroy, J. W., and Pietroniro, A. (2009). Parameterizing redistribution and sublimation of blowing snow for hydrological models: tests in a mountainous subarctic catchment. Hydrological Processes, 23(18), 2570-2583

Mary, A., Dumont, M., Dedieu, J.P., Durand, Y., Sirguey, P. (2013). Intercomparison of retrieval algorithms for the specific surface area of snow from near-infrared satellite data in mountainous terrain, and comparison with the output of a semi-distributed snowpack model. The Cryosphere 7 (2), 741.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., et al. (2013). The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. Geosci Model Dev 6, 929–960.

Marsh, C. B., Pomeroy, J. W., and Spiteri, R. J. (2012). Implications of mountain shading on calculating energy for snowmelt using unstructured triangular meshes. Hydrological Processes, 26(12), 1767-1778.

McCartney SE, Carey SK, Pomeroy JW. (2006). Intra-basin variability of snowmelt water balanced calculations in a subarctic catchment. Hydrological Processes 20: 1001–1016.

McCreight, J.L., Slater, A.G., Marshall, H.P., and Rajagopalan, B. (2012). Inference and uncertainty of snow depth spatial distribution at the kilometre scale in the Colorado Rocky Mountains: The effects of sample size, random sampling, predictor quality, and validation procedures. Hydrological Processes. 28 (3), 933-957.

Meusburger, K., Leitinger, G., Mabit, L., Mueller, M.H., Walter, A., and Alewell, C. (2014). Soil erosion by snow gliding – a first quantification attempt in a subalpine area in Switzerland. Hydrol. Earth Syst. Sci. *18*, 3763–3775.

Mott, R., Schirmer, M., Bavay, M., Grünewald, T., and Lehning, M. (2010). Understanding snow-transport processes shaping the mountain snow-cover. The Cryosphere 4, 545–559.

Nester, T., Kirnbauer, R., Parajka, J., and Blöschl, G. (2012). Evaluating the snow component of a flood forecasting model. Hydrol. Res. *43*, 762–779.

Oreiller, M., Nadeau, D.F., Minville, M., and Rousseau, A.N. (2014). Modelling snow water equivalent and spring runoff in a boreal watershed, James Bay, Canada. Hydrol. Process. *28*, 5991–6005.

Orth, R., Staudinger, M., Seneviratne, S.I., Seibert, J., and Zappa, M. (2015). Does model performance improve with complexity? A case study with three hydrological models. J. Hydrol. 523, 147–159.

Parajka, J., and Blöschl, G. (2008). Spatio-temporal combination of MODIS images – potential for snow cover mapping. Water Resour. Res. 44 (3).

Pomeroy, J.W., Hedstrom, N. and Parviainen, J. (1999). The snow mass balance of Wolf Creek: effects of snow, sublimation and redistribution. In Wolf Creek Research Basin: Hydrology, Ecology, Environment, Pomeroy JW, Granger R (eds). Environment Canada: Saskatoon; 15–30

Pomeroy, J., Essery, R., and Toth, B. (2004). Implications of spatial distributions of snow mass and melt rate for snow-cover depletion: Observations in a subarctic mountain catchment. Ann. Glaciol. 38, 195–201.

Pomeroy, J.W., Bewley, D.S., Essery, R.L.H., Hedstrom, N.R., Link, T., Granger, R.J., Sicart, J.E., Ellis, C.R. and Janowicz, J.R.. (2006). Shrub tundra snowmelt. Hydrological Processes 20: 923–941.

Pomeroy, J., Fang, X., and Ellis, C. (2012). Sensitivity of snowmelt hydrology in Marmot Creek, Alberta, to forest cover disturbance. Hydrol. Process. *26*, 1891–1904.

Quéno, L., Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Dumont, M., and Karbou, F. (2016). Snowpack modelling in the Pyrenees driven by kilometric-resolution meteorological forecasts. The Cryosphere *10*, 1571–1589.

Rabatel, A., Dedieu, J.-P., and Vincent, C. (2005). Using remote-sensing data to determine equilibrium-line altitude and mass-balance time series: validation on three French glaciers, 1994–2002. J. Glaciol. *51*, 539–546.

Rabatel, A., Letréguilly, A., Dedieu, J.-P., and Eckert, N. (2013). Changes in glacier equilibrium-line altitude in the western Alps from 1984 to 2010: evaluation by remote sensing and modeling of the morpho-topographic and climate controls. The Cryosphere *7*, 1455–1471.

Rabatel, A., Dedieu, J.-P., and Vincent, C. (2016). Spatio-temporal changes in glacier-wide mass balance quantified by optical remote-sensing on 30 glaciers in the French Alps for the period 1983-2014. J. Glaciol., 62 (236), 1153-1166. doi: 10.1017/jog.2016.113.

Raleigh, M.S., Lundquist, J.D., and Clark, M.P. (2015). Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. Hydrol Earth Syst Sci *19*, 3153–3179.

Réveillet, M., Vincent, C., Six, D., and Rabatel, A. (2017). Which empirical model is best suited to simulating glacier mass balances? J. Glaciol., 63 (237), 39-54. doi: 10.1017/jog.2016.110.

Réveillet, M., Six, D., Vincent, C., Rabatel., A., Dumont, M., Lafaysse, M., Morin, S., Vionnet, V., Litt, M.(submitted). Relative performance of empirical and physical models in assessing seasonal and annual glacier surface mass balance in the French Alps. The Cryosphere.

Revuelto, J., López-Moreno, J.I., Azorin-Molina, C., and Vicente-Serrano, S.M. (2014). Topographic control of snowpack distribution in a small catchment in the central Spanish Pyrenees: Intra- and inter-annual persistence. Cryosphere *8*, 1989–2006.

Revuelto, J., Vionnet, V., López-Moreno, J.I., Lafaysse, M., and Morin, S. (2016a). Combining snowpack modeling and terrestrial laser scanner observations improves the simulation of small scale snow dynamics. J. Hydrol. 291–307.

Revuelto, J., Jonas, T., and López-Moreno, J.-I. (2016b). Backward snow depth reconstruction at high spatial resolution based on time-lapse photography. Hydrol. Process. 30, 2976–2990.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V. (2010). The AROME-France Convective-Scale Operational Model. Mon. Weather Rev. 139, 976–991.

Schirmer, M., and Jamieson, B. (2015). Verification of analysed and forecasted winter precipitation in complex terrain. The Cryosphere *9*, 587–601.

Schirmer, M., Wirz, V., Clifton, A., and Lehning, M. (2011). Persistence in intra-annual snow depth distribution: 1.Measurements and topographic control. Water Resour. Res. *47*, W09516.

Schön, P., Prokop, A., Vionnet, V., Guyomarc'h, G., Naaim-Bouvet, F., and Heiser, M. (2015). Improving a terrain-based parameter for the assessment of snow depths with TLS data in the Col du Lac Blanc area. Cold Reg. Sci. Technol. *114*, 15–26.

Schweizer, J., Kronholm, K., Jamieson, J.B., and Birkeland, K.W. (2008). Review of spatial variability of snowpack properties and its importance for avalanche formation. Cold Reg. Sci. Technol. *51*, 253–272.

Scipión, D.E., Mott, R., Lehning, M., Schneebeli, M., and Berne, A. (2013). Seasonal small-scale spatial variability in alpine snowfall and snow accumulation. Water Resour. Res. *49*, 1446–1457.

Seidel, F.C., Rittger, K., Skiles, S.M., Molotch, N.P., and Painter, T.H. (2016). Case study of spatial and temporal variability of snow cover, grain size, albedo and radiative forcing in the Sierra Nevada and Rocky Mountain snowpack derived from imaging spectroscopy. The Cryosphere *10*, 1229–1244.

Seyfried, M.S., and Wilcox, B.P. (1995). Scale and the Nature of Spatial Variability: Field Examples Having Implications for Hydrologic Modeling. Water Resour. Res. *31*, 173–184.

Sirguey, P., Mathieu, R., Arnaud, Y., Kahn, M.M., and Chanussot, J. (2008). Improving MODIS spatial resolution for snow mapping using wavelet fusion and ARSIS concept. IEEE Geosci. Remote Sens. Lett. *5*, 78–82.

Sirguey, P., Mathieu, R., and Arnaud, Y. (2009). Subpixel monitoring of the seasonal snow cover with MODIS at 250 m spatial resolution in the Southern Alps of New Zealand: Methodology and accuracy assessment. Remote Sens. Environ. 113, 160–181.

Sirguey, P. (2016) MODImLAB user´s manual V.1.2.5.d, July 2016. Copyright © Pascal Sirguey, Univerity of Otago

Six, D., and Vincent, C. (2014). Sensitivity of mass balance and equilibrium-line altitude to climate change in the French Alps. J. Glaciol. *60*, 867–878.

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

Sturm, M., and Wagner, A.M. (2010). Using repeated patterns in snow distribution modeling: An Arctic example. Water Resour. Res. *46*, W12549.

Tacnet, J.-M., Dezert, J., Curt, C., Batton-Hubert, M., and Chojnacki, E. (2014). How to manage natural risks in mountain areas in a context of imperfect information? New frameworks and paradigms for expert assessments and decision-making. Environ. Syst. Decis. *34*, 288–311.

Thirel, G., Salamon, P., Burek, P., and Kalas, M. (2013). Assimilation of MODIS snow cover area data in a distributed hydrological model using the particle filter. Remote Sens. 5, 5825–5850.

Trujillo, E., Ramírez, J.A., and Elder, K.J. (2007). Topographic, meteorologic, and canopy controls on the scaling characteristics of the spatial distribution of snow depth fields. Water Resour. Res. 43, W07409.

Verfaille, D., Déqué, M., Morin, S. and Lafaysse, M. (2017). The method ADAMONT v1.0 for statistical adjustment of climate projections applicable to energy balance and surface models. Geoscientific Model Development, 10, 4257-4283

Viani, A., Condom, T., Vincent, C., Rabatel, A., Bacchi, A., Sicart, J.E., Revuelto, J., Six, D., and Zin, I. (~~submitted~~2018). Glacier-wide summer surface mass balance reconstruction: hydrological balance applied on Argentière and Mer de Glace drainage basins (Mont Blanc, France). Journal of Glaciology.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E., and Willemet, J.-M. (2012). The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2. Geosci. Model Dev. *5*, 773–791.

Vionnet, V., Martin, E., Masson, V., Guyomarc'h, G., Naaim-Bouvet, F., Prokop, A., Durand, Y., and Lac, C. (2013). Simulation of wind-induced snow transport in alpine terrain using a fully coupled snowpack/atmosphere model. Cryosphere Discuss. 7, 2191–2245.

Vionnet, V., Martin, E., Masson, V., Guyomarc'h, G., Naaim-Bouvet, F., Prokop, A., Durand, Y., and Lac, C. (2014). Simulation of wind-induced snow transport and sublimation in alpine terrain using a fully coupled snowpack/atmosphere model. The Cryosphere *8*, 395–415.

Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Quéno, L., Seity, Y., and Bazile, E. (2016). Numerical Weather Forecasts at Kilometer Scale in the French Alps: Evaluation and Application for Snowpack Modeling. J. Hydrometeorol. *17*, 2591–2614.

Vionnet, V., Guyomarc'h, G., Lafaysse. M., Naaim-Bouvet. F., Giraud, G. and Deliot, Y (2018). Operational implementation and evaluation of a blowing snow scheme for avalanche hazard forecasting. Cold Reg. Sci. Technol. 147, 1-10

Viviroli, D., Dürr, H.H., Messerli, B., Meybeck, M., and Weingartner, R. (2007). Mountains of the world, water towers for humanity: Typology, mapping, and global significance. Water Resour. Res. *43*, W07447.

Weusthoff, T., Ament, F., Arpagaus, M., and Rotach, M.W. (2010). Assessing the Benefits of Convection-Permitting Models by Neighborhood Verification: Examples from MAP D-PHASE. Mon. Weather Rev. *138*, 3418–3433.

Winstral, A., Marks, D., and Gurney, R. (2012). Simulating wind-affected snow accumulations at catchment to basin scales. Adv. Water Resour. 55, 64-79

Wipf, S., Stoeckli, V., and Bebi, P. (2009). Winter climate change in alpine tundra: plant responses to changes in snow depth and snowmelt timing. Clim. Change *94*, 105–121.

Figures

Figure 1: Upper Arve catchment study area. The white shaded area shows the extent of the glaciers in 2012 (Gardent *et al*., 2014). The inner maps show various magnifications of the Alps and the location of the Arve valley within the mountain range. The red points show the position of the five Météo-France stations located in the study area.

Figure 2: Schematic representation of the approaches used to account for mountain spatial heterogeneity when simulating snowpack dynamics.

Figure 3: Glacier SMB measurement locations for ablation and accumulation areas in the Mer de Glace and Argentière glaciers.

Figure 4: Observed (black squares) and simulated (red lines) snow depth evolution for the 2007–08 (upper panel) and 2012–13 (bottom panel) snow seasons. The elevations of the stations are: Chamonix: 1025 m.a.s.l.; Le Tour: 1470 m.a.s.l.; La Flegere: 1850 m.a.s.l.; Lognan: 1970 m.a.s.l.; and Aiguilles Rouges: 2365 m.a.s.l.

Figure 5: Spatial distribution of the UWS MODImLab product (equivalent to the SCA distribution), and the simulated snow depth obtained using the distributed approach (the purple color shows the snow depth values exceeding the 0.15 m threshold) for 24 July 2008.

Figure 6: Temporal evolution of the SCA (2004–2014) based on semi-distributed and distributed simulations and MODIS sensor observations. The vertical bars associated with the MODIS observations show the uncertainty associated with cloud presence for days having ≤ 20% snow cover.

Figure 7: Observed and simulated SCA evolution for a period of low level snowpack accumulation (2006–2008; upper panel) and a period of high level snowpack accumulation (2011–2013 lower panel). The vertical bars for the MODIS observations show the uncertainty associated with cloud presence for days having ≤ 20% snow cover. Red and blue shading for the distributed and semi-distributed SCA simulations show the uncertainty associated with various snow depth thresholds for determining whether a pixel was snow covered. The lower limit of the shading represents the SCA evolution for a 0.1 m threshold, the upper limit of the shading represents a 0.2 m snow depth threshold, and the middle line represents a 0.15 m snow depth threshold.

Figure 8: Evolution of the SCA in relation to north and south aspect for the 2006–2008 (upper panel; low level of snowpack accumulation) and 2011–2013 (lower panel; high level of snowpack accumulation) snow seasons. Vertical bars for the MODIS observations show the uncertainty associated with cloud presence for days having ≤ 20% snow cover. Red and blue shading for the distributed and semi-distributed SCA

simulations show the uncertainty associated with various snow depth thresholds for determining whether a pixel was snow covered. The lower limit of the shading represents the SCA evolution for a 0.1 m threshold, the upper limit of the shading represents a 0.2 m snow depth threshold, and the middle line represents a 0.15 m snow depth threshold.

Figure 9: Jaccard index and ASSD values for low level (2006–07 and 2007–08) and high level (2011–12 and 2012–13) snow accumulation seasons.

Figure 10: Temporal evolution of the observed and simulated (semi-distributed and distributed) SMB for the Argentière glacier for the four 300-m elevation bands for the period 1994–2013. The points show the average observation and simulation values for the same measurement locations, and the vertical bars show the standard deviations for those values.

Figure 11: Temporal evolution of the observed and simulated (semi-distributed and distributed) SMB for the Mer de Glace glacier for the seven 300-m elevations bands for the period 1994–2013. The points show the average observation and simulation values for the same measurement locations, and the vertical bars show the standard deviations for those values.

Figure 12: Altitudinal dependence of the observed and simulated (semi-distributed and distributed) SMB for two snow seasons (2007–08: low level snow accumulation; and 2012–13: high level snow accumulation) at the Mer de Glace glacier.

Figure 13: Observed and simulated evolution of the ELA for the five glaciers during the study period, based on the same dates as those for the satellite image acquisition.

Tables

Observatory	RMSE [cm]	Bias[cm]	Period	Num. Obs.
Chamonix	23.3	12.1	1983-2015	6704
Le Tour	29.6	13.0	1985-2015	6323
Nivose Aiguilles Rouges	66.6	49.4	1983-2015	5902
La Flegere	45.0	-19.1	2003-2015	1231
Lognan	20.8	1.9	1994-2015	5964

Table 1: Error statistics (bias and RMSE) between simulated and in situ snow depth observations for the five meteorological stations in the study area for periods for which observations were available. The locations of the stations are shown in Figure 1.

Threshold		Distributed approach			Semi-Distributed approach		
SCA [0,1]	SD [m]	R2	RMSE	MAE	R2	RMSE	MAE
0.25	0.1	0.803	13.84	9.85	0.790	15.48	10.36
	0.15	0.807	13.75	9.54	0.793	15.04	9.79
	0.2	0.806	13.79	9.60	0.789	16.41	12.05
0.35	0.1	0.821	12.64	8.36	0.809	14.31	9.79
0.35	**0.15**	**0.828**	**12.51**	**8.24**	**0.815**	**13.59**	**9.60**
	0.2	0.815	12.86	8.54	0.811	14.90	10.49
0.45	0.1	0.812	13.47	9.33	0.798	15.29	10.47
	0.15	0.813	13.69	9.58	0.805	14.31	9.81
	0.2	0.813	13.38	9.24	0.80	16.29	11.17

Table 2: UWS threshold selection ~~performance~~ for various snow thicknesses selected as thresholds for the 2008–09 and 2009–10 snow seasons for distributed and semi-distributed simulations. Bold values indicate the selected snow depth and SCA thresholds.

~~Period~~Exposition	**Approach**	**R^2**	**MAE**	**RMSE**
~~Entire period (2001-2015)~~Whole catchment	Semi-distributed	~~0.815~~0.819±0.022	**~~10.47~~10.40±0.39**	**~~15.28~~15.2±0.71**
	Distributed	~~0.822~~0.821±0.021	**~~8.35~~8.34±0.30**	**~~12.64~~12.6±0.77**
~~2006-07 to 2007-08~~North ~~Exposition~~aspect	Semi-distributed	~~0.744~~0.721±0.036	**~~10.756~~10.09±0.52**	**~~16.903~~15.98±0.88**
	Distributed	~~0.756~~0.726±0.035	**~~8.74~~7.57±0.38**	**~~14.82~~12.73±0.97**
~~2011-12 to 2012-13~~South ~~Exposition~~aspect	Semi-distributed	~~0.881~~0.858±0.018	**~~11.56~~10.17±0.31**	**~~15.58~~14.78±0.64**
	Distributed	~~0.895~~0.857±0.016	**~~7.99~~9.83±0.39**	**~~11.10~~13.19±0.70**

Table 3: R^2, MAE and RMSE average and standard deviations values from the 100 sample bootstrapping for the observed and simulated SCA (based on the distributed and semi-distributed approaches) for the entire time period with SCA observation (2001-2015). Results for the entire study area and for North and South Expositions are presented. Error metrics in bold note p-values of the T-student test lower than 0.01 (99% confidence interval for rejecting null hypothesis).

Period	**Approach**	**R^2**	**MAE**	**RMSE**
~~2006-07 to 2007-08~~Whole catchment	Semi-distributed	0.744	10.756	16.903
	Distributed	0.756	8.74	14.82
~~2011-12 to 2012-13~~Northern aspect	Semi-distributed	0.58~~0.881~~	11.26~~11.56~~	18.36~~15.58~~
	Distributed	0.59~~0.895~~	8.61~~7.99~~	15.62~~11.10~~
Southern aspect	Semi-distributed	0.80	10.17	16.48
	Distributed	0.815	10.34	16.21

Table 4: RMSE, MAE and R^2 values for the observed and simulated SCA (based on the distributed and semi-distributed approaches) for ~~various~~2006-2008 time periods for the whole catchment, Northern aspect (N, NE, NW).and Southern aspect (S, SE, SW).~~those parts of the study area having a northern aspect (N, NE, NW).~~

Period	Approach	R^2	MAE	RMSE
Whole catchment~~Entire period~~ ~~(2001–2015)~~	Semi-distributed	0.881~~0.71~~	11.56~~10.12~~	15.58~~16.04~~
	Distributed	0.895~~0.72~~	7.99~~7.60~~	11.10~~12.84~~
Northern aspect~~2006–07 to 2007–08~~	Semi-distributed	0.82~~0.58~~	11.30~~11.26~~	16.38~~18.36~~
	Distributed	0.84~~0.59~~	7.79~~8.61~~	11.69~~15.62~~
Southern aspect~~2011–12 to 2012–13~~	Semi-distributed	0.902~~0.82~~	10.98~~11.30~~	15.09~~16.38~~
	Distributed	0.905~~0.84~~	8.25~~7.79~~	11.81~~11.69~~

Table 5: RMSE, MAE and R^2 values for the observed and simulated SCA (based on the distributed and semi-distributed approaches) for various2011-2013 time period for the whole catchment, Northern aspect (N, NE, NW).and Southern aspect (S, SE, SW).

Period	Approach	Jaccard	ASSD
Entire period (2001–2015)	Semi-distributed	0.817	0.912
	Distributed	0.832	0.975
2006–07 to 2007–08	Semi-distributed	0.783	0.920
	Distributed	0.801	0.952
2011–12 to 2012–13	Semi-distributed	0.826	0.897
	Distributed	0.836	0.952

Table 6: Average values of the Jaccard index and ASSD values for each simulation approach for various time periods.

Period	Approach	Jaccard Index		ASSD	
		JFM	MJJ	JFM	MJJ
2006–07	Semi-distributed	0.9535	0.802	0.687	1.152
	Distributed	0.9557	0.823	0.704	1.104
2007–08	Semi-distributed	0.950	0.793	0.717	1.062
	Distributed	0.951	0.809	0.724	1.043
2011–12	Semi-distributed	0.968	0.756	0.711	0.983
	Distributed	0.967	0.754	0.734	0.994
12012–13	Semi-distributed	0.980	0.790	0.199	1.271
	Distributed	0.990	0.799	0.198	1.250

Table 7: Average values of the Jaccard index and ASSD for each simulation approach for the maximum (JFM) and minimum (MJJ) snow accumulation periods.

Glacier	Period	Approach	RMSE	MAE	R2
Arg	WSMB	Semi-distributed	**0,56±0.027~~0.53~~**	**0,43±0.022~~0.42~~**	**0,54±0.065~~0.537~~**
		Distributed	**0,51±0.028~~0.52~~**	**0,42±0.023~~0.40~~**	**0,50±0.073~~0.51~~**
	SSMB	Semi-distributed	**0,96±0.057~~0.96~~**	**0,78±0.045~~0.78~~**	**0,75±0.031~~0.72~~**
		Distributed	**0,77±0.049~~0.76*~~**	**0,62±0.047~~0.61~~**	**0,84±0.019~~0.84~~**
	ASMB	Semi-distributed	**1,21±0.059~~1.21~~**	**0,99±0.055~~0.99~~**	0,72±0.021~~**0.71**~~
		Distributed	**1,18±0.062~~1.05~~**	**0,909±0.054~~0.85~~**	0,71±0.055~~**0.78**~~
Mdg	WSMB	Semi-distributed	**0,73±0.031~~0.72~~**	0,57±0.024~~**0.56**~~	**0,64±0.041~~0.64~~**
		Distributed	**0,76±0.026~~1.57~~**	0,58±0.027~~**1.15**~~	**0,59±0.045~~0.83~~**
	SSMB	Semi-distributed	**1,47±0.093~~1.46~~**	**1,18±0.083~~1.17~~**	**0,746±0.049~~0.75~~**
		Distributed	**1,19±0.069~~1.19~~**	**0,86±0.057~~0.86~~**	**0,86±0.014~~0.86~~**
	ASMB	Semi-distributed	**1,74±0.095~~1.72~~**	**1,36±0.075~~1.33~~**	**0,76±0.041~~0.75~~**
		Distributed	**1,57±0.088~~1.57~~**	**1,16±0.069~~1.15~~**	**0,838±0.020~~0.83~~**

Table 8: R^2, MAE and RMSE average and standard deviations values from the 100 sample bootstrapping for the observed and simulated SMB for Mer de Glace (Mdg) and Argentière (Arg) glaciers (based on the distributed and semi-distributed approaches. Error metrics in bold note p-values of the T-student test lower than 0.01 (99% confidence interval for rejecting null hypothesis).
~~RMSE, MAE, R^2 values for the slope and intersection in linear adjustments between the observed and simulated SMB for Mer de Glace (Mdg) and Argentière (Arg) glaciers.~~

Glacier	Approach	Avg Dif	Std. Dev (Differences)	Slope	R2
Mdg	Semi-distributed	155.11	69.62	0.715	0.420
	Distributed	88.57	48.90	0.869	0.627
Les	Semi-distributed	158.34	101.84	0.188	0.102
	Distributed	110.73	109.67	0.560	0.586
Tal	Semi-distributed	105.14	59.25	0.4936	0.2336
	Distributed	80.12	41.87	0.766	0.476
Tour	Semi-distributed	105.14	59.25	0.339	0.528
	Distributed	84.33	68.71	0.625	0.715
Arg	Semi-distributed	63.89	42.87	0.270	0.103
	Distributed	54.52	31.85	0.578	0.381

Table 9: Average differences, standard deviations, slope of the linear adjustment, and R2 values for the observed and simulated ELA for Mer de Glace (Mdg), Leschaux (Les), Talefre (Tal), Tour and Argntière (Arg) glaciers.