In search of operational snow model structures for the future – comparing four snow models for 17 catchments in Norway

Thomas Skaugen, Hanneke Luijting, Tuomo Saloranta, Dagrun Vikhamar-Schuler and Karsten Müller

ABSTRACT

In order to use the best suited snow models to investigate snow conditions at ungauged sites and for a changed climate, we have tested four snow models for 17 catchments in Norway. The Crocus and seNorge models are gridded whereas the Distance Distribution Dynamics (DDD) model with its two versions, DDD_CX and DDD_EB, is catchment based. Crocus and DDD_EB use energy balance for estimating snowmelt and SeNorge and DDD_CX use temperature-index methods. SeNorge has calibrated the temperature-index against observed snowmelt, whereas DDD_CX has calibrated the temperature-index against runoff. The models use gridded temperature and precipitation at 1 h resolution for the period 2013–2016. Crocus needs additional forcing from a numerical weather prediction model, whereas DDD_EB calculates the energy-balance elements by using proxy models forced by temperature and precipitation. The threshold temperature for solid and liquid precipitation is common for all the models and equal to 0.5 °C. No corrections of precipitation or temperature are allowed. The snow simulations are validated against observed snow water equivalent (SWE) and against satellite derived snow covered area (SCA). SeNorge and DDD_EB perform best with respect to both SWE and SCA suggesting model structures suited for describing snow conditions at ungauged sites and for a changed climate.

Key words | climate change, operational hydrology, prediction in ungauged basins, SCA, snow models, SWE

Thomas Skaugen (corresponding author)
Tuomo Saloranta
Karsten Müller
Hydrology Department,
Norwegian Water Resources and Energy Directorate,
P.O. Box 5091, Maj, 0301 Oslo,
Norway
E-mail: *ths@nve.no*

Hanneke Luijting
Dagrun Vikhamar-Schuler
The Norwegian Meteorological Institute,
P.O. Box 43, Blindern, 0313 Oslo,
Norway

INTRODUCTION

The seasonal snow cover has a significant impact on nature and human society. In Norway, where 30% of the annual precipitation falls as snow (Saloranta 2012), snow and snowfall affect hydropower planning and production, flood forecasting, transport and traffic, biological systems and tourism. The general public, the hydropower industry and public authorities are increasingly in need of more detailed information (in time and space) on snow states and the forecasting of these. Often, the information is needed for sites where measurements do not exist or are sparse. In addition, the need for information of snow conditions in a changing climate may test empirical snow model structures beyond the regime for which they were designed (i.e. calibrated).

Broadly, we can distinguish between two types of operational snow models: (1) empirical and usually calibrated models which are parsimonious both with respect to structural complexity, parameters and forcing (i.e. temperature) and (2) physically based models. A typical example of a calibrated empirical model is the temperature-index ('degree-day') model for snowmelt (Omhura 2001), where snowmelt

doi: 10.2166/nh.2018.198

is a linear function of the difference between air temperature and a (often calibrated) temperature threshold under which there is no snowmelt. Applied in hydrological models, the temperature-index factor is often calibrated against runoff, and may hence account for many other processes in addition to that of snowmelt. In general, if parameters in the snow models are not calibrated against snow, the snow simulations may compensate for errors in model structure, meteorological and runoff data. A consequence of using heavily calibrated hydrological models is that the snow in hydrological models and the snow in nature might be two different things. The value of the snow measurements hence decreases since the collected snow data cannot be used for updating and their representativeness is highly uncertain. Calibration is, of course, a challenge when the task at hand is to predict for ungauged basins and for a changed climate since no data are available. However, if the calibrated relationships represent simple, well known physical processes, do not suffer from equifinality issues and sufficient data is available to establish the relations, such procedures may still be useful.

The physically based models typically solve the surface energy balance in order to estimate snowmelt and other snow states. Given sufficient input data such as short- and longwave radiation, wind, humidity, precipitation and surface and air temperature, physically based models (for example Crocus; Brun *et al.* 1992) are in little need of calibration, but one is on the other hand faced with the difficulty of providing such data for every site of interest.

Information of snow conditions at ungauged sites and for a changing climate hence call for models that: (1) are efficient, i.e. run for areas of catchment/sub-catchment size on an operationally useful temporal scale; (2) are parsimonious with respect to forcing data, i.e. only need precipitation and temperature; and (3) have model parameters that are identifiable, are physically meaningful and constant in time and for different climates.

Through the last decades, many snow model intercomparison projects have been conducted (Slater *et al.* 2001; Strasser *et al.* 2002; Etchevers *et al.* 2004; Rutter *et al.* 2009; Essery *et al.* 2013). These projects compare models for different purposes, such as land surface schemes in numerical weather prediction models, or models for hydrology or avalanche forecasting. The models compared are of different complexity and common conclusions are: (1) that there is no 'best' model (Rutter *et al.* 2009; Essery *et al.* 2013) and (2) that model complexity does not necessarily translate into better performance (Strasser *et al.* 2002; Essery *et al.* 2013). The comparisons have been carried out for one site (Slater *et al.* 2001; Strasser *et al.* 2002; Essery *et al.* 2013), or several sites (Etchevers *et al.* 2004; Rutter *et al.* 2009) where snow information is abundant.

In the literature, we find examples of snow models focused exclusively on the dynamics of snow covered area (SCA) (Mir *et al.* 2015; Pardo-Iguzquiza *et al.* 2017) or on snow water equivalent (SWE) (Sexstone & Fassnacht 2014; Collados-Lara *et al.* 2017). The present study is novel in that it does the comparison and validation for a number of catchments for areal values of SCA and for several measurements, located at different elevations of SWE. We believe this to be an important contribution since operational hydrology needs good estimates of snow over areas of some extent spanning different elevations and landscape types (i.e. catchments with both forests and alpine areas). In addition, such a comparison takes into account the effect of the upscaling of model parameters. Upscaling may limit the usefulness of complex, data-demanding models which may prove superior at well instrumented points, but whose performance with regionalized parameters over a catchment is less known (Kirchner 2006).

In this study, we will compare the accumulation and melting of snow from four different snow models, DDD_CX (Skaugen & Weltzien 2016), DDD_EB (Skaugen & Saloranta 2015), SeNorge (Saloranta 2012, 2014), and Crocus (Brun *et al.* 1992; Vionnet *et al.* 2011). SeNorge and Crocus are dedicated snow models and DDD_CX and DDD_EB represent snow modules implemented in a hydrological model (the Distance Distribution Dynamics (DDD) model) (Skaugen & Onof 2014; Skaugen & Mengistu 2016). These models hence represent different operational needs, but they share a common goal, namely to simulate snow and snow conditions as realistically as possible. In the suite of models we find both gridded- and catchment-based models. The models further include melt algorithms which are empirical and calibrated, and more physically based. The models hence address to various degrees the prerequisites of the future snow model listed above. Since no models are calibrated either to the SCA or SWE

observations, the model comparison will inform on which model structures are best suited for operational snow modelling under a changed climate and for predictions in ungauged basins.

The snow models

The four models in the comparison differ in the algorithms for snowmelt. Common features of the models are, however, that redistribution by wind and interaction with vegetation are not taken into account. These processes are clearly important in a country like Norway, which is characterized by large areas of forests and mountains where wind is important for altering the snow distribution over many scales (Melvold & Skaugen 2013; Skaugen & Weltzien 2016), but are believed to be more dominant on spatial scales smaller than that which is used here ($\geq$1 km^2).

The DDD_CX and DDD_EB models

The DDD model is a conceptual rainfall-runoff model used by the Norwegian flood forecasting since 2013. DDD is developed with the aim of keeping the number of calibrated parameters to a minimum. The subsurface and runoff dynamics do not use parameters estimated through calibration against runoff, but are estimated using Geographical Information Systems (GIS) and recession analysis. The model keeps track of the moisture input for 10 elevations zones of equal area, so the spatial distribution, accumulation and melt of snow is carried out independently for each elevation zone. The spatial distribution of snow is parameterized from observed spatial variability of precipitation described in Skaugen & Weltzien (2016) and is hence not calibrated against runoff. DDD has been fitted with two different melt algorithms. Melt algorithm (1) is a classic temperature-index ('degree-day') model where the temperature index factor, CX (mm/°C, Δt) is calibrated against runoff.

$$M = CX \times T_a \qquad T_a > 0 \tag{1}$$

where M (mm/Δt) is the melted amount and T_a (°C) is the air temperature. This model is denoted DDD_CX. Melt algorithm (2) is described in some detail since it includes new sub-algorithms and is an energy-balance approach which is forced entirely by precipitation and temperature:

$$\mathrm{M} = \frac{(SW + L_a - L_t + H + LE + G + R - CC)}{(\lambda_F \rho_\mathrm{w})} \tag{2}$$

where M is the change in the snowpack's water equivalent (m), SW (kJm^{-2}) is the net incident solar (short wave) radiation, L_a (kJm^{-2}) is the atmospheric long wave radiation, L_t (kJm^{-2}) is the terrestrial long wave radiation, H (kJm^{-2}) is the sensible heat exchange, LE (kJm^{-2}) is the energy flux associated with the latent heats of vaporization and condensation at the surface, G (kJm^{-2}) is ground heat conduction to the bottom of the snowpack, R (kJm^{-2}) is heat added by precipitation, CC (kJm^{-2}) is the change of snowpack heat storage, λ_F is the latent heat of fusion ($\lambda_F = 335$ kJkg^{-1}) and ρ_w (1000 kgm^{-3}) is the density of water. The energy balance elements are computed using information of location and Julian day (for short wave radiation) and from using algorithms found in the substantial literature on energy-balance proxy models (e.g. Walter *et al.* 2005; Tarboton & Luce 1996, etc.). The collection of proxy models for energy balance elements is largely described in Skaugen & Saloranta (2015). Tvedalen (2015), however, pointed out several weaknesses in Skaugen & Saloranta (2015), especially concerning the estimation of atmospheric long wave radiation (L_a) or its emissivity, ε_a, cloud cover, Cl and relative humidity, Rh. These algorithms have been revised and calibrated against observed data at the Filefjell research station for snow (1,000 m a.s.l.) run by the Norwegian Water Resources and Energy Directorate (NVE) (see Figure 1 below). This model is denoted DDD_EB.

The new algorithms which differs from those described in Skaugen & Saloranta (2015) are: atmospheric emissivity,

$$\varepsilon_a = (1.0 + 0.0025{*}\mathrm{Ta}){*}(0.75 + 0.25\mathrm{Cl}). \tag{3}$$

When compared with observed data of atmospheric longwave radiation, $L_a = \varepsilon_a \sigma (T_a(K) - 273.4)^4$, where σ is the Stefan-Boltzman constant, $\sigma = 5.67 \times 10^{-11}$ (kJm^{-2}K^{4}s^{-1}), Equation (3) has half the root mean square error (RMSE) to that of the previously used expression of emissivity from Unsworth & Monteith (1975).

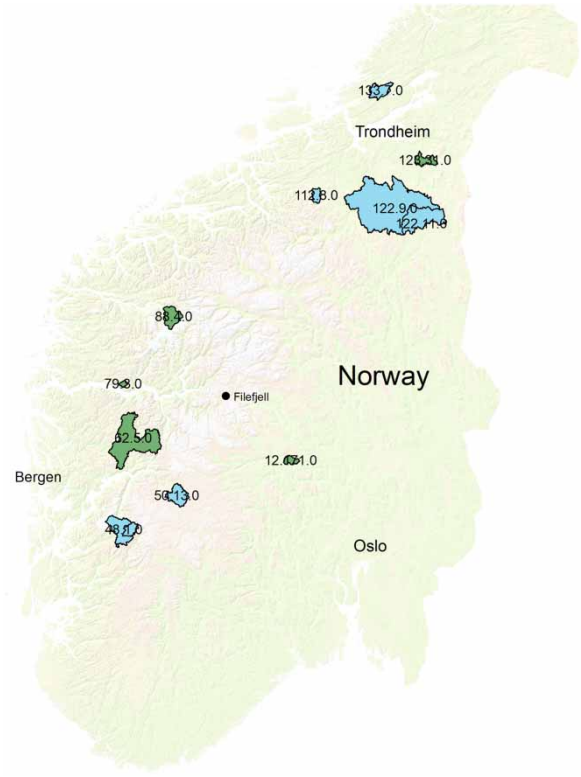

Figure 1 | Location of the study catchments in central and southern Norway and Filefjell research station. All 17 catchments are used for validation of SCA simulations, whereas catchments in light colour indicate catchments where also SWE simulations were assessed. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/nh.2018.198.

Cl is related to precipitation rate P (mm/Δt), in the following manner:

$$\begin{aligned} &\text{if } P > 1\,, \quad Cl = 1.0, \\ &\text{if } 0 < P \leq 1\,, \quad Cl = U(0.7, 1) \\ &\text{if } P = 0\,, \quad Cl = U(0.1, 0.7) \end{aligned} \tag{4}$$

where $U(min, max)$ is a random number drawn from the uniform distribution with parameters min and max parameters as indicated in Equation (4).

Rh is modelled similarly to Herrero & Polo (2012):

$$\begin{aligned} &\text{if } Cl \geq 0.7, \quad Rh = \frac{(Cl + 1.667)}{2.667}, \\ &\text{if } 0.6 < Cl \leq 0.7, \quad Rh = 0.8349 \\ &\text{if } Cl < 0.6, \quad Rh = \frac{(Cl + 1.267)}{2.667} \end{aligned} \tag{5}$$

with wind speed u, and air pressure P_a, are constants in DDD_EB having values $u = 2.5\ (\mathrm{m\ s^{-1}})$ and $P_a = 101.1$ (HPa).

The SeNorge snow model

The SeNorge snow model uses gridded observations of precipitation and temperature as input and provides a 1×1 km grid of snow conditions for all of Norway. The model consists of two sub-models, the SWE sub-model for snowpack water balance, and the snow compaction and density sub-model for converting SWE to snow depth. The model handles separately the ice and water fractions of the total SWE and keeps track of the accumulation and melting of snow. The snowpack can retain liquid water from snowmelt and rain up to a fraction of its ice content, while the excess goes to runoff. The melt rate is modelled with the extended temperature-index algorithm (Hock 2003), and is a function of air temperature and solar radiation:

$$M = b_0\, T_a + c_0 S*, \tag{6}$$

where b_0 and c_0 are estimated using multiple regression analysis against observed melt rates from the network of snow pillows operated by NVE and $S*$ is the potential extra-terrestrial solar radiation on a horizontal normalized by the maximum value at the latitude 60°N (for more details, see Saloranta 2014). The average grid cell snowmelt is further adjusted by the simulated fraction of SCA within each grid cell. The SCA algorithm assumes that snow is distributed according to a uniform distribution. The two parameters defining this distribution (minimum, maximum) are derived from the highest SWE reached so far in the snow season and a variation parameter F. Default values for F are set to 0.25 and 0.5 below and above the treeline, respectively (expert judgement).

The Crocus snow model

Crocus is a detailed snowpack model originally developed for avalanche forecasting. Crocus simulates the energy exchange within the snowpack and at its boundaries, the absorption of solar radiation with snow depth, the phase changes between solid and liquid water, water flow through

the snowpack, mass exchanges due to snow melt and metamorphism of snow. Snow albedo depend on the wavelength of incoming radiation and the type size and age of the snow surface (Strasser *et al.* 2002). Crocus uses up to 50 layers in order to simulate the internal processes and physical properties of snow and is clearly the most complex model in the ensemble. In this study, Crocus is run regionally on a 0.01° (approximately 1 km) grid within the *SURFEX* (Surface Externaliseé) interface (Masson *et al.* 2013). Once the snowpack reaches a threshold of 1 kg m^{-2} SWE within a gridpoint in Crocus, the fractional snow cover fraction is assumed to be 1. In practice this means that a grid cell is either completely covered with snow or snow free.

Model forcing and validation data

All the models are forced with 1 hourly gridded temperature and precipitation data, provided by MET Norway for the period of 1st September 2013–31st July 2016. The dataset is based on observations from the climate database of the Norwegian Meteorological Institute (eKlima 2017). For gridded precipitation and temperature, a modified Optimal Interpolation scheme is applied (Lussana *et al.* 2018). The number of observations varies in time and there are regions in Norway, typically at high elevations where the density of meteorological stations is low, which leads to an underestimation for these areas. Lussana *et al.* (2018) provide uncertainty assessments for the 24 h product to be ±20–30%, for precipitation, depending on intensity, and +0.8–2.4 °C for temperature. In Norway the number of precipitation stations measuring at a 1 hourly interval is much lower than for 24 hours. Hence, the interpolated 1 hourly precipitation grid is corrected such that, for each grid value, the hourly values summed over 24 hours matches the 24 hourly value.

From optical MODIS (Moderate Resolution Imaging Spectroradiometer) satellite scenes, a time series of SCA has been retrieved for elevation bands for the study catchments. Each 500 × 500 meter grid cell is assigned a SCA value between 0 and 100% snow coverage using a method based on the Norwegian linear reflectance to snow cover (NLR) algorithm (Solberg *et al.* 2006). The input to the NLR algorithm is the normalized difference snow index (NDSI) signal (Salomonson & Appel 2004). Each catchment has 10 elevation zones of equal area. In total 691 satellite scenes over the period 2014–2015 are used which give us 7,601 data points (i.e. catchment mean and 10 elevation zones). Only scenes with less than 10% cloud cover are used.

SWE data are collected routinely at the peak of the accumulation season and sometimes also at mid-winter by hydropower companies. The usual procedure is to measure snow depth every 10–20 meters at a snow course of length 1–2 km, at a certain elevation, and snow density at 2 or 3 representative points (e.g. where snow depth is close to the average). For this study we use the average SWE value measured in the actual, or in a nearby catchment. 193 such averaged SWE values, collected during the period 2014–2016 are used in this study.

The hydrological models are calibrated against hourly discharge data provided by the HYDRAII database of NVE. The measured water stage is transformed to runoff using estimated stage-runoff relationships. Many of the catchments have problems of ice clogging the cross-section where the water stage is measured during winter. The observed daily runoff is corrected manually if suspicious values due to ice clogging appear. The hourly values are then corrected automatically so that hourly values summed over 24 hours match the corrected 24-hour value. This production chain produces uncertainty at every step, so that winter runoff at mountainous catchments in Norway must be treated with caution. In addition to precipitation and temperature, the Crocus model needs additional forcing data (relative humidity, incoming short wave and long wave radiation, wind speed and wind direction). Forcing for these variables were derived from the numerical weather prediction model AROME MetCoOp (AROME) used operationally by MET Norway (Luijting *et al.* 2017).

Study area

The model comparison is carried out for catchments located in central- and mid-Norway. 17 catchments are evaluated for the models ability to simulate SCA, whereas a subset of 10 catchments are evaluated for SWE. The gridded models, SeNorge and Crocus, simulate regionally or nationally for areas which include these catchments, whereas DDD_EB and DDD_CX simulate for the catchments and their

elevation zones. Figure 1 shows the location of the catchments and Table 1 provides geographical information in addition to average annual maximum SWE. The average annual maximum SWE is the simulated catchment value over 58 years as simulated by the HBV model (Hydrologiska Byråns Vattenbalans model; Bergström 1992) which is, together with DDD, used at the Norwegian flood forecasting service at NVE. The HBV model is calibrated against daily runoff and precipitation is adjusted too so that the volume error is minimized (no bias). The values of average annual maximum SWE are hence not necessarily the truth, but indicate the snow regime associated with each catchment.

Set up of model comparison

All models use a threshold temperature of $T_x = 0.5°C$ to separate between solid and liquid precipitation. When applicable (for SeNorge and DDD_CX), the threshold temperature for snowmelt is set equal to $T_S = 0$. All models use the hourly gridded observations of precipitation and temperature as input data. SeNorge aggregated the 1 hourly data to 3 hourly data. This, because SeNorge is an operational tool providing gridded snow information for all of Norway and it is not set up to run on a 1 hourly time interval since such a fine temporal resolution does not, at the moment, meet any operational needs.

Table 1 | Geographical information of the study catchments

Catchment	Area (km^2)	Mean elevation (m)	Min-max elevation (m)	Max average SWE (mm)
2.268*	796.3	1,466	475–2,462	1,076
12.171	79.8	903	780–1,203	406
48.1*	470.2	1,091	87–1,652	1,953
48.5*	120.5	1,231	595–1,635	2,123
50.13*	262.6	1,248	1,010–1,536	1,291
62.5	1,092.0	868	47–1,602	1,353
75.23*	45.9	1,147	17–1,467	1,472
77.3*	110.9	1,002	395–1,600	1,886
79.3	30.0	820	288–1,346	1,776
83.2	508.1	841	144–1,635	1,591
88.4	234.9	1,339	52–2,076	1,558
109.9	745.4	1,347	556–2,283	604
112.8*	86.23	816	467–1,591	1,292
122.11*	654.2	843	285–1,284	994
122.9*	3,086.4	734	54–1,326	790
123.31	145.0	578	200–1,166	1,421
133.7*	205.9	349	87–627	947

*Catchments marked with an asterisk have measurements of SWE within or close to the catchment. Maximum averaged SWE is the simulated catchment value over 58 years as simulated by the HBV model used in the Norwegian flood forecasting service at NVE.

During the calibration of the hydrological model DDD_CX, there was no possibility to adjust precipitation and temperature in order to improve calibration results or to better close the water balance. In using such a set-up for comparison, we believe that we can focus the comparison on the snow algorithms of the different models and better diagnose the results. In addition, such a set-up mimics the situation faced when we need to estimate the snow conditions at ungauged sites and for a changed climate where only scenarios of meteorological forcing is at hand.

With imperfect input, such as underestimated precipitation, we expect the models to underestimate SWE and possibly also SCA. For DDD_CX this is hazardous since through calibration, all model parameters will be optimized in order to try to make up for the deficiency provided by the underestimation of precipitation. The number of parameters to be calibrated against runoff for DDD_CX is six. Table 2 show the calibration parameters of DDD_CX in addition to parameters common to all the models and we note that two of the calibration parameters are associated with snow, *Pro* and *CX*, but through the calibration process these parameters may have attained values not necessarily associated with snow processes.

DDD_CX was calibrated using the Particle Swarm Optimization and Dimensioned Search (PPSO) method (https://www.rforge.net/ppso/) for the period 1st September 2014–31st December 2015. As can be seen from Table 2, DDD_EB uses model parameters calibrated with DDD_CX, except, for the parameter CX which obviously does not apply when using the energy-balance approach for snowmelt.

RESULTS

Simulating runoff

Table 3 shows the result for the calibration of DDD_CX and comparison with DDD_EB. Crocus and SeNorge are

Table 2 | Calibration parameters in DDD_CX and parameters common for several models

Parameter name	Parameter function	Value, range and mean
Pro(%)[a,b]	Max liquid water content in snow (fraction of SWE)	1–15 (6)
CX (mm/°C, Δt)[a]	Temperature index for snow melt	0.1–0.44 (0.22)
CGlac (mm/°C, Δt)[a,b]	Temperature index for ice melt	0.27–0.56 (0.44)*
Cea (mm/°C, Δt)[a,b]	Temperature index for evapotranspiration	0.001–0.011 (0.007)
Rv (m/s)[a,b]	Celerity for water transport in river	0.3–3.59 (1.89)
Gtcel (%)[a,b]	Saturation threshold for overland flow	86–99 (96)
T_x (°C)[a,b,c,d]	Temperature threshold for solid and liquid precipitation	0.5
T_S (°C)[a,c]	Temperature threshold for snow melt	0.0

Notes: The range and mean (in brackets) are based upon 17 catchments, except for the temperature index for glacier melt, CGlac. Only three catchments have a glacier fraction above 5% and the range and mean (marked with *) are based upon these catchments. [a]DDD_CX, [b]DDD_EB, [c]SeNorge and [d]Crocus model uses the parameter.

Table 3 | Skill scores for the calibration of DDD_CX and comparison with DDD_EB 1 h, TX = 0.5, TS = 0, and the period is 1.9.2014–31.12.2015

Catchment	KGE_CX	KGE_EB	Bias_CX	Bias_EB
2.268*	0.823	0.86	0.968	0.694
12.171	0.685	0.77	0.97	1
48.1*	0.822	0.825	0.92	0.882
48.5*	0.768	0.758	0.798	0.795
50.13*	0.825	0.801	0.842	0.852
62.5	0.795	0.817	0.899	0.905
75.23*	0.564	0.523	0.593	0.614
77.3*	0.664	0.692	0.865	0.781
79.3	0.599	0.595	0.684	0.677
83.2	0.614	0.641	0.783	0.742
88.4	0.906	0.824	0.987	0.857
109.9	0.822	0.647	1.086	1.219
112.8*	0.733	0.66	0.801	0.804
122.11*	0.586	0.76	0.794	0.838
122.9*	0.618	0.821	0.988	1.004
123.31	0.524	0.494	0.573	0.588
133.7*	0.572	0.45	0.702	0.693
Mean	0.701	0.702	0.838	0.820

*Catchments marked with an asterisk have measurements of SWE within or close to the catchment. Bias is measured as the ratio between mean simulated runoff and mean observed runoff.

dedicated snow models and are therefore not validated against runoff. We can note from the mean values of the statistics that the two models give very similar results with the Kling-Gupta efficiency criterion (KGE; Kling *et al.* 2012) and bias. The estimate of bias, the ratio between simulated and observed runoff shows that, with a few exceptions, the meteorological grids seem to underestimate precipitation with, on average, nearly 20%. This finding corresponds with those of Lussana *et al.* (2018) as reported above. The skill scores are, in general, not very high, although 5(6) out of the 17 catchments have KGE values over 0.8 for DDD_CX (DDD_EB). Several reasons for this can be put forward and lack of ability to correct the precipitation input to close the water balance together with high uncertainty in the runoff data are the two most prominent candidates.

Simulating SWE

Ten of the catchments (marked with an asterisk in the tables and with blue colour in Figure 1) have SWE measurements within or close to the catchments. Of the catchments with snow measurements, all catchments underestimate the runoff and some to a quite serious degree, with the exception of catchment 122.9. This is the case especially for catchments 133.7 and 75.23, which also have very low KGE. Figure 2 shows the mean absolute error (MAE):

$$MAE = \frac{\sum |SWE_s - SWE_o|}{n}, \tag{7}$$

where SWE_s and SWE_o are simulated and observed SWE, respectively, and n is the number of observations. In Figure 2, MAE of simulated SWE by the different models is organized by elevation (m a.s.l.) (Figure 2(a)) and by catchment (Figure 2(b)). We note from Figure 2(a) that the error clearly increases with elevation. This is partly reflected in Figure 2(b) where we find the highest errors for catchments with high mean elevation and also high mean annual maximum SWE (see Table 1). The exception is catchment 2.268 which is located at high elevation but have relatively small

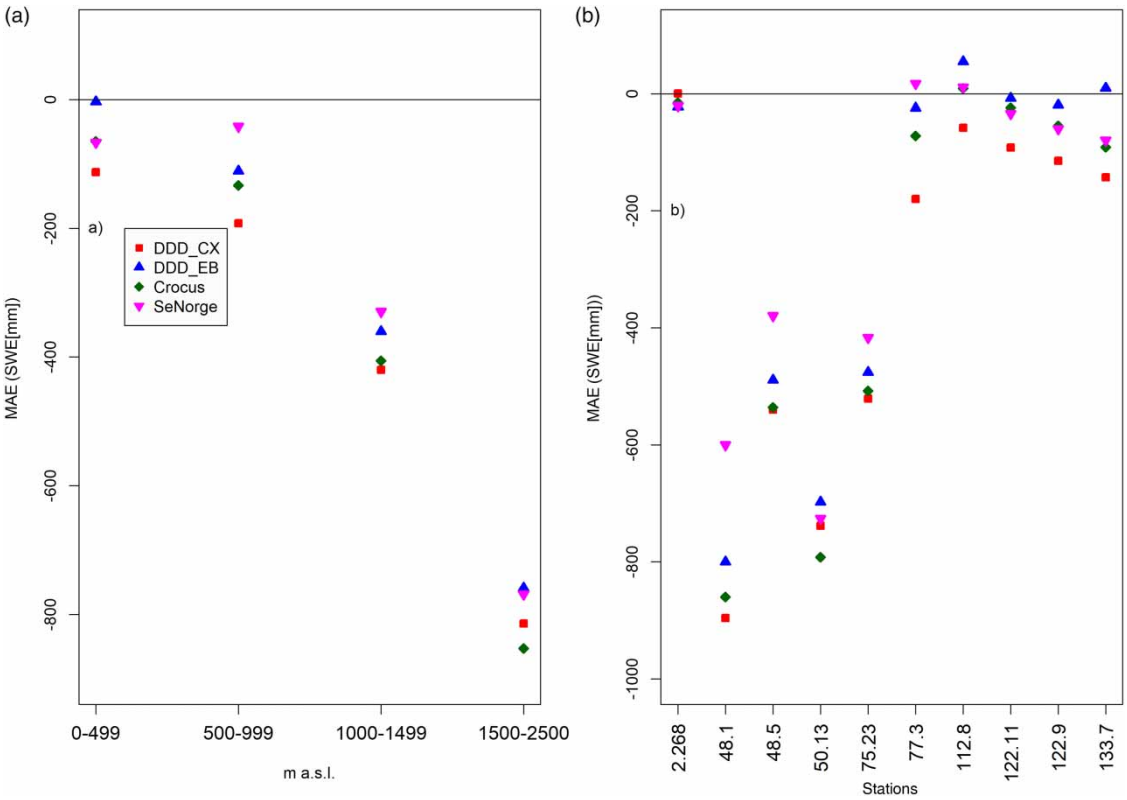

Figure 2 | Mean absolute error (MAE) of SWE simulations organized by (a) elevation and by (b) catchment.

errors by all models. However, the mean annual maximum SWE for 2.268 is only 50% to that of, for example 48.1 and 48.5, so one would expect smaller errors.

The pattern shown in Figure 2 is repeated for the RMSE of simulated SWE (not shown).

$$RMSE = \left(\frac{\sum (SWE_s - SWE_o)^2}{n} \right)^{0.5} \quad (8)$$

Figure 3 shows simulated and observed SWE for two elevation zones for the catchment 48.5. From Tables 1 and 3 we see that the catchment 48.5 is characterized by high mean elevation (1,231 m a.s.l.), high average annual maximum SWE (2,123 mm) and an underestimation of runoff by approximately 20%. The elevation zones represents elevations of 1,174–1,231 m a.s.l. (Figure 3(a)) and 1,232–1,278 m a.s.l. (Figure 3(b)). In general we see that simulated SWE is less than the observed SWE and is consistent with the reported underestimation of the precipitation grid. It must be noted that the observation is based upon one single snow course and must be considered to be an uncertain estimate of SWE in that elevation zone. SeNorge is closer to the observed SWE than the other models. The other models are quite similar, although Crocus underestimates the most. Figure 4 shows simulated and observed SWE for two elevation zones for the catchment 122.9. The elevation zones represent elevations of 734–811 m a.s.l. (Figure 4(a)) and 812–877 m a.s.l. (Figure 4(b)). Catchment 122.9 has a mean elevation of 734 m a.s.l., average annual maximum SWE of 790 mm and a very small bias in the hydrological simulations. We observe that the simulated and observed SWE are comparable for most years. Crocus, DDD_EB and SeNorge are quite similar, whereas DDD_CX underestimates SWE for these elevation zones and for the catchments in general (see Figure 2(b)).

Simulating SCA

From the 691 scenes of satellite derived SCA, we could abstract between 25 (133.7) up to 53 (48.5) values of

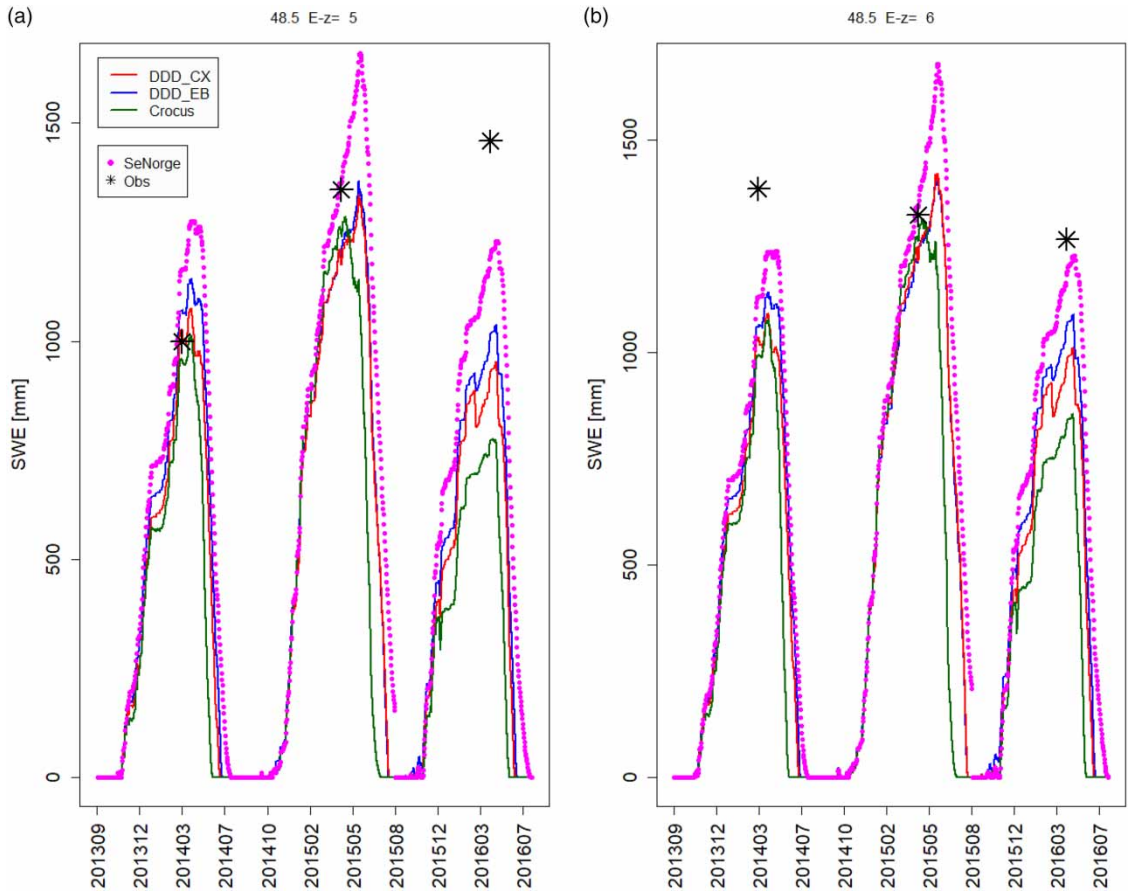

Figure 3 | (a) Simulated and (b) observed SWE for two elevation zones for catchment 48.5.

SCA for each catchment and their elevation zones for the two snow seasons 2013/14 and 2014/15. Figure 5 shows the MAE for SCA organized by elevation and by catchment. We can see that the errors in estimating SCA are not sensitive to elevation. What is striking is the clear difference in performance between Crocus and DDD_CX on the one hand and SeNorge and DDD_EB on the other hand. The two latter models have very similar performance with MAE quite close to zero and low RMSE (figure not shown). Figures 6 and 7 show simulated and observed catchment values of SCA for the two catchments 48.5 and 122.9. From Figure 6 (catchment 48.5) we see that both DDD_EB and SeNorge simulates the catchment SCA very well. DDD_CX underestimates the SCA in 2014 but has a similar performance to DDD_EB in 2015. For both years Crocus has a too steep decline in SCA and starts the melting too early. This latter feature is common for all the models except DDD_EB. From Figure 7 (catchment 122.9) we see that the observed SCA has a more varied behaviour during the snow season which is not captured very well by the models. DDD_EB is overestimating SCA, whereas SeNorge, Crocus and DDD_CX underestimate SCA in the melting period of 2014. All models capture the timing of melt-out reasonably well.

Summary of results

Table 4 shows the overall results of validating the models against SWE and SCA. For both validations the SeNorge and DDD_EB emerge as the best models. We see for example that on average, the error of SeNorge for SWE is 31% less than that of DDD_CX.

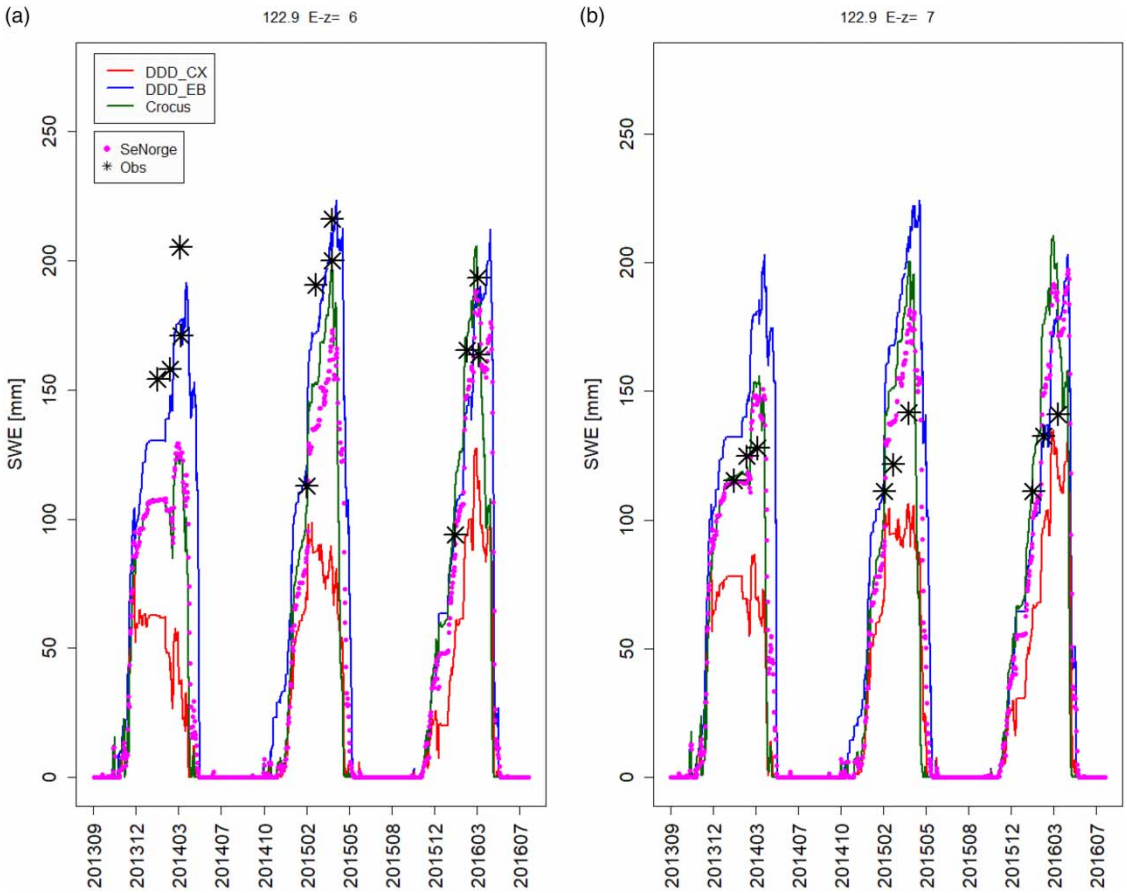

Figure 4 | (a) Simulated and (b) observed SWE for two elevation zones for catchment 122.9.

DISCUSSION

The aim of this study is to learn what principles in snow modelling model are best suited for operational snow modelling under a changed climate and for predictions in ungauged basins. When validating the models against observed SWE and SCA, SeNorge and DDD_EB, quite convincingly stand out as best suited, and especially so for the validation against observed SCA. These two models represent the most commonly used snow melt algorithms; the temperature index (although extended with a radiation term) and the energy balance approach (i.e. Equations (6) and (2)). The inferior models, DDD_CX and Crocus also use these snow melt algorithms, so, clearly, the choice of empirical/physically based approach has not been the key issue for differences in results, but more on how these methods are applied. We have set rather strict rules for the model comparison set-up, with the intention of isolating the workings of the models' snow algorithms. In addition, such a set-up is intended to be as close to situations of predicting for ungauged basins and for effects of climate change as possible.

Crocus is quite demanding in respect of forcing variables. In this study, these forcing variables are abstracted from AROME. One could argue that using the interpolated meteorological grids of observed precipitation and temperature instead of the physically consistent precipitation and temperature from AROME would provide better results. However, in a recent study by Luijting *et al.* (2017), simulated snow depths were compared against observed and it was quite clearly shown that using Crocus with the interpolated meteorological grids performed much better than with only using forcing from AROME. In this study, Crocus was originally intended to be the benchmark model, the role Crocus

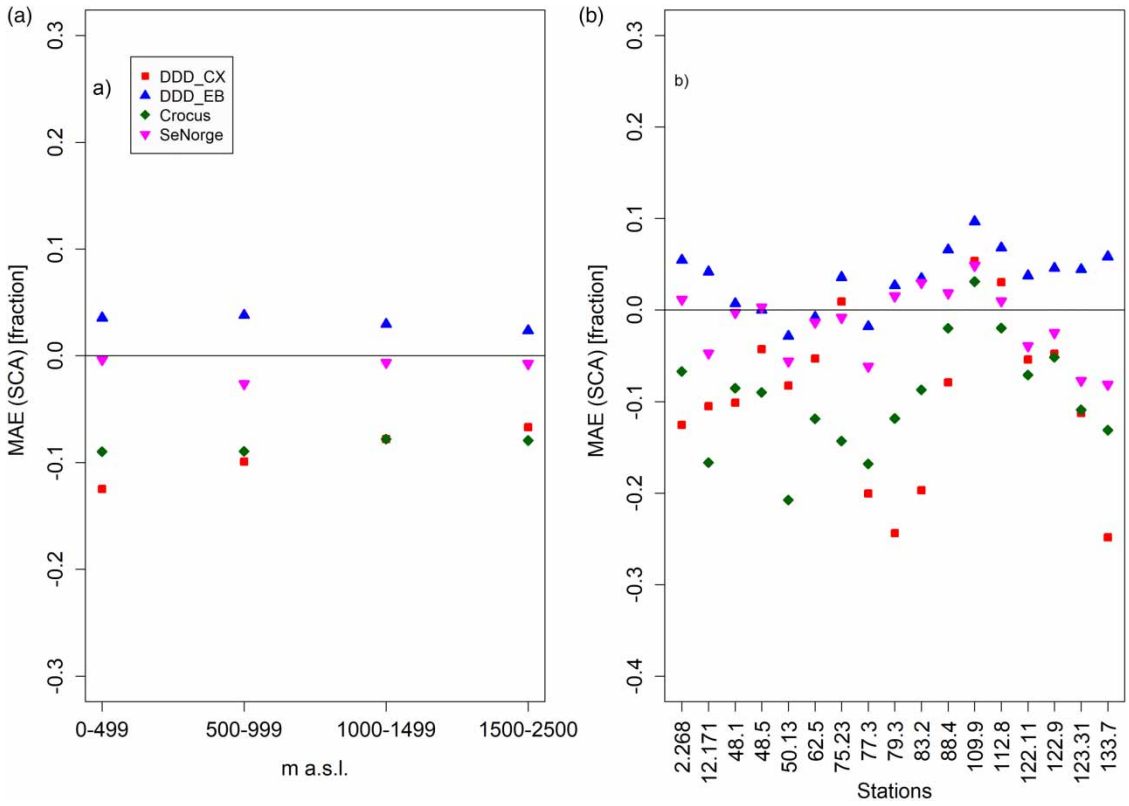

Figure 5 | MAE for estimated SCA organized by (a) elevation and by (b) catchment.

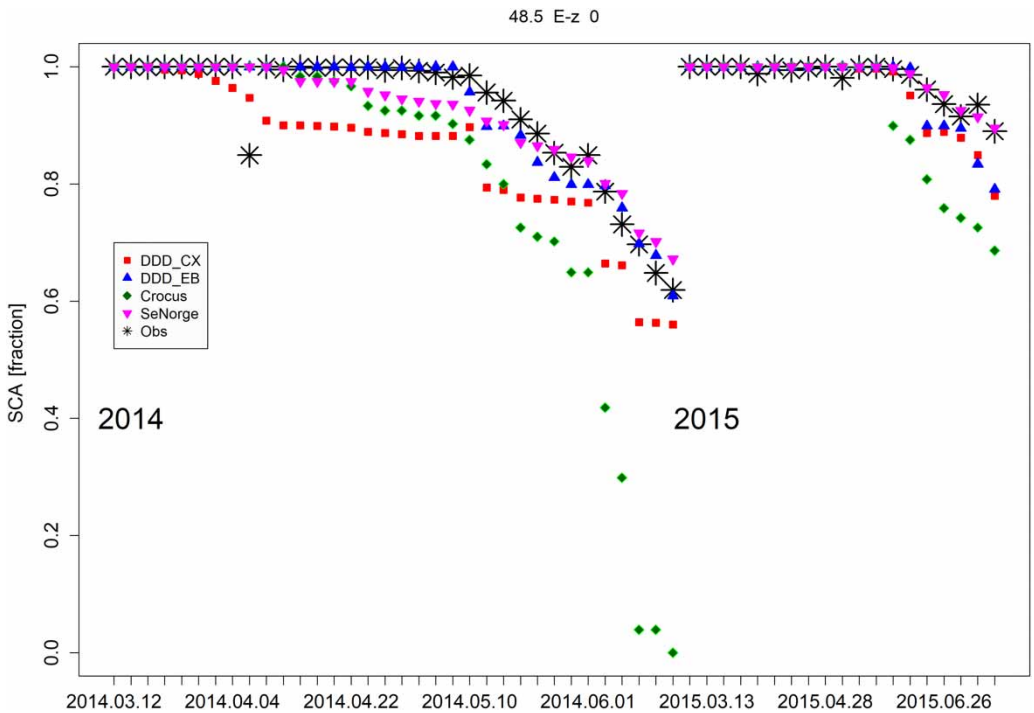

Figure 6 | Catchment values of simulated and observed SCA for catchment 48.5 (120.5 km^2) for 2014 and 2015.

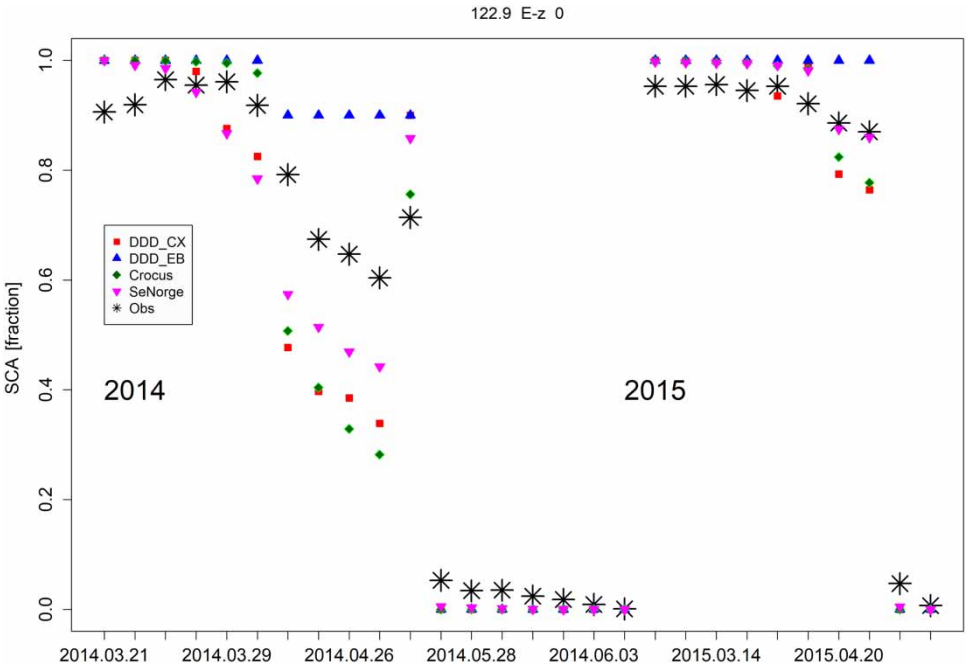

Figure 7 | Catchment values of simulated and observed SCA for catchment 122.9 (3,086.4 km^2) for 2014 and 2015.

Table 4 | Summary of results

	SWE		SCA	
Model	**MAE (mm)**	**RMSE (mm)**	**MAE (%)**	**RMSE (%)**
SeNorge	−260.0 (69%)	328.1 (80%)	−2.0 (18%)	10.0 (50%)
DDD_EB	−302.2 (80%)	364.2 (89%)	4.0 (36%)	12.0 (60%)
Crocus	−343.8 (91%)	389.3 (95%)	−9.0 (85%)	13.0 (65%)
DDD_CX	−378.9 (100%)	409.6 (100%)	−11.0 (100%)	20.0 (100%)

Notes: The MAE and RMSE values are averaged over all observations and catchments. The number in brackets represent the percentage of error compared to the worst performing model (DDD_CX).

has had in many model comparison studies (Strasser *et al.* 2002; Essery *et al.* 2013). However, this study differs from previous ones in that we have run Crocus regionally and conducted the validation on spatially averaged values from elevation zones of catchments and entire catchments, and not at points as has been customary. The spatial distribution of snow over an area has been the focus of several investigations and is seen as crucial for the dynamics of snowmelt, the evolution of SCA and for properly accounting for the energy fluxes in land-surface schemes (Essery & Pomeroy 2004; Helbig *et al.* 2015). Crocus, being a point model, does not aim especially at representing a realistic spatial distribution of snow, which is in contrast to the other models.

It may be instructive to look at the behaviour of the models in more detail in order to better understand why a well-established model such as Crocus shows deficiencies in simulating at a regional scale. Figure 8 shows accumulated melt rate for the catchment 122.9 for the winter 2013/14 (Figure 8(a)) and the winter 2014/15 (Figure 8(b)).

We see from Figure 8 that Crocus tends to start melting early for both years, and has finished melting earlier than all other models. In order to further investigate the models' discrepancies we have, in Figures 9 and 10, plotted net radiation and albedo as simulated by Crocus and DDD_EB (DDD_CX and SeNorge do not simulate such elements being temperature-index based models). Figure 9 shows that the general pattern of net radiation for DDD_EB and Crocus is quite similar except that the net radiation of Crocus increases much earlier than DDD_EB. This can be explained by the differences in albedo shown

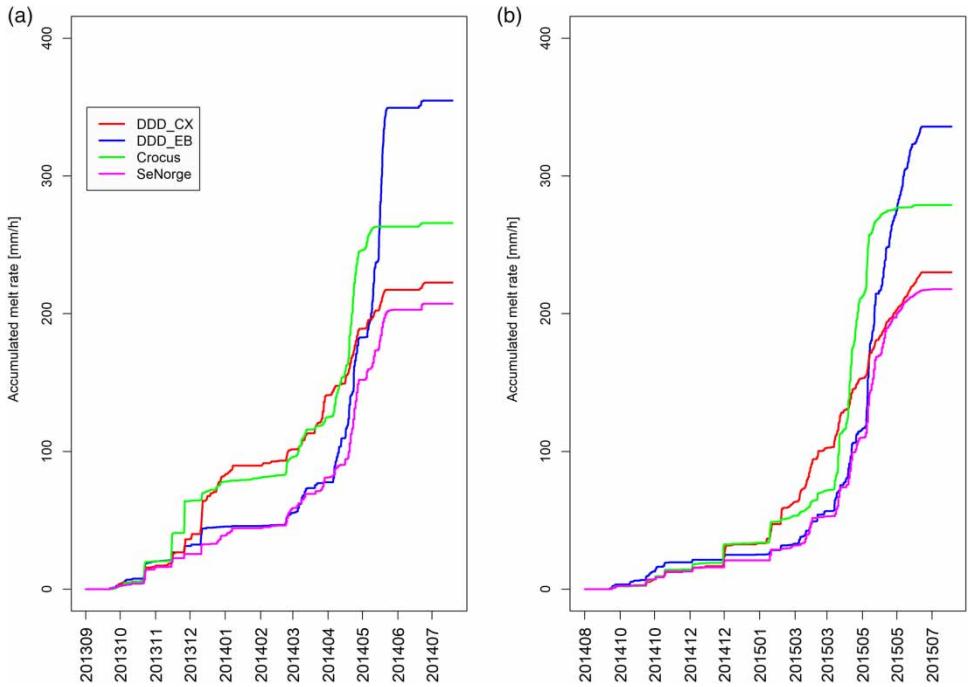

Figure 8 | Accumulated melt rate (mm h^{-1}) for catchment 122.9 for the winters (a) 2013/14 and (b) 2014/15. Horizontal lines at the end of the season indicate melt-out.

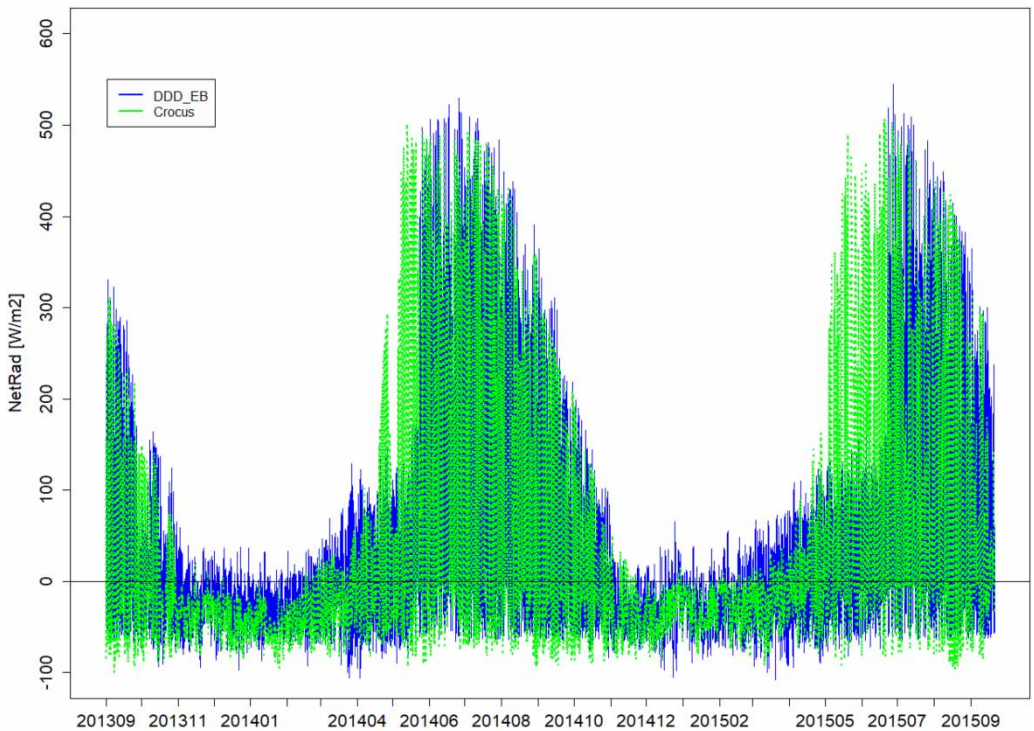

Figure 9 | Net-radiation for elevation zone 10 (1,018–1,326 m a.s.l.) for the winters 2013/14 and 2014/15 for catchment 122.9 as simulated by Crocus and DDD_EB.

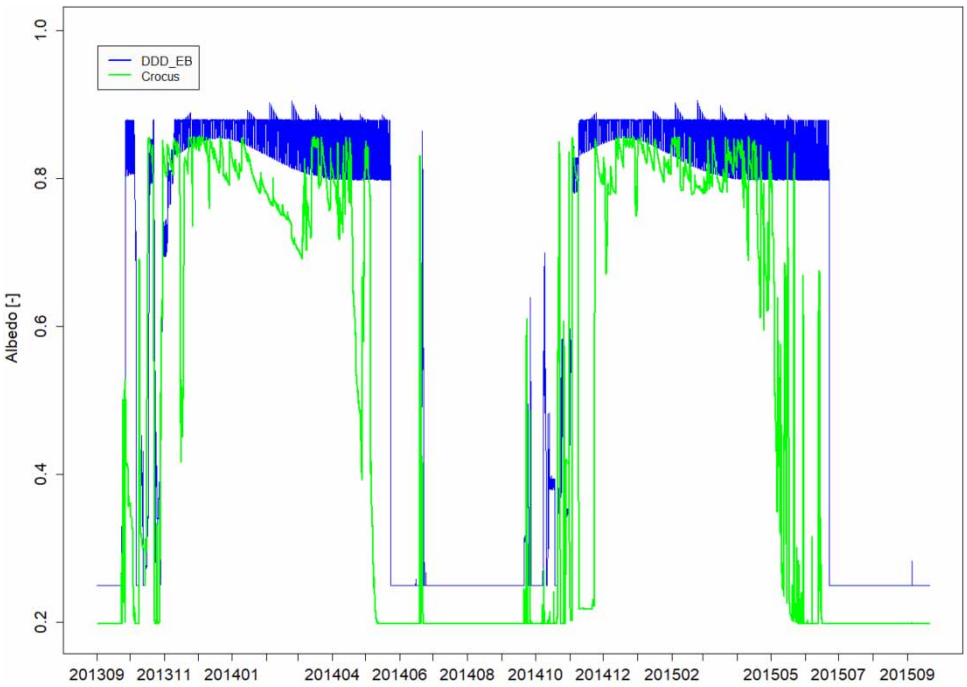

Figure 10 | Albedo for elevation zone 10 (1,018–1,326 m a.s.l.) for winters 2013/14 and 2014/15 for the catchment 122.9 as simulated by Crocus and DDD_EB.

in Figure 10 and suggests that the difference in net radiation is caused by increased SW radiation of Crocus. The loss in albedo of Crocus is consistent with the higher fraction of snow-free areas found early in the season for the Crocus simulations. We also note that, in general, the albedo of Crocus is lower than that of DDD_EB which suggests increased melting. At this point, we do not find any obvious reasons as to why the melt starts so early for Crocus and it is beyond the scope of this study to investigate and perhaps trace errors to the AROME input. This study represents a first attempt at regional simulations by Crocus in Norway, and further investigations will have to address questions such as the need for a more detailed land-surface description and the scaling up of model parameters. Figure 10 also shows that the albedo of DDD_EB displays some strange temporal fluctuations. The albedo seems to fluctuate mainly between 0.8–0.87. DDD_EB uses the albedo algorithms of the Utah Energy balance Snow Accumulation and Melt Model (Tarboton & Luce 1996), and clearly the implementation needs adjustment. The most probable cause is associated with the illumination angle of radiation. The effects of the errors, however, are not large enough to seriously affect the results of SWE and SCA simulations.

The results of DDD_CX present a very convincing case of the problems encountered in calibrating a parameter (CX) as a member of a set of parameters that simultaneously are to be calibrated against a variable (i.e. runoff). The relationship between parameter and variable may vary over time and, in addition, non-physical combinations of parameters may work together to produce good results for the variable. The result can be that we may have sensible, reliable results as well as results where, for example, the simulation of snow variables have been sacrificed in order to provide good results for runoff. In Figures 2 and 5 we see that DDD_CX quite often show the poorest performance for simulating SWE and SCA, and the summary of results in Table 4 points at DDD_CX as the inferior model for simulating both SWE and SCA. Figure 4 shows how SWE is consistently underestimated by DDD_CX compared to the other models and to observations, whereas in Figure 3, the DDD_CX simulations do not differ much from the other models. The process parameterization itself (temperature-index) appears to be quite capable of providing good results for the simulation of snowmelt (see also the results of the SeNorge model), but the procedure of calibrating the temperature-index (as part of a parameter set) against runoff

instead of against snow observations is hazardous and cannot be recommended. It is a paradox that such a calibration set-up has been the usual practice in operational hydrology in the Nordic countries for decades, although a wealth of snow measurements exists which could be used for calibration of the temperature-index directly.

DDD_EB shows surprisingly robust results across the 17 (10) catchments for SCA and SWE, when we take into consideration that only precipitation and temperature is used to force the energy balance calculations. The catchments represent a variety of climatic regimes of Norway from relatively dry inland catchments (2.268 and 12. 171), moist lowland catchments (122.9, 133.7) and to alpine and moist (located on the west coast) catchments. It is promising, with respect to the overall aim of providing model structures useful for predicting in ungauged basins and for a changed climate, that the proxy models used to calculate the energy balance seem to work well for such different areas. We can also note, from Table 3 that there is no significant difference in skill for simulating runoff between DDD_CX and DDD_EB, indicating that the introduction of the proxy models for energy balance calculations, and the associated loss of one calibration parameter (CX), did not deteriorate the runoff predictions of the DDD model.

The SeNorge model shows the best performance of all the models. The fact that it is run at a 3-hourly resolution instead of 1-hourly as the other models, may account for the results being less sensitive to high variability in the meteorological input (precipitation and temperature) but the validation against data for which it has not been calibrated shows that the model is able to capture the essential snow characteristics at different regions and in different climates in Norway. The results of SeNorge also demonstrate that simple empirical relations such as the extended temperature index method work well and can be used regionally if calibrated against relevant data (snow melt measurements).

Summarizing model limitations

The Crocus model is primarily an alpine model and needs additional forcing data for simulating snow. In this study, Crocus simulates regionally over many different landscape types and the additional forcing, provided by the numerical weather prediction model AROME, may not be of sufficient quality to give us the full potential of the Crocus model. The temperature-index approach in the DDD_CX model represents a very common approach of simulating snowmelt in operational hydrological models. It has been clearly demonstrated in this study that calibrating the temperature index factor for snowmelt against runoff is not optimal for simulating the coverage and dynamics of SWE. The proxy models for estimating the energy balance in DDD_EB have shown quite acceptable results although the approach is coarse, i.e. neglecting spatial variability in topography and landscape types and temporal variability in wind speed and air pressure. There might be dependencies between the choice of the constants in the in the model and the chosen temporal resolution (1 h) that have to be investigated. The SeNorge model applies the extended temperature-index method which includes a radiation term. The snow melt algorithm is calibrated against observed snow melt rates which is both reasonable and gives good results. In addition, the models discriminate between areas below and above the tree-line. The temperature-index factor is dependent on the chosen temporal resolutions and needs, in principle, recalibration for other resolutions. There is a potential for improving the model by applying a more realistic spatial distribution of snow.

CONCLUSIONS

The overarching aim of this study is to provide suggestions for snow model structures suited for predicting snow conditions at ungauged sites and under a changed climate. Common to the results of several other such model comparisons we find that well-established empirical parameterizations give as good (and poor) results as more physically based relationships (Essery *et al.* 2013). Some quite clear conclusions can, however, be drawn from this exercise.

Empirical relationships, such as the temperature index method, work well, also regionally if the temperature index is calibrated against the variable it is supposed to represent, namely the melting of snow as a function of temperature. The SeNorge model has shown itself to be the best model when validated against SWE and SCA. The temperature-index based method calibrated against runoff

(DDD_CX), had the poorest results when validated against SWE and SCA.

Proxy modelling of the energy balance to simulate snowmelt (DDD_EB) works surprisingly well, and almost as good as SeNorge when validated against SWE and SCA. It is especially promising with respect to predictions in ungauged basins that no extra tuning was performed for the different catchments. The only difference in parameters from one catchment to the next (besides those calibrated with DDD_CX) was that the location of the centre of the catchment was provided so that the SW calculations could be performed.

Both SeNorge and DDD_EB show promise as candidates for model structures suited for simulating snow for ungauged basins and for a changed climate. The model parameters in the snow module of DDD_EB need no calibration, and the consistently good results of SeNorge over all the catchment indicate that the calibration of the extended temperature index melt model is physically meaningful and thereby possibly independent of climate. Running Crocus regionally, probably need some scheme for realistically representing the spatial distribution of SWE. The temperature-index method for snowmelt calibrated against runoff, cannot be recommended.

ACKNOWLEDGEMENT

The support of the Norwegian Research Council, Glommen and Laagen Brukseierforening, E-CO, Trøndelag Energiverk and HYDRO Energi through the project 'Better SNOW models for predictions of natural Hazards and HydropOWer applications' (SNOWHOW) (244153/E10), is gratefully acknowledged.

REFERENCES

Bergström, S. 1992 *The HBV Model – its Structure and Applications*. SMHI Reports Hydrology No. 4. Swedish Meteorological and Hydrological Institute, Norrköping, Sweden.

Brun, E., David, P., Sudul, M. & Brunot, G. 1992 A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting. *Journal of Glaciology* **38**, 13–22.

Collados-Lara, A.-J., Pardo-Igúzquiza, E. & Pulido-Velazquez, D. 2017 Spatiotemporal estimation of snow depth using point data from snow stakes, digital terrain models, and satellite data. *Hydrological Processes* **31**, 1966–1982. https://doi.org/10.1002/hy1165.

eKlima 2017 Free access to weather and climate data from Norwegian Meteorological Institute from historical data to real time observations. http://eklima.met.no/ (Accessed online 15 September 2017).

Essery, R. & Pomeroy, J. 2004 Implications of spatial distributions of snow mass and melt rate for snow-cover depletion: theoretical considerations. *Ann. Glaciol.* **38**, 261–265.

Essery, R., Morin, S., Lejeune, Y. & Menard, C. 2013 A comparison of 1701 snow models using obervations from an alpine site. *Adv. Water Resources* **55**, 131–148. doi:10.1016/j.advwatres.2012.07.013.

Etchevers, P., Martin, E., Brown, R., Fierz, C., Lejeune, Y., Bazile, E., Boone, A., Dai, Y.-J., Essery, R., Fernandez, A., Gusev, Y., Jordan, R., Koren, V., Kowalczyk, E., Nasonova, N. O., Pyles, R. D., Schlosser, A., Shmakin, A. B., Smirnove, T. G., Strasser, U., Verseghy, D., Yamzaki, T. & Yang, Z.-L. 2004 Validation of the Energy budget of an alpine snowpack simulated by several snow models (SnowMIP project). *Ann. Glaciol.* **38**, 150–158.

Helbig, N., van Herwijnen, A., Magnusson, J. & Jonas, T. 2015 Fractional snow-covered area parameterization over complex topography. *Hydrol. Earth Syst. Sci.* **19**, 1339–1351. doi:105194/hess-19-1339-2015.

Herrero, J. & Polo, M. J. 2012 Parameterization of atmospheric longwave emissivity in a mountainous site for all sky conditions. *Hydrol. Earth Syst. Sci.* **16**, 3139–3147. doi:10.5194/hess-16-3139-2012.

Hock, R. 2003 Temperature index melt modelling in mountain areas. *J. Hydrol.* **282**, 104–115.

Kirchner, J. W. 2006 Getting the right answers for the right reasons: linking measurements, analyses and models to advance the science of hydrology. *Water Resour. Res.* **42**, W03S04. doi:10.1029/2005WR004362.

Kling, H., Fuchs, M. & Paulin, M. 2012 Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.* **424**, 264–277. doi:10.1016/j.jhydrol.2012.01.011.

Luijting, H., Vikhamar-Schuler, D., Aspelien, T. & Homleid, M. 2017 Forcing the SURFEX/Crocus snow model with combined hourly meteorological forecasts and gridded observations in southern Norway. *The Cryosphere Discuss.* https://doi.org/10.5194/tc-2017-220.

Lussana, C., Saloranta, T., Skaugen, T., Magnussson, J., Tveito, O. E. & Andersen, J. 2018 Senorge2 daily precipitation, an observational gridded datatset over Norway from 1957 to the present day. *Earth System Science Data* **10**, 235–249. https://doi.org/10.5194/essd-10-235-2018.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L.,

Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V. & Voldoire, A. 2013 The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geoscientific Model Development* **6**, 929–960. https://doi.org/10.5194/gmd-6-929-2013, https://www.geosci-model-dev.net/6/929/2013/.

Melvold, K. & Skaugen, T. 2013 Multiscale spatial variability of lidar-derived and modeled snow depth on Hardangervidda, Norway. *Ann. Glaciol.* **54**. doi:10.3189/2013AoG62A161.

Mir, R. A., Jain, S. K., Saraf, A. K. & Goswami, A. 2015 Accuracy assessment and trend analysis of MODIS-derived data on snow-covered areas in the Sutlej basin, Western Himalayas. *International Journal of Remote Sensing* **36**, 3837–3858.

Omhura, A. 2001 Physical basis for the temperature-based melt-index method. *Journal of Applied Meteorology* **40**, 753–761.

Pardo-Iguzquiza, E., Collados-Lara, A. J. & Pulido-Velazquez, D. 2017 Estimation of the spatiotemporal dynamics of snow cover area by using cellular automata models. *Journal of Hydrology* **550**, 230–238.

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., Barr, A., Bartlett, P., Boone, A., Deng, H., Douville, H., Dutra, E., Elder, K., Ellis, C., Feng, X., Gelfan, A., Goodbody, A., Gusev, Y., Gustafsson, D., Hellström, R., Hirabayashi, Y., Hirota, T., Jonas, T., Koren, V., Kuragina, A., Lettenmaier, D., Li, W.-P., Luce, C., Martin, E., Nasonova, O., Pumpanen, J., Pyles, R. D., Samuelsson, P., Sandells, M., Schädler, G., Shmakin, A., Smirnova, T. G., Stähli, M., Stöckli, R., Strasser, U., Su, H., Suzuki, K., Takata, K., Tanaka, K., Thompson, E., Vesala, T., Viterbo, P., Wiltshire, A., Xia, K., Xue, Y. & Yamazaki, T. 2009 Evaluation of forest snow processes models (SnowMIP2). *Journal of Geophysical Research* **114** (D06111). doi:10.1029/2008JD011063.

Salomonson, V. V. & Appel, I. 2004 Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote Sensing of Environment* **89**, 351–360. https://doi.org/10.1016/j.rse.2003.10.016.

Saloranta, T. M. 2012 Simulating snow maps for Norway: description and statistical evaluation of the seNorge snow model. *The Cryosphere* **6**, 1323–1337. https://doi.org/10.5194/tc-6-1323-2012, https://www.the-cryosphere.net/6/1323/2012/.

Saloranta, T. M. 2014 *New Version (v.1.1.1) of the SeNorge Snow Model and Snow Maps for Norway*. Rapport 6-2014, Norwegian Water Resources and Energy Directorate, Oslo, Norway, pp. 30. Available at: http://webby.nve.no/publikasjoner/rapport/2014/rapport2014_06.pdf.

Sexstone, G. A. & Fassnacht, S. R. 2014 What drives basin scale spatial variability of snowpack properties in northern Colorado? *Cryosphere* **82**, 329–344.

Skaugen, T. & Mengistu, Z. 2016 Estimating catchment scale groundwater dynamics from recession analysis-enhanced constraining of hydrological models. *Hydrol. Earth. Syst. Sci.* **20**, 4963–4981. doi:10.5194/hess-20-4963-2016.

Skaugen, T. & Onof, C. 2014 A rainfall runoff model parameterized form GIS and runoff data. *Hydrol. Process.* **28**, 4529–4542. doi:10.1002/hyp.9968.

Skaugen, T. & Saloranta, T. 2015 *Simplified Energy-Balance Snowmelt Modeling*. NVE-rapport 31-2015. Norwegian Water Resources and Energy Directorate, Oslo, Norway.

Skaugen, T. & Weltzien, I. H. 2016 A model for the spatial distribution of snow water equivalent parameterised from the spatial variability of precipitation. *The Cryosphere* **10**, 1947–1963. doi:10.5194/tc-10_1947_2016.

Slater, A. G., Schlosser, C. A., Desborough, C. E., Pitman, A. J., Henderson-Sellers, A., Robock, A., Vinnikov, K. Y., Mitchell, K., Boone, A., Braden, H., Chen, F., Cox, P. M., deRosnay, P., Dickinson, R. E., Dal, J.- Y., Duan, Q., Entin, J., Etchevers, P., Gedney, N., Gusev, Y. M., Habets, F., Kim, J., Koren, V., Kowalczyk, E. A., Nasonova, O. N., Noilhan, J., Schaake, S., Shmakin, A. B., Smirnova, T. G., Verseghy, D., Wetzel, P., Xue, Y., Yang, Z.-L. & Zeng, Q. 2001 The representation of snow in land surface schemes: results from PILPS 2(d). *J. Hydrometeorology* **2**, 7–25. doi:10.1175/1525-7541(2000)002<0007:TROSIL>2.0.CO;2.

Solberg, R., Koren, H. & Amlien, J. 2006 *A Review of Optical Snow Cover Algorithms*. SAMBA/40/06. Norwegian Computing Centre, Norway.

Strasser, U., Etchevers, P. & Lejeune, Y. 2002 Inter-comparison of two snow models with different complexity using data from an alpine site. *Nordic Hydrology* **3381**, 15–26.

Tarboton, D. G. & Luce, C. H. 1996 *Utah Energy Balance Snow Accumulation and Melt Model (UEB)*. Computer model and technical description and user guide. UTAH Water Research Laboratory, Utah State University, Utah, USA and USDA Forest Research Service, Intermountain Research Station.

Tvedalen, A. K. 2015 *Snow Melt: Evaluation of an Energy Balance Model*. MSci thesis, University of Oslo, Norway. https://www.duo.uio.no/handle/10852/45502.

Unsworth, M. H. & Monteith, J. L. 1975 Long-wave radiation at the ground. I. Angular distribution of incoming radiation. *Quart. J.R. Met. Soc.* **101**, 13–24.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Moigne, P. L., Martin, E. & Willemet, J.-M. 2011 The detailed snowpack scheme Crocus and its implementation in SURFEX v7. *Geoscientific Model Development Discussion* **4**, 2365–2415.

Walter, M. T., Brooks, E. S., McCool, D. K., King, L. G., Molnau, M. & Boll, J. 2005 Process-based snowmelt modeling: does it require more input data than temperature index modeling? *J. Hydrol.* **300**, 65–75.

First received 13 December 2017; accepted in revised form 3 May 2018. Available online 6 June 2018