


Impact of snow distribution modelling for runoff predictions

Ilaria Clemenzi ^{a,*}, David Gustafsson^a, Wolf-Dietrich Marchand^b, Björn Norell^c, Jie Zhang^d, Rickard Pettersson^d and Veijo Allan Pohjola^d

^a Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

^b Norwegian Water Resources and Energy Directorate (NVE), Trondheim, Norway

^c Vattenregleringsföretagen, Östersund, Sweden

^d Department of Earth Sciences, Uppsala University, Uppsala, Sweden

*Corresponding author. E-mail: ilaria.clemenzi@smhi.se

 IC, 0000-0001-7140-6647

ABSTRACT

Snow in the mountains is essential for the water cycle in cold regions. The complexity of the snow processes in such an environment makes it challenging for accurate snow and runoff predictions. Various snow modelling approaches have been developed, especially to improve snow predictions. In this study, we compared the ability to improve runoff predictions in the Överuman Catchment, Northern Sweden, using different parametric representations of snow distribution. They included a temperature-based method, a snowfall distribution (SF) function based on wind characteristics and a snow depletion curve (DC). Moreover, we assessed the benefit of using distributed snow observations in addition to runoff in the hydrological model calibration. We found that models with the SF function based on wind characteristics better predicted the snow water equivalent (SWE) close to the peak of accumulation than models without this function. For runoff predictions, models with the SF function and the DC showed good performances (median Nash–Sutcliffe efficiency equal to 0.71). Despite differences among the calibration criteria for the different snow process representations, snow observations in model calibration added values for SWE and runoff predictions.

Key words: hydrological modelling, model calibration, mountainous catchment, snow modelling, snow spatial distribution, snowmelt runoff

HIGHLIGHTS

- Models with a snow distribution based on wind and topography in addition to precipitation and temperature improved snow predictions.
- Models with a snow distribution based on wind and topography could use snow information and perform similarly to models with a depletion curve for runoff.
- The robustness of model calibration increased by including spatially distributed snow observations in addition to runoff data.

1. INTRODUCTION

Snow is a key component of the water cycle in cold regions. Snow, accumulated in the mountains, stores large amounts of winter precipitation that is reallocated in time to spring and summer runoff, influencing streamflow variability in the downstream regions. The knowledge of the amount of winter snow, snowmelt timing and runoff is essential for the sustainable use of this resource for human consumption, agriculture and hydropower (Viviroli *et al.* 2011; Sturm *et al.* 2017), ecological services (Wipf & Rixen 2010; Boelman *et al.* 2019) and flood control.

Accurate predictions of snow in this resource are challenging because processes at different scales distribute this resource unevenly (Clark *et al.* 2011). Orographic effects on precipitation, wind and vegetation induce snow accumulation in topographic depressions and erosion in wind-swept areas (Liston & Sturm 1998; Winstral *et al.* 2002). Spatially variable air temperature and exposure to solar radiation control the snow energy budget and influence snowmelt patterns and runoff dynamics (DeBeer & Pomeroy 2010; Grünwald *et al.* 2010). Depending on the modelling purpose and data availability, snow spatial distribution in mountain catchments has been represented in different ways. Snowpack energy balance models complemented with snow transport algorithms were used to explicitly simulate all physical processes, including the effect of wind redistribution on snow depth patterns (Lehning *et al.* 2006; Liston & Elder 2006). These models can

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

reproduce snow distribution at relatively very fine (<100 m) spatial scales (Lehning *et al.* 2008; Mott *et al.* 2010; Marsh *et al.* 2020). However, they require spatially detailed meteorological (e.g., air temperature, wind and solar radiation) and snow observations, which are difficult to obtain and scarce in the mountains. A way to model the snow spatial distribution using only precipitation and air temperature consisted of distributing snowfall or rainfall and melt rates as a function of elevation, i.e., with the temperature-index model. Given its simplicity, this approach was found particularly suitable to be integrated into models for runoff predictions with snow processes represented at the scale of landscape units by lumping similar land uses, soils and elevations (Lindström *et al.* 1997; Viviroli *et al.* 2009; Seibert & Vis 2012). Using a simple model structure, Girons Lopez *et al.* (2020) evaluated the effect of modifying different temperature-index models, discovering that few modifications were most valuable.

Based on snow depth or snow water equivalent (SWE) observations, different studies have explored the link and the predictive power of topographic characteristics (e.g., elevation, slope, aspect and curvature) alone and with wind data to snow spatial distribution (Winstral *et al.* 2002, 2013; Grünewald *et al.* 2013; Helbig *et al.* 2015; Skaugen & Melvold 2019). Topographic characteristics are, in fact, readily extractable from digital elevation models. The combination of different topography-related factors was meant to account for the integrated effect of snow accumulation and melt processes on snow spatial distribution.

The predictive power of the topography-related factors was found to increase with the spatial aggregation of the snow observations (Grünewald *et al.* 2013; Helbig *et al.* 2015). When applied at different alpine sites, they were found to be site-specific and dependent on the spatial resolution of snow data. Studies on the relation between snow amount and topography-related factors mainly provided information on snow spatial variability at the peak of the accumulation but did not evaluate the results for runoff predictions.

Next to the spatial distribution of snow amount, snow cover variability was also explored and used to model the effect of uneven snow spatial distribution on melt dynamics (Luce *et al.* 1998; Luce & Tarboton 2004; Pomeroy *et al.* 2004). Snow cover depletion curves were found suitable to implicitly model snow processes by linking the temporal evolution of SWE to the snow-covered area on larger model units, e.g., sub-kilometre, than fine grid (Skaugen & Randen 2013; Frey & Holzmann 2015). The shapes of SWE distributions for such relations were assumed a priori and consisted, for example, of two-parameter log-normal or gamma distributions (Skaugen & Randen 2013; Frey & Holzmann 2015). In such approaches, the SWE dynamic was linked to precipitation forcing. The distribution parameters were estimated through model calibration and validation on runoff and fractional snow cover data, respectively.

In addition to the different model designs, the key point is, in fact, the availability of data to estimate or validate snow model parameters. Most studies used runoff and satellite data to estimate these parameters through calibration and to validate the evolution in time of the snow-covered area (Freudiger *et al.* 2017). Although spatially distributed, satellite snow cover data do not provide quantitative information about snow amount, while runoff provides integrated information on all catchment hydrological processes. Only a few studies used SWE in addition to runoff observations for parameter estimations (Whitaker *et al.* 2003; Girons Lopez *et al.* 2020; Nemri & Kinnard 2020). However, these studies mainly used SWE to calibrate one snow modelling approach, i.e., the snow routines based on temperature-index models.

Attempts to model snow redistribution processes for a more realistic representation of snow in mountainous terrain for runoff predictions are still needed, as noted by Freudiger *et al.* (2017) and Frey & Holzmann (2015).

In line with the need to improve the representation of the snow distribution in mountain catchments for runoff predictions, the aim of this paper is two-fold. On one hand, we investigate the impact of different snow model structures on SWE and runoff predictions. As a novel component, a snowfall distribution (SF) function based on wind shelter factors and wind direction data was implemented in the semi-distributed hydrological model HYPE (hydrological predictions for the environment, Lindström *et al.* 2010). On the other hand, assess the value of using distributed snow observations in addition to runoff and non-distributed snow observations in hydrological model calibration using snow survey observations and runoff observations from the Lake Överuman Catchment, Northern Sweden.

2. STUDY SITE

The Överuman Catchment is located in the north-west part of Sweden at the border with Norway and is the headwater of the River Umeälven, which drains into the Bothnian Bay (Figure 1). The Överuman Catchment extends over an area of 652 km² with its outlet through the Överuman Lake (85 km²), used as a hydropower reservoir. Elevation ranges between 524 and 1,575 m above sea level with an average elevation of 834 m. The catchment area mostly consists of bare rock, moss, heathers and shrubs (56%), while forests

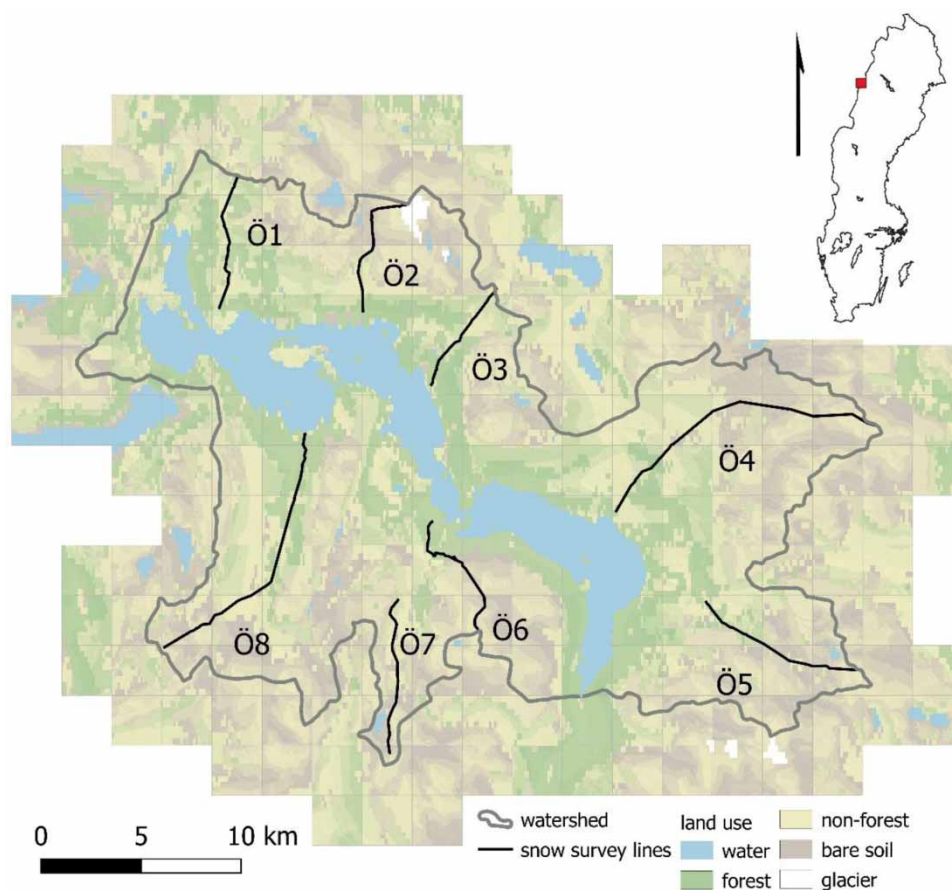


Figure 1 | Map of the Överuman Catchment. Catchment boundary is shown with a thick grey line and represents the hydrological model unit in the lumped model setup. Square grids ($2.5 \times 2.5 \text{ km}^2$) in light grey represent the hydrological model units in the gridded model setup. The snow survey lines (Ö1–Ö8) are represented by thick black lines.

cover about 31% of the lower areas along the lake. Sparse birch forests are at higher elevations and slightly denser birch forests are in the lowest areas along the large lakes. The terrain varies from moderate to very steep in some areas.

3. DATA

3.1. Meteorological forcing and runoff data

Meteorological forcing and runoff data covered the periods 2016–2020. The meteorological forcing data consist of precipitation, air temperature and fields of wind direction. Precipitation data were provided by a gridded product based on gauge data, Precipitation Temperature Hydrologiska Byråns Vattenmodell (PTHBV) (Johansson & Chen 2003). The PTHBV product is available at a spatial resolution of $4 \times 4 \text{ km}^2$ and a daily time resolution. Hourly air temperature and wind direction data were provided by the MESAN meteorological analysis of the AROME meteorological model and observations (Häggmark *et al.* 2000). The spatial resolution of air temperature and wind direction data is $2.5 \times 2.5 \text{ km}^2$. They were aggregated to a daily time scale, the same temporal resolution of precipitation data. Precipitation data were resampled to the same air temperature and wind direction data spatial resolution. Both products are currently used by the Swedish Meteorological and Hydrological Institute (SMHI 2022).

The daily average runoff at the catchment outlet is measured as the residual between reservoir storage change and reservoir outflow. These data were provided by the water regulation company, Vattenregleringsföretagen AB.

3.2. Snow data

Snow surveys were performed in the Överuman Catchment close to the snow accumulation peak in the years 2017–2020 along survey lines distributed in the catchment area (Figure 1). Snow depth measurements were acquired along the survey

lines using a ground penetrating radar (GPR) system (Sensor & Software, Noggin Plus with a 500 MHz antenna) connected to a global positioning system (Trimble) pulled by a snowmobile. SWE was obtained by combining the two-way travel time (TWT) of the radar signal, between the snow surface and the ground, with point density measurements. The latter were taken approximately every kilometre along the survey lines (Marchand 2003). The final spatial resolution between SWE measurements was about 1 m distance. SWE estimates were aggregated to the spatial resolution of 25 m distance and used for model calibration and validation.

3.3. Topographic and vegetation data

A digital elevation model (DEM) with a resolution of 25×25 m was obtained by bilinear resampling of a 32×32 m DEM provided by ArcticDEM (Porter *et al.* 2018). The digital elevation model was used to extract topographic characteristics of the catchment (i.e., elevation, aspect) and to calculate the wind shelter factor (Winstral *et al.* 2002). Elevation and aspect were used to discretize the catchment landscape in classes of the model domain. The wind shelter factor was used to model the snow spatial variability induced by wind transport processes in the catchment (Section 4.1). Land-use characteristics were derived from the European Space Agency (ESA) land cover climate change initiative data aggregated into surface water, forest, non-forest, bare soil and glacier classes.

4. METHODS

4.1. HYPE model setup

This study used the semi-distributed hydrological model HYPE (Lindström *et al.* 2010) to simulate SWE and runoff in four hydrological years (2016–2020). The model consists of several subroutines where the different hydrological processes are described. The landscape can be divided into different units. Each is further divided into classes representing different soil and land use properties. Model parameters link landscape characteristics to the physical processes through the model parameters. Two different model spatial configurations were considered, lumped and gridded (Figure 1). In the lumped configuration, the Överuman Catchment consisted of one model unit and meteorological data were aggregated at the catchment scale. In the gridded configuration, the catchment was divided into 2.5×2.5 km² units preserving the spatial variability of meteorological forcing at that spatial scale. Each unit was further divided into three elevation zones, four aspect zones (north, east, south and west facing slopes), five land use classes (water, forest, non-forest, bare soil and glacier) and one soil class. In general, in the HYPE model, hydrological processes in the classes are simulated independently within each model unit, except for the SF model introduced in this study, which will be described in more detail in the next section. Runoff from the land classes is aggregated into water classes representing river and lake networks in each unit. It is further routed into the river classes in the next downstream model unit.

4.2. Snow processes in HYPE

In the HYPE model, four different snow models were considered: base, depletion curve (DC), snowfall distribution (SF) and the SF with the DC (Table 1). In all of the model configurations, precipitation (P) is distributed between snowfall, P_s and rainfall, P_r , as a linear function of air temperature (T) within a temperature interval:

$$P_r = P \cdot a_{\text{rain}} \quad (1)$$

$$P_s = P \cdot (1 - a_{\text{rain}}) \quad (2)$$

$$a_{\text{rain}} = \begin{cases} 0 & T < ttpd - 1 \\ 1 & T > ttpd + 1 \\ \frac{T - (ttpd - 1)}{2} & \text{else} \end{cases} \quad (3)$$

where a_{rain} is the proportion of rainfall and $ttpd$ (°C) is the mid-point of the temperature interval with mixed rainfall/snowfall precipitation (Table 2). For this study, the width of the temperature interval with mixed precipitation was fixed to 2 °C, but it may also be changed using additional parameters that were not used here. The correction of air temperature as a function of elevation will to some extent generate the variability of snowfall and rainfall fraction between the land surface classes of a model unit depending on their elevation differences.

Table 1 | Model type and snow process representation

Model type	Snow process representation in land classes
Base	Snow accumulation: snowfall/rainfall partition by elevation-temperature changes Snow melt: proportional to air temperature
Depletion curve (DC)	Snow accumulation: snowfall/rainfall partition by elevation-temperature changes Snow melt: proportional to air temperature and scaled by DC (snow-covered area – SWE relation)
Snowfall distribution (SF)	Snow accumulation: (1) snowfall/rainfall partition by elevation-temperature changes and (2) SF function Snow melt: proportional to air temperature
Snowfall distribution and depletion curve (SF + DC)	Snow accumulation: (1) snowfall/rainfall partition by elevation-temperature changes and (2) SF function Snow melt: proportional to air temperature and scaled by DC (snow-covered area – SWE relation)

Table 2 | Model parameters considered in the calibration strategy

Parameter	Description	Min-Max	Unit
Process above ground			
<i>ttpd</i>	Snowfall/rainfall threshold	[−2,0 4,0]	°C
<i>cevpcorr</i>	Correction factor for evaporation	[−1,0 1,0]	–
<i>wsfscale</i>	Scaling factor for SF	[0,01 0,1]	–
<i>sfamax</i>	Maximum amount of SF	[5,0 10,0]	–
Land processes			
<i>cmlt</i> (forest, non-forest, bare soil)	Snow melt factor	[2 5]	°C mm day ^{−1}
<i>wcep</i>	Effective porosity	[0,01 0,2]	–
<i>sracs</i> (forest, non-forest, bare soil)	Recession coefficient for surface runoff	[0,001 0,5]	day
<i>rrcs1</i>	Recession coefficient for uppermost soil layer	[0,2 0,7]	day ^{−1}
<i>rrcs2</i>	Recession coefficient for lowest soil layer	[0,03 0,2]	day ^{−1}
<i>mper1</i>	Maximum percolation capacity from soil layer 1 to soil layer 2	[10 100]	mm day ^{−1}
<i>mper2</i>	Maximum percolation capacity from soil layer 2 to soil layer 3	[10 100]	mm day ^{−1}

For each model unit and land surface class, the snowfall is accumulated in a single snow layer, characterized by its SWE (mm), (optionally) depth (cm) and density (kg m^{−3}). Snow melt is calculated by a temperature-index model, with an air temperature threshold, *ttpd*, optionally treated as a free parameter, but here fixed at 0 °C to initiate snow melt and release water for soil infiltration. First, the potential amount of melting snow, *pmelt*, is calculated as proportional to the air temperature through the parameter *cmlt* (°C mm day^{−1}). The actual melt is then based on the available SWE, *swe* (mm):

$$pmelt = cmlt \cdot (T - ttpd) \quad (4)$$

$$melt = \min(pmelt, swe) \quad (5)$$

In this model version, snow is assumed to cover the entire land class, if present.

The model, including a snow DC to calculate the snow-covered area separately for each land class, is referred to as the DC model (Table 1). The snow DC is used to scale the estimated snow melt and evaporation.

In this case, the snow-covered area in each land class, fsc , is related to the SWE, swe , following a hyperbolic function:

$$fsc = fsc_{\max} \cdot \tanh \frac{swe}{10} \quad (6)$$

where fsc_{\max} is the maximum value set equal to one. The snow-covered area is then used to scale the potential melt following:

$$pmelt = cmlt \cdot (T - tmp) \cdot fsc \quad (7)$$

The effective melt is then calculated as in the base model.

In this study, an explicit SF function was introduced to represent the influence of local topography on the spatial variability of snow.

The SF function based on the wind shelter factor (Winstral *et al.* 2002) was introduced. The wind shelter factor was calculated on a $25 \times 25 \text{ m}^2$ resolution DEM for eight wind directions and five searching distances (from 150 up to 2,500 m). A preliminary investigation of the predictive power of the wind shelter factor and the GPR-derived SWE data (Section 3.2) at 25 m spatial resolution indicated a correlation between the two variables (Figure S1). To further investigate the type of relationship between SWE data and wind shelter factor that could have been used to model the snow distribution, we derived the ratio between the SWE in each GPR point and the mean SWE along the survey line. A log-linear function best represented the relation between the snow spatial distribution and the wind shelter factor (Figure S2). This optimal searching distance of 300 m was chosen as a trade-off among the different analysed searching distances based on the correlation between GPR-derived SWE and wind shelter factor for the eight wind directions (Figure S1). The wind shelter factors were then aggregated on the land classes of the model.

Snowfall within a model land class was, thus, distributed by the wind shelter factor following the relation:

$$\frac{P_{s \text{ class}}}{P_{s \text{ mean}}} = w_N \cdot 10^{(wsf \text{ scale} \cdot WSF_{\text{class}})} \quad (8)$$

where $P_{s \text{ mean}}$ is the mean snowfall in the model unit at that time step and $P_{s \text{ class}}$ and WSF_{class} are the corrected snowfall and wind shelter factors for the land class, respectively. The $wsf \text{ scale}$ is a scaling parameter and w_N is a wind shelter weight calculated separately so that the sum of the snowfall correction ratios within a model unit (N) adds up to one. The parameter $sfdmax$ is an upper limit to the SF ratio ($P_{s \text{ class}}/P_{s \text{ mean}}$). The model including the SF function is referred to as the SF model (Table 1). We finally considered a model with the combination of the SF function and DC (SF + DC; Table 1).

4.3. Model calibration and validation

The model calibration was based on the generalized likelihood uncertainty estimation (GLUE) technique (Beven & Binley 1992) and a split-sample approach (Klemes 1986). The GLUE method was implemented as follows: a Monte Carlo sampling from a parameter space was used to generate a prior distribution of model realizations (Beven & Freer 2001). Pseudo-likelihood transformations of performance criteria (goodness-of-fit between observations and predictions) were introduced and used to obtain the posterior distributions, rescaling the prior distribution with the pseudo-likelihood values of each model realization as weight. This rescaling was performed by resampling the prior distribution (with replacement) with the estimated pseudo-likelihood values as a probability. The pseudo-likelihood transformations consisted of scaling performance criteria linearly between zero and one. For this purpose, a minimum and a maximum threshold were defined for each performance criterion. In this way, a strict subjective decision of a behavioural cut-off value that excludes the model realizations below that threshold, as often done in the implementations of the GLUE methodology, is avoided. In addition, the pseudo-likelihood transformations help to combine different performance criteria to rescale the prior distribution into the posterior distribution of model realizations. Due to the limited snow data available, model calibration and validation can be sensitive to a specific year. This limitation is overcome with the split-sample approach that divides the study period into an equal number of years both for calibration and validation (cross-calibration/validation).

Around 10,000 simulations were performed with a Monte Carlo sampling from a feasible parameter space with uniform distribution for both the lumped and gridded HYPE models. Model parameters and their ranges were defined based on previous HYPE model applications in cold environments (Strömqvist *et al.* 2012; Gelfan *et al.* 2017), as reported in Table 2.

Multiple criteria were considered for model calibration and validation with two hydrological variables, SWE and runoff, including the temporal and spatial information content of the observations: the Nash–Sutcliffe efficiency (q_1 , Nash & Sutcliffe 1970), the volume error (Madsen 2000), calculated between computed and observed runoff (q_2), the volume error calculated between computed and observed catchment SWE mean (s_1) and the Pearson correlation coefficient (Freedman *et al.* 2007) calculated between the computed and observed SWE spatial frequency distributions (s_2). The performance criteria were transformed into pseudo-likelihood functions and linearly rescaled between zero and one using a chosen minimum (x_{\min}) and a maximum (x_{\max}) threshold for each performance criterion (x):

$$f_x = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad \text{where } x = q_1, q_2, s_1, s_2 \quad (9)$$

The minimum and maximum thresholds were 0.5–1 for q_1 , $\pm 20\%$ for q_2 and s_2 and above 0.5 for s_2 performance criteria.

The different pseudo-likelihood functions were then combined to generate multi-likelihood functions with regard to model performance criteria for runoff and SWE:

$$f_q = f_{q_1} \cdot f_{q_2} \quad (10)$$

$$f_s = f_{s_1} \cdot f_{s_2} \quad (11)$$

$$f_{q+s} = f_q \cdot f_s \quad (12)$$

Posterior distributions were then generated by resampling model realizations (with replacement) with a probability given by the likelihood estimated as described above.

The 4-year study period was divided into two equally long 2-year periods. Information from the two different periods was combined to generate two separate final posterior distributions (one posterior distribution for calibration and one for validation).

The correlation of the SWE distributions, s_2 and the runoff Nash–Sutcliffe efficiency, q_1 , were then used to evaluate the different models about catchment runoff and snow distribution for the different performance criteria.

4.4. Statistical analysis

The posterior distributions of the model's performances in snow distribution (s_2) and catchment runoff (q_1) were tested for significant differences. For each of the four considered snow models (base, DC, SF and SF + DC, Section 4.2), the non-parametric Kruskal–Wallis test was used to detect whether significant differences occurred among all the posterior distributions. The post-hoc Dunn's test (Dunn 1964) was then used to identify the distribution pairs that were significantly different at the 95% significance level.

This statistical analysis was performed for both lumped and gridded HYPE model configurations in cross-calibration and cross-validation periods.

5. RESULTS

5.1. Model performance in terms of snow distribution

The results of the HYPE model performances in simulating SWE distributions differed among the snow models (Table 1) and between model spatial configurations in cross-calibration (Figure 2). The base model's performance with the lumped configuration showed high variability depending on the criteria used in calibration (Figure 2(a)). The highest model performances were found for the snow distribution criteria (s_2). The performance of the DC model with the lumped configuration showed similar results to the base model (Figure 2(a)). In contrast with the base and DC models, the performance of the SF model showed low variability among the different criteria. The highest model performances were still found for the snow distribution criterion (s_2) alone and in combination with the runoff criteria (q). Similar results were found for the model SF + DC. The base and DC models performed the worst among all the models with the lumped model configuration. In contrast, the SF and the SF + DC performed best in simulating the SWE distribution (Figure 2(a)). This result was confirmed when the snow distribution simulated by the four snow models was compared with the observations (Figure 3). The SF and SF + DC models were able to better capture the observed snow distributions compared to the base and the DC models in both calibration years 2018 and 2020. In both years, the snow distributions simulated by the base and DC

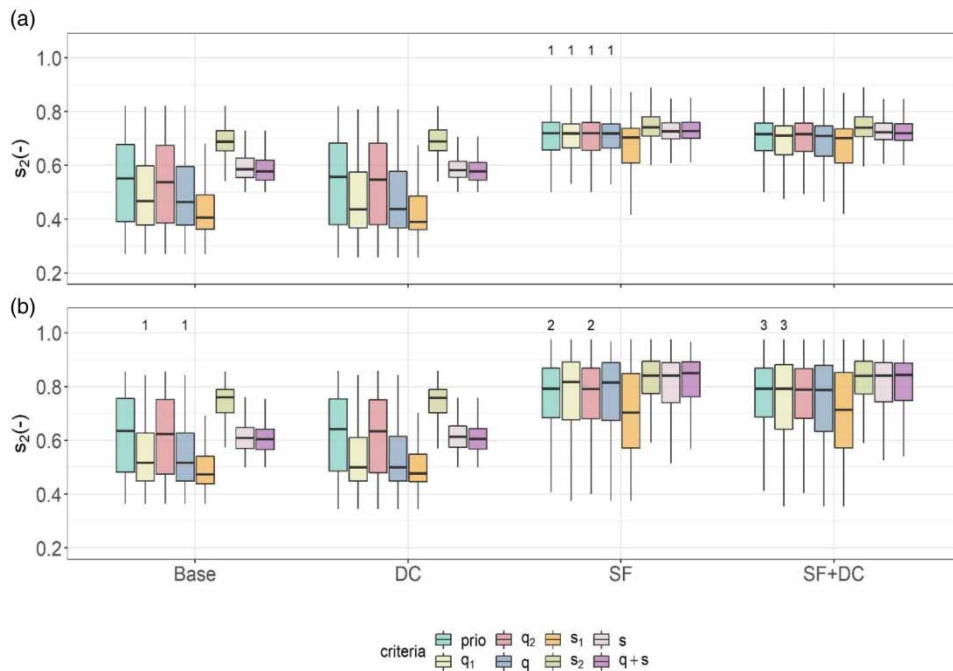


Figure 2 | Boxplot of the model performances in terms of snow distribution (s_2) for prior (prio) and posterior distributions with the different calibration criteria (q_1 to $q + s$). Model performances are reported for the four snow models (base, DC, SF and SF + DC) for the HYPE lumped (a) and gridded (b) model configurations. Numbers in each panel indicate distributions not significantly different models at the 95% level.

models were skewed compared to the snow distribution simulated by the other two models and the observations. The peak of the snow distribution was simulated by the base and DC models towards the highest SWE values compared to observations, especially in the year 2020 (Figure 3, lower panel). This SWE overestimation did not occur in the observed snow distribution.

The results with the gridded model configuration were similar to those with the lumped model configuration (Figure 2(b)). However, the performance of the SF and the SF + DC models showed a higher spread than in the lumped configuration (Figures 2(a) and (b)). The best model performances were found when models were calibrated based on the snow distribution (s_2). The combination with the runoff criteria (q), in particular with Nash–Sutcliffe efficiency (q_1), did not decrease the model's performance (Figure 2(b)).

Changes in model performances between cross-calibration and cross-validation are shown as relative percentages for the four snow models (base, DC, SF, SF + DC) and model configurations (lumped and gridded) in Figure 4. Small relative performance changes ($< \pm 2\%$) of the SF and SF + DC were found from cross-calibration to cross-validation in the lumped and gridded configurations. This result indicated similar performance in snow distribution from cross-calibration to cross-validation. The performance changes of the base and the DC models were slightly bigger than those of the SF and SF + DC models. The performance changes of the base and DC models with the calibration criteria varied more from cross-calibration to cross-validation compared to the other two models. This result is more marked with the lumped (Figure 4(a)) than the gridded model configuration (Figure 4(b)).

5.2. Model performance in terms of catchment runoff

Similar to the results of snow distributions, the HYPE model performances in simulating catchment runoff differed among snow models (Figure 5).

However, this difference is less marked than the model performances in snow distribution (Figure 2). In contrast with the results of snow distribution, similar model performances were found with both the lumped (Figure 5(a)) and gridded (Figure 5(b)) configurations for a given snow model.

Model performances also differed with calibration criteria, although these differences are less marked than for snow distribution (Figures 2 and 5). The performance of the base model was high when the model was calibrated with the runoff (q) and snow volume (s_1) criteria, not with the snow distribution criteria (s_2). The results of the DC model were similar to

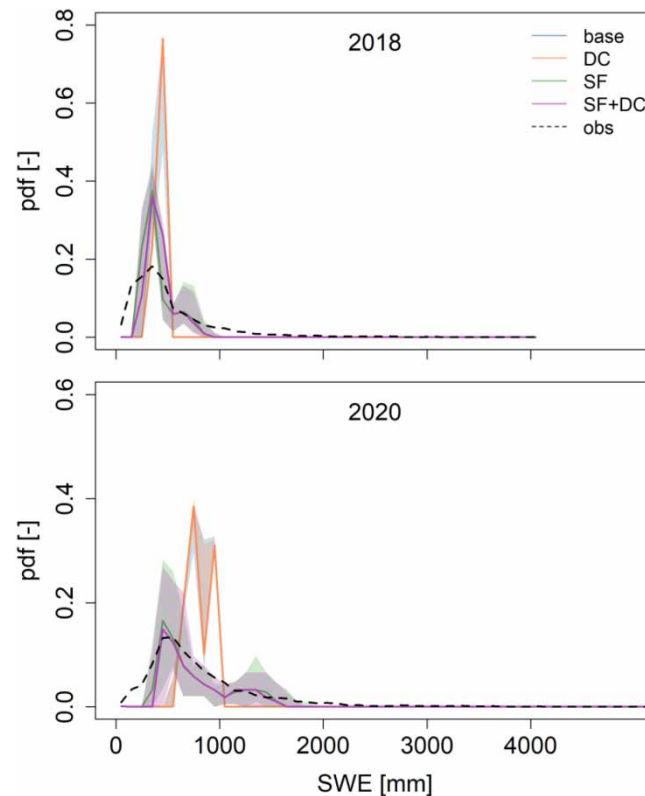


Figure 3 | Observed and simulated SWE distribution with the calibration criteria $q + s$ for the four snow models (base, DC, SF and SF + DC) for the HYPE lumped configuration.

the base model. The performance of the SF model varied less with the calibration criteria compared to the performances of the base and the DC models. In particular, the model performance did not decrease with snow distribution observations as for the other two models. The best performances of the SF were found when calibrated with snow and runoff ($q + s$). Similar results were found for the SF + DC model.

This result was confirmed when the runoff simulated by the four snow models was compared with the observations (Figure 6). All models captured the observations, particularly the spring flood in both calibration years 2018 and 2020. However, the SF and SF + DC models were able to better capture the hydrograph recession curves compared to the base and the DC models in June 2018 and August 2020, respectively. This ability is more evident in 2020, a year characterized by more snow than in 2018 (Figure 3).

Changes in model performances between cross-calibration and cross-validation are shown as a relative percentage for the different models and model configurations in Figure 7. Small relative changes ($< \pm 2\%$) in the SF and SF + DC model performances were found in simulating catchment runoff from cross-calibration to cross-validation both in the lumped (Figure 7(a)) and gridded configurations (Figure 7(b)).

The performance of the base and DC models with the lumped configuration slightly decreased in cross-validation compared to cross-calibration (Figure 7(a)). This decrease in model performance occurred with calibration criteria based on snow distribution observations (s_2). Low performances were already found with the same criteria in cross-calibration (Figure 5(a)).

6. DISCUSSION

When comparing models with different snow process representations, the best model performances in terms of snow distribution were obtained with the explicit SF function (median among criteria equal to 0.72 in Figure 8(a)), compared to the other models, e.g., base and DC (median among criteria equal to 0.57 in Figure 8(a)). This result means that the model with the SF

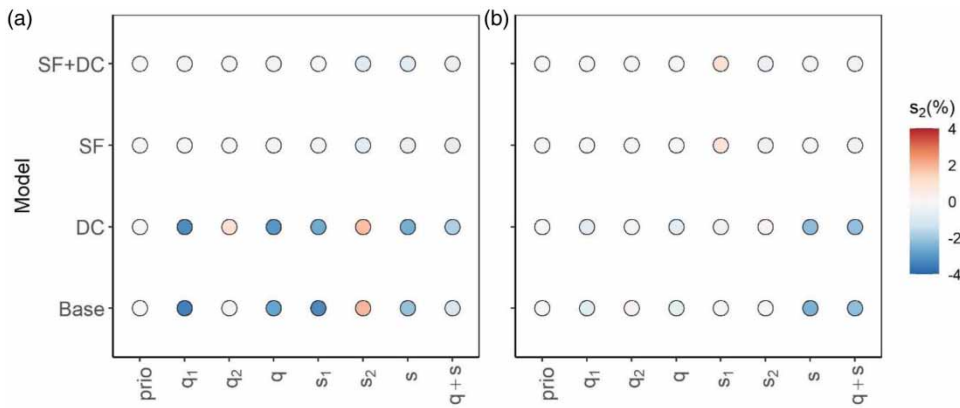


Figure 4 | Relative percentage change of model performances in terms of snow distribution between cross-calibration and cross-validation for prior (prio) and posterior distributions with the different calibration criteria (q_1 to $q + s$). Relative percentage changes are reported for the four snow models (base, DC, SF and SF + DC) for the HYPE lumped (a) and gridded (b) model configurations.

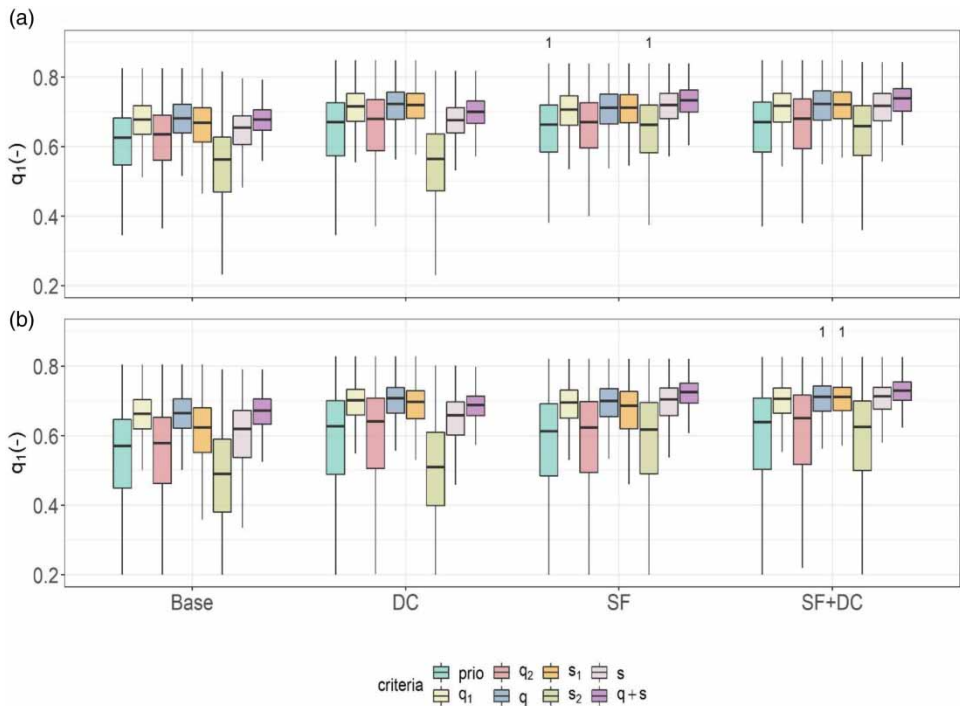


Figure 5 | Boxplot of the model performance in terms of catchment runoff (q_1) for prior (prio) and posterior distributions with the different calibration criteria. Model performances are reported for the different snow models (base, DC, SF and SF + DC) and for the HYPE lumped (a) and gridded (b) model configurations. Numbers in each panel indicate the distributions not significantly different at the 95% level.

function was able to better capture the snow distribution in the catchment close to the snow accumulation peak than the models where snow distribution was only generated with precipitation partition with temperature elevation. In contrast with the other two models, the SWE in a land class depends on the degree of exposure to wind direction through the wind shelter factor in the models with the SF function. Simulated SWE can thus vary among different elevation classes and within the same elevation class. Similar to previous studies that explored the predictive power of the wind shelter factor at the peak of accumulation in mountain areas (Schirmer *et al.* 2011; Grünewald *et al.* 2013), this factor could explain part of the snow distribution at the catchment scale. More recently, the study by Skaugen & Melvold (2019) showed instead that this factor was unfavourably compared with other terrain parameters, e.g., the squared slope, to explain the snow depth

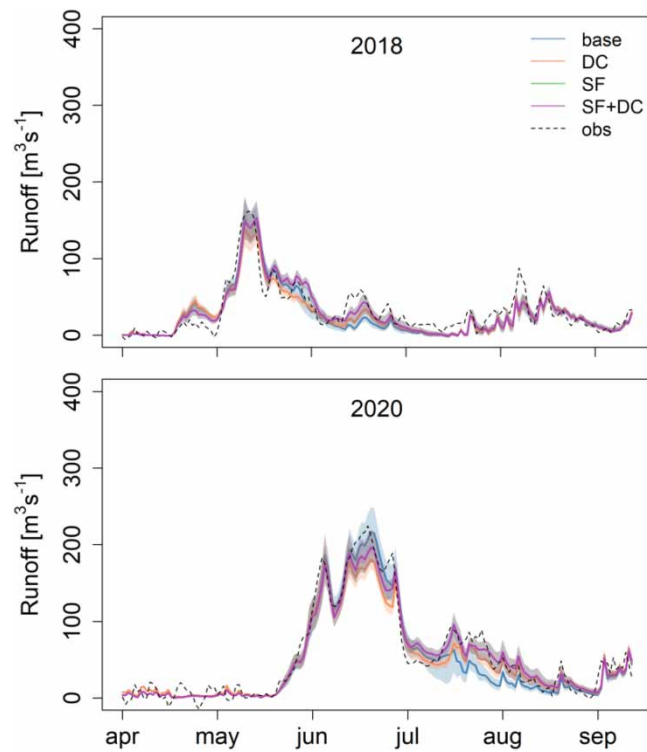


Figure 6 | Observed and simulated catchment runoff with the calibration criteria $q + s$ for the four snow models (base, DC, SF and SF + DC) for the HYPE lumped configuration.

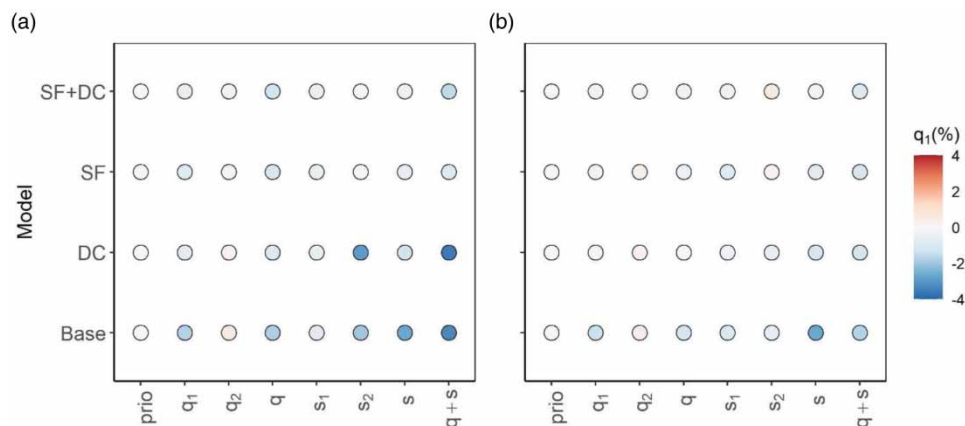


Figure 7 | Relative percentage changes of model performances in terms of catchment runoff between cross-calibration and cross-validation for prior (prio) and posterior (q_1 to $q + s$) distributions in cross-calibration. Relative percentage changes are reported for the four snow models (base, DC, SF and SF + DC) for the HYPE lumped (a) and gridded (b) configurations.

distribution at the peak of accumulation. In contrast with previous studies, the wind shelter factor in the SF function was not fixed for a prevailing wind direction but calculated for all possible wind directions and aggregated on the land use classes (Section 4.2). This dynamical integration had the advantage of considering the meteorological variability between snowfall events within and across snow seasons.

The use of distributed meteorological forcing (gridded configuration) did not increase the performances of models based only on precipitation partition with elevation, e.g., base and DC models (median among criteria equal to 0.60 in Figure 8(a)). The performances of models with the SF function increased (median among criteria equal to 0.84 in Figure 8(a)). At the same

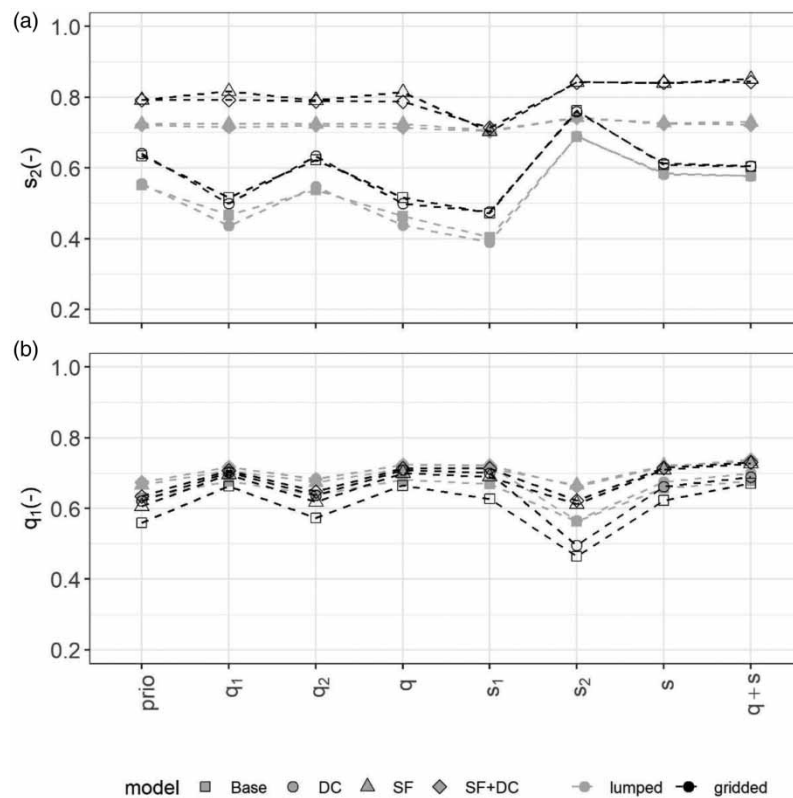


Figure 8 | Medians of the model performances in terms of snow distribution (a) and catchment runoff (b) for prior (prio) and posterior (q_1 to $q+s$) distributions in cross-calibration. Performance indicators are reported for the four snow distribution models (base, DC, SF and SF + DC) in the HYPE lumped (grey) and gridded (black) configurations.

time, the larger spread of the model performances found with the gridded (Figure 2(b)) compared to the lumped configuration (Figure 2(a)) indicated higher model uncertainty. Since the model calibration with lumped and gridded configurations was performed with the same approach, this uncertainty might result from the uncertainty in the gridded meteorological data. This uncertainty is mainly related to the precipitation data that originate from a topography-driven spatial interpolation of point gauge data relatively sparse in the model area (Johansson & Chen 2003). The distributed meteorological data provide some additional information. However, the uncertainty in model forcing for single years or grid cells can lead to increased model uncertainty, i.e., a large spread of the model performances. On the other hand, the assessment of model performance change from cross-calibration to cross-validation indicated a more stable model performance with the gridded models (Figure 4(b)) than in the lumped models (Figure 4(a)) – almost irrespective of snow model types or calibration strategies. The models without SF function, i.e., base and DC, are the most unstable from calibration to validation among the different criteria, especially for the lumped configuration (Figure 4(a)).

Regarding catchment runoff, high model performances (median among the criteria equal to 0.71 in Figure 8(b)) were found for models with the SF function (SF and SF + DC models). The base and DC models showed similar performances (median among the criteria equal to 0.71 in Figure 8(b)), except for the decreased performance for the snow distribution criteria (s_2). The models with the SF function (SF and SF + DC models) better captured both the runoff dynamics and the snow distribution in the catchment compared to the others. In contrast with the model's performances in snow distribution, the gridded configuration did not improve the runoff predictions. However, the model performance is still more stable from calibration to validation than in the lumped (Figure 7).

The analysis of the model performances thus suggests that, beyond spatial model configurations, models with the snowfall function are more suitable to predict both SWE and runoff models than those without this function. This finding is in line with the previous modelling attempts to improve snow volume estimations and thus snowmelt magnitude and timing by including snow redistribution for runoff simulations in mountainous catchments (Frey & Holzmann 2015; Freudiger *et al.* 2017).

Similar to the previous studies (Girons Lopez *et al.* 2020; Nemri & Kinnard 2020) that used snow observations in model calibration and showed an overall positive impact, in this study, model performances did not decrease in terms of runoff. In addition, models with SF function proved to be more stable across calibration criteria for snow and runoff predictions and able to include the information from distributed snow observations.

For operational applications where the objective is spring flood forecasting (e.g., for hydropower), an implicit way to represent snow distribution, such as the DC, could also be adopted.

Investigating how to use the information from the distributed snow observations to improve SWE simulations with the DC, understanding whether these results are site-specific, consistent across more snow seasons and how to transfer to other mountainous areas are relevant to the research objectives of future work.

7. CONCLUSIONS

In this study, we used different approaches to represent the spatial distribution of snow in a hydrological model and investigated their impact on SWE and runoff predictions. At the same time, we also investigated the impact of criteria such as snow volume and spatial distribution, in addition to runoff only in model calibration. Models including an SF function based on wind characteristics better predicted SWE distribution at the accumulation peak than models without.

For runoff predictions, models using the SF function and the DC showed good performances. In addition, the gridded model configuration showed similar model performances to the lumped model. To predict SWE and runoff, models with the SF function are more suitable than those without. Despite differences among the calibration criteria for the different snow process representations, the inclusion of snow observations in model calibration provided added value for snow and runoff predictions. As a general conclusion, the use of explicit representation of snow distribution, either by a spatially gridded model configuration or by the explicit SF function, provided stable improvements of the runoff simulation compared to the base model with a lumped configuration.

ACKNOWLEDGEMENTS

The authors would like to thank the editor, Bjørn Kløve, Pertti Ala-aho and an anonymous reviewer for comments and suggestions that helped to improve the manuscript significantly. The authors would like also to thank Energimyndigheten, Sweden, for supporting this study (Grant No. 46424-1).

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Beven, K. & Binley, A. 1992 *The future of distributed models: model calibration and uncertainty prediction*. *Hydrological Processes* **6**, 279–298. <https://doi.org/10.1002/hyp.3360060305>.
- Beven, K. & Freer, J. 2001 *Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology*. *Journal of Hydrology* **249**, 11–29. [https://doi.org/10.1016/S0022-1694\(01\)00421-8](https://doi.org/10.1016/S0022-1694(01)00421-8).
- Boelman, N. T., Liston, G. E., Gurarie, E., Meddens, A. J. H., Mahoney, P. J., Kirchner, P. B., Bohrer, G., Brinkman, T. J., Cosgrove, C. L., Eitel, J. U. H., Hebblewhite, M., Kimball, J. S., LaPoint, S., Nolin, A. W., Pedersen, S. H., Prugh, L. R., Reinking, A. K. & Vierling, L. A. 2019 *Integrating snow science and wildlife ecology in Arctic-Boreal North America*. *Environmental Research Letters* **14**, 010401. <https://doi.org/10.1088/1748-9326/aaec1>.
- Clark, M., Hendrikx, J., Slater, A., Kavetski, D., Anderson, B., Cullen, N., Kerr, T., Hreinsson, E. & Woods, R. 2011 *Representing spatial variability of snow water equivalent in hydrologic and land-surface models: a review*. *Water Resources Research* **47**, W07539.
- DeBeer, C. M. & Pomeroy, J. W. 2010 *Simulation of the snowmelt runoff contributing area in a small alpine basin*. *Hydrology and Earth System Sciences* **14**, 1205–1219. <https://doi.org/10.5194/hess-14-1205-2010>.
- Dunn, O. J. 1964 *Multiple comparisons using rank sums*. *Technometrics* **6**, 241–252. <https://doi.org/10.2307/1266041>.
- Freedman, D., Pisani, R. & Purves, R. 2007 *Statistics: Fourth International Student Edition, International Student Edition*. W.W. Norton & Company, New York.

- Freudiger, D., Kohn, I., Seibert, J., Stahl, K. & Weiler, M. 2017 Snow redistribution for the hydrological modeling of alpine catchments. *WIREs Water* **4**, e1232. <https://doi.org/10.1002/wat2.1232>.
- Frey, S. & Holzmann, H. 2015 A conceptual, distributed snow redistribution model. *Hydrology and Earth System Sciences* **19**, 4517–4530. <https://doi.org/10.5194/hess-19-4517-2015>.
- Gelfan, A., Gustafsson, D., Motovilov, Y., Arheimer, B., Kalugin, A., Krylenko, I. & Lavrenov, A. 2017 Climate change impact on the water regime of two great Arctic Rivers: modeling and uncertainty issues. *Climatic Change* **141**, 499–515. <https://doi.org/10.1007/s10584-016-1710-5>.
- Girons Lopez, M., Vis, M. J. P., Jenicek, M., Griessinger, N. & Seibert, J. 2020 Assessing the degree of detail of temperature-based snow routines for runoff modelling in mountainous areas in central Europe. *Hydrology and Earth System Sciences* **24**, 4441–4461. <https://doi.org/10.5194/hess-24-4441-2020>.
- Grünewald, T., Schirmer, M., Mott, R. & Lehning, M. 2010 Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment. *The Cryosphere* **4**, 215–225. <https://doi.org/10.5194/tc-4-215-2010>.
- Grünewald, T., Stötter, J., Pomeroy, J. W., Dacic, R., Moreno Baños, I., Marturià, J., Spross, M., Hopkinson, C., Burlando, P. & Lehning, M. 2013 Statistical modelling of the snow depth distribution in open alpine terrain. *Hydrology and Earth System Sciences* **17**, 3005–3021. <https://doi.org/10.5194/hess-17-3005-2013>.
- Hägmark, L., Ivarsson, K.-I., Gollvik, S. & Olofsson, P.-O. 2000 Mesan, an operational mesoscale analysis system. *Tellus A: Dynamic Meteorology and Oceanography* **52**, 2–20. <https://doi.org/10.3402/tellusa.v52i1.12250>.
- Helbig, N., van Herwijnen, A., Magnusson, J. & Jonas, T. 2015 Fractional snow-covered area parameterization over complex topography. *Hydrology and Earth System Sciences* **19**, 1339–1351. <https://doi.org/10.5194/hess-19-1339-2015>.
- Johansson, B. & Chen, D. L. 2003 The influence of wind and topography on precipitation distribution in Sweden: statistical analysis and modelling. *International Journal of Climatology* **23**, 1523–1535. <https://doi.org/10.1002/joc.951>.
- Klemes, V. 1986 Operational testing of hydrological simulation models. *Hydrological Sciences Journal* **31**, 13–24. <https://doi.org/10.1080/02626668609491024>.
- Lehning, M., Völsch, I., Gustafsson, D., Nguyen, T. A., Stähli, M. & Zappa, M. 2006 ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology. *Hydrological Processes* **20**, 2111–2128. <https://doi.org/10.1002/hyp.6204>.
- Lehning, M., Löwe, H., Ryser, M. & Raderschall, N. 2008 Inhomogeneous precipitation distribution and snow transport in steep terrain. *Water Resources Research* **44**, 1–19. <https://doi.org/10.1029/2007WR006545>.
- Lindström, G., Johansson, B., Persson, M., Gardelin, M. & Bergström, S. 1997 Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology* **201**, 272–288. [https://doi.org/10.1016/S0022-1694\(97\)00041-3](https://doi.org/10.1016/S0022-1694(97)00041-3).
- Lindström, G., Pers, C., Rosberg, J., Strömqvist, J. & Arheimer, B. 2010 Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology Research* **41**, 295–319. <https://doi.org/10.2166/nh.2010.007>.
- Liston, G. E. & Elder, K. 2006 A distributed snow-evolution modeling system (SnowModel). *Journal of Hydrometeorology* **7**, 1259–1276. <https://doi.org/10.1175/JHM548.1>.
- Liston, G. E. & Sturm, M. 1998 A snow-transport model for complex terrain. *Journal of Glaciology* **44**, 498–516.
- Luce, C. H. & Tarboton, D. G. 2004 The application of depletion curves for parameterization of subgrid variability of snow. *Hydrological Processes* **18**, 1409–1422. <https://doi.org/10.1002/hyp.1420>.
- Luce, C. H., Tarboton, D. G. & Cooley, K. R. 1998 The influence of the spatial distribution of snow on basin-averaged snowmelt. *Hydrological Processes* **12**, 1671–1683.
- Madsen, H. 2000 Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology* **235**, 276–288. [https://doi.org/10.1016/S0022-1694\(00\)00279-1](https://doi.org/10.1016/S0022-1694(00)00279-1).
- Marchand, W.-D. 2003 *Application and Improvement of a Georadar System to Assess Areal Snow Distribution for Advances in Hydrological Modeling (Dr.ing.-avhandling 2003:64)*. Fakultet for ingeniørvitenskap og teknologi. Institutt for vann og miljøteknikk, Trondheim.
- Marsh, C. B., Pomeroy, J. W., Spiteri, R. J. & Wheeler, H. S. 2020 A finite volume blowing snow model for use with variable resolution meshes. *Water Resources Research* **56**, e2019WR025307. <https://doi.org/10.1029/2019WR025307>.
- Mott, R., Schirmer, M., Bavay, M., Grünewald, T. & Lehning, M. 2010 Understanding snow-transport processes shaping the mountain snow-cover. *The Cryosphere* **4**, 545–559.
- Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *Journal of Hydrology* **10**, 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Nemri, S. & Kinnard, C. 2020 Comparing calibration strategies of a conceptual snow hydrology model and their impact on model performance and parameter identifiability. *Journal of Hydrology* **582**, 124474. <https://doi.org/10.1016/j.jhydrol.2019.124474>.
- Pomeroy, J., Essery, R. & Toth, B. 2004 Implications of spatial distributions of snow mass and melt rate for snow-cover depletion: observations in a subarctic mountain catchment. *Annals of Glaciology* **38**, 195–201. <https://doi.org/10.3189/172756404781814744>.
- Porter, C., Morin, P., Howat, I., Noh, M.-J., Bates, B., Peterman, K., Keeseey, S., Schlenk, M., Gardiner, J., Tomko, K., Willis, M., Kelleher, C., Cloutier, M., Husby, E., Foga, S., Nakamura, H., Platson, M., Wethington Jr., M., Williamson, C., Bauer, G., Enos, J., Arnold, G., Kramer, W., Becker, P., Doshi, A., D'Souza, C., Cummins, P., Laurier, F. & Bojesen, M. 2018 *ArcticDEM, Version 3*. <https://doi.org/10.7910/DVN/OHHUKH>.

- Schirmer, M., Wirz, V., Clifton, A. & Lehning, M. 2011 Persistence in intra-annual snow depth distribution: 1. Measurements and topographic control. *Water Resources Research* **47**. <https://doi.org/10.1029/2010WR009426>.
- Seibert, J. & Vis, M. J. P. 2012 Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences* **16**, 3315–3325. <https://doi.org/10.5194/hess-16-3315-2012>.
- Skaugen, T. & Melvold, K. 2019 Modeling the snow depth variability with a high-resolution lidar data set and nonlinear terrain dependency. *Water Resources Research* **55**, 9689–9704. <https://doi.org/10.1029/2019WR025030>.
- Skaugen, T. & Randen, F. 2013 Modeling the spatial distribution of snow water equivalent, taking into account changes in snow-covered area. *Annals of Glaciology* **54**, 305–313. <https://doi.org/10.3189/2013AoG62A162>.
- SMHI 2022 Available from: <https://www.smhi.se/data/utforskaren-oppna-data> (accessed 25 August 2022).
- Strömquist, J., Arheimer, B., Dahné, J., Donnelly, C. & Lindström, G. 2012 Water and nutrient predictions in ungauged basins: set-up and evaluation of a model at the national scale. **57**, 229–247. <https://doi.org/10.1080/02626667.2011.637497>.
- Sturm, M., Goldstein, M. A. & Parr, C. 2017 Water and life from snow: a trillion dollar science question. *Water Resources Research* **53**, 3534–3544. <https://doi.org/10.1002/2017WR020840>.
- Viviroli, D., Zappa, M., Gurtz, J. & Weingartner, R. 2009 An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools. *Environmental Modelling & Software* **24**, 1209–1222. <https://doi.org/10.1016/j.envsoft.2009.04.001>.
- Viviroli, D., Archer, D. R., Buytaert, W., Fowler, H. J., Greenwood, G. B., Hamlet, A. F., Huang, Y., Koboltschnig, G., Litaor, M. I., López-Moreno, J. I., Lorentz, S., Schädler, B., Schreier, H., Schwaiger, K., Vuille, M. & Woods, R. 2011 Climate change and mountain water resources: overview and recommendations for research, management and policy. *Hydrology and Earth System Sciences* **15**, 471–504. <https://doi.org/10.5194/hess-15-471-2011>.
- Whitaker, A., Alila, Y., Beckers, J. & Toews, D. 2005 Application of the distributed hydrology soil vegetation model to Redfish Creek, British Columbia: model evaluation using internal catchment data. *Hydrological Processes* **17**, 199–224. <https://doi.org/10.1002/hyp.1119>.
- Winstral, A., Elder, K. & Davis, R. E. 2002 Spatial snow modeling of wind-redistributed snow using terrain-based parameters. *Journal of Hydrometeorology* **3**, 524–538. [https://doi.org/10.1175/1525-7541\(2002\)003<0524:SSMOWR>2.0.CO;2](https://doi.org/10.1175/1525-7541(2002)003<0524:SSMOWR>2.0.CO;2).
- Winstral, A., Marks, D. & Gurney, R. 2013 Simulating wind-affected snow accumulations at catchment to basin scales. *Advances in Water Resources* **55**, 64–79.
- Wipf, S. & Rixen, C. 2010 A review of snow manipulation experiments in Arctic and alpine tundra ecosystems. *Polar Research* **29**, 95–109. <https://doi.org/10.1111/j.1751-8369.2010.00153.x>.

First received 1 April 2022; accepted in revised form 17 January 2023. Available online 7 March 2023