



Loss-compensated snowfall scaling using large-area lidar observations improves snow depth simulations over multiple years

Jens Oprel^{1,2}, Tobias Jonas¹, Andreas S. Stordal³, Felix N. Matt⁴, Gaute Lappegard⁴,
Manuela I. Brunner^{1,2,5}, and Jan Magnusson¹

¹WSL Institute for Snow and Avalanche Research SLF, Davos Dorf, Switzerland

²Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

³Energy & Technology, NORCE Research, Bergen, Norway

⁴Statkraft AS, Oslo, Norway

⁵Climate Change, Extremes and Natural Hazards in Alpine Regions Research Center CERC, Davos Dorf, Switzerland

Correspondence: Jens Oprel (jens.oprel@slf.ch)

Abstract. In many regions, reliable runoff forecasts over spring and summer depend on accurate predictions of snowmelt volume and timing, which in turn require snow models that realistically represent peak snow amounts as well as their spatial distribution across the basin. Observation-based snowfall scaling methods can reduce biases in the snowfall input data and improve the representation of spatial variability in the modelled snowpack. These methods adjust the forcing data of snow models so that the simulations match snow depth observations, e.g. from high-resolution lidar scans. Our results show that existing snowfall scaling methods cannot reliably incorporate lidar observations over two large study regions in South Norway. In particular, existing methods fail to produce correct results in areas where snow has melted before the lidar observation. To overcome this limitation, we introduce the loss-compensated scaling method, which accounts for melt and sublimation losses that occur prior to the lidar observation. Simulations with this method reliably reproduce the assimilated lidar snow depth observations. Furthermore, we demonstrate that, with this method, information from a single lidar scan can substantially improve model performance even in years without such observations, reducing the simulated bias in snow depth by at least 67 % and the RMSE by at least 50 %. Thus, our improved method provides a reliable way to integrate lidar scans into physics-based snow models, enhancing model performance over multiple years, even in large study areas where low-lying regions start to melt before the basin reaches peak snow accumulation.

1 Introduction

Reliable predictions of runoff volume and timing are crucial for flood forecasting and the optimal operation of hydropower plants (Anghileri et al., 2016; Bruland, 2012; Davidson-Chaput et al., 2025; Nester et al., 2012). During the melt season, these variables are strongly influenced by snow and its spatial distribution (Clark et al., 2011; Griessinger et al., 2019; Lundquist and Dettinger, 2005). Therefore, representative estimates of snow amounts at the peak of winter provide valuable information on the total melt volume in spring. In addition, information on the spatial distribution of snow can help predict snowmelt timing: a more spatially variable snowpack leads to a longer melt season with lower peak flows than a more uniform snowpack (Brauchli



et al., 2017; Pfohl et al., 2026; Sextstone et al., 2020). Spatial variations in snow depth and snow water equivalent typically form during the accumulation rather than during the melt phase (Egli et al., 2012). Weather patterns influence snow amounts and the large-scale spatial variability of accumulation processes, whereas preferential deposition, wind redistribution, and avalanching are the main drivers of small-scale variability (Freudiger et al., 2017). Capturing both large- and small-scale variability of snow depth improves short- and long-term runoff predictions, especially in mountain regions (Warscher et al., 2013).

Snow estimates are often derived using distributed snow models that simulate snow depth in a grid, using meteorological data such as temperature, radiation, and precipitation as inputs. Such snow models often fail to capture both total snow volumes and the spatial variability of snow cover at the end of the accumulation season, due to large errors in forcing data, deficiencies in model structure, and parameter uncertainties (Günther et al., 2019). To accurately predict snow volume, such models require reliable snowfall input data from weather models or observations (Raleigh et al., 2015). Such observations are often highly uncertain because measuring snowfall accurately is difficult (Wolff et al., 2015), which also affects the output of weather models that assimilate such uncertain observations .

To predict small-scale spatial variability in snow cover, some models explicitly represent preferential deposition and redistribution processes (Berg et al., 2024; Gerber et al., 2019; Quéno et al., 2024; Rudisill et al., 2024). However, these approaches are often ineffective because of substantial simplifications of the relevant processes and, again, the difficulty of obtaining accurate forcing data, such as detailed wind fields (Gerber et al., 2019; Rudisill et al., 2024). Furthermore, the computational cost of such models is often prohibitive of operational forecasts over large areas, as they must run at high resolution to accurately represent the governing processes. Thus, current modelling approaches alone cannot consistently reproduce real-world snow cover amounts and spatial variability, which can be problematic for runoff prediction.

To reduce errors in snow modelling, simulations are commonly combined with observations using data assimilation techniques such as the particle filter or the ensemble Kalman filter (Andreadis and Lettenmaier, 2006; Leisenring and Moradkhani, 2011). These methods update model states, parameters, or forcing data whilst accounting for model and observation uncertainties. They have been applied in snow modelling to correct simulations by assimilating point observations of snow depth (Magnusson et al., 2017; Oberrauch et al., 2024) or snow cover fraction retrievals from satellites (Aalstad et al., 2018; Andreadis and Lettenmaier, 2006; Baba et al., 2018). Airborne lidar surveys of snow depth provide another promising observational data source for correcting snow simulations because these measurements can cover complete basins at high spatial resolution with low observation errors (Painter et al., 2016). As airborne lidar observations are considerably more accurate than prior model results, less complex data assimilation techniques than filters, such as direct insertion, could be justified (Hedrick et al., 2018; Revuelto et al., 2016). In direct insertion, model states are replaced by observations—in this case, of snow depth. This typically improves model results after the time of insertion, but cannot improve model results for other seasons. This can be problematic because airborne lidar surveys are expensive and, in some regions, cannot be conducted every year due to persistent adverse weather. Since small-scale patterns in snow distribution tend to repeat between years (Deems et al., 2008; Pflug et al., 2021), we hypothesize that it should be possible to leverage information from occasional lidar scans to improve snow-hydrological forecasting, even in years without such observations.



One promising yet simple data assimilation approach that adjusts model results throughout one or more entire snow seasons is to adjust the snowfall input to the model using lidar scans of snow depth. With this approach, the snowfall is scaled by a factor such that the simulated snow depth matches the observed snow depth at the time of the scan. Such snowfall scaling methods aim to correct errors in the input data and to implicitly account for preferential deposition and snow redistribution processes. Vögeli et al. (2016) proposed two methods to derive snowfall scaling factors from a lidar scan: the direct scaling method and the iterative scaling method. The direct scaling method derives scaling factors from observations alone and does not require an initial model run. This method has been applied with various physics-based snow models to better represent small-scale snow depth patterns (Brauchli et al., 2017; Crumley et al., 2024; Haberkorn et al., 2017; Hale et al., 2023; Kiewiet et al., 2022; Pflug et al., 2021). The direct scaling method has also been applied to correct snowfall in years without lidar observations by transferring scaling factors across years (Hale et al., 2023; Kiewiet et al., 2022; Pflug et al., 2021; Vögeli et al., 2016). However, the direct scaling method does not correct for biases in the snowfall forcing (Brauchli et al., 2017; Pflug et al., 2021). In contrast, the iterative scaling method corrects for such biases, requiring an initial model run to scale the snowfall input based on the ratio between the observed and modelled snow depth. This method is therefore better suited than the direct scaling method for correcting snow simulations in basins where the precipitation input data are highly uncertain.

The existing snowfall scaling methods described above have largely been successful but have mainly been applied and tested in small ($< 150 \text{ km}^2$) Alpine study regions (Koch et al., 2024; Pulka et al., 2024; Richter et al., 2021). In such small catchments, it is often possible to acquire lidar surveys at a time when the snow cover is largely unaffected by snowmelt. However, in large regions with considerable elevation differences, it is rarely possible to acquire basin-wide lidar observations of snow depth unaffected by snow losses near peak accumulation, as lower elevations may already have experienced snowmelt whilst high-elevation areas have not yet reached peak accumulation. Since existing snowfall scaling methods can yield unreliable scaling factors in regions with substantial snowmelt (Pflug et al., 2021), it remains unclear to what extent the methods of Vögeli et al. (2016) can reliably be applied in large mountainous watersheds.

The main aim of this study is to improve snowfall scaling methods for operational use in large watersheds. We address this aim in three steps. First, we evaluate how effectively existing snowfall scaling methods integrate high-resolution lidar snow depth observations into a physics-based snow model. In doing so, we identify limitations in large study regions with considerable elevation differences and where melt occurs before the lidar observation. Second, to address these limitations, we develop a new loss-compensated scaling method that adjusts snowfall input in a robust and computationally efficient manner across a range of model resolutions, even in regions where substantial melt has occurred prior to the lidar acquisition. Third, we study the extent to which scaling factors from a single lidar scan can be transferred in time to improve results in other years and which scaling approach is best suited to achieve this transferability. The study is conducted in two large catchments in South Norway where full-area airborne lidar observations of snow depth are available for three winters (2021/2022, 2023/2024 and 2024/2025).



2 Methods and data

2.1 Study regions

90 In this study, we focus on the two study regions, Ulla-Førre and Songa, in South Norway (see Fig. 1). The Ulla-Førre region covers 757 km² with elevations ranging from 550 m to 1550 m, whilst Songa covers 586 km² with elevations ranging from 965 m to 1602 m. The regions host several hydropower plants, which together account for around 4 % of the Norwegian electricity production (Statkraft AS, 2025; IEA, 2025). Both regions have a predominantly rocky, open, and mountainous terrain shaped by glaciers during the last ice age. Ulla-Førre has a large inland water fraction (30 %), partly owing to the vast
 95 hydropower reservoir Blåsjø, whilst the inland water fraction is smaller in Songa (12 %). In Ulla-Førre, most of the area lacks vegetation and only the low-lying western parts are covered by shrubs. As in Ulla-Førre, vegetation in Songa is confined to the lower elevations in the east, whilst the higher western parts are rocky. In both regions, forest cover is limited (3 %). A distinctive feature of Ulla-Førre is its high mountain plateaus dissected by elongated trenches (see Fig. 1b).

South Norway exhibits a clear west-to-east precipitation gradient, with wetter climate towards the ocean in the west and
 100 drier conditions inland. The average annual precipitation in Ulla-Førre is approximately 2600 mm, whilst Songa is considerably drier, with around 1600 mm. Roughly half of this precipitation falls as snow, producing substantial snow accumulation throughout winter. In both regions, the accumulation season begins in October and snow persists into late summer, until July or August. Reliable estimates of both snow volume and its spatial distribution are therefore crucial for planning hydropower production in these regions.

105 2.2 Lidar data collection

In both study regions, airborne lidar scanning was used to obtain snow depth maps at a spatial resolution of 3 m, covering the full extent of each region. Elevation data were collected during snow-free and snow-covered periods using a Riegl VQ-1560 II-S laser scanning system (RIEGL Laser Measurement Systems GmbH, 2020) mounted on a fixed-wing aircraft. The resulting point clouds were processed using the methods outlined by Painter et al. (2016). Snow depth was derived on a 3 by 3 m grid
 110 from the difference in elevation between the snow-covered and snow-free surveys. Next, data over inland water were removed based on the outlines of natural and regulated lakes from the N50 Map Data (Norwegian Mapping Authority, 2026b). Finally, the data were upsampled to aligned model grids of 90, 300, and 900 m resolution by taking the mean of all available pixels within each model grid cell. We also store the number of pixels that contributed to this average such that area-weighting can be applied during evaluation (see Section 2.5). We chose not to remove pixels with negative snow depths from the 3 m resolution dataset
 115 since doing so would introduce a positive bias to the aggregated snow depths. Negative values in the aggregated snow depths are not used during assimilation and evaluation. Snow depth after the accumulation season was derived from scans in March or April for three years in Ulla-Førre and two years in Songa. An additional measurement is available in Ulla-Førre in June for one year. Specifically, in Ulla-Førre the snow free scan was obtained on 2021-08-28 and snow-covered scans on 2022-03-19, 2022-06-05, 2024-04-22, and 2025-03-14. For Songa, the snow free scan was obtained on 2024-09-01 and snow-covered scans

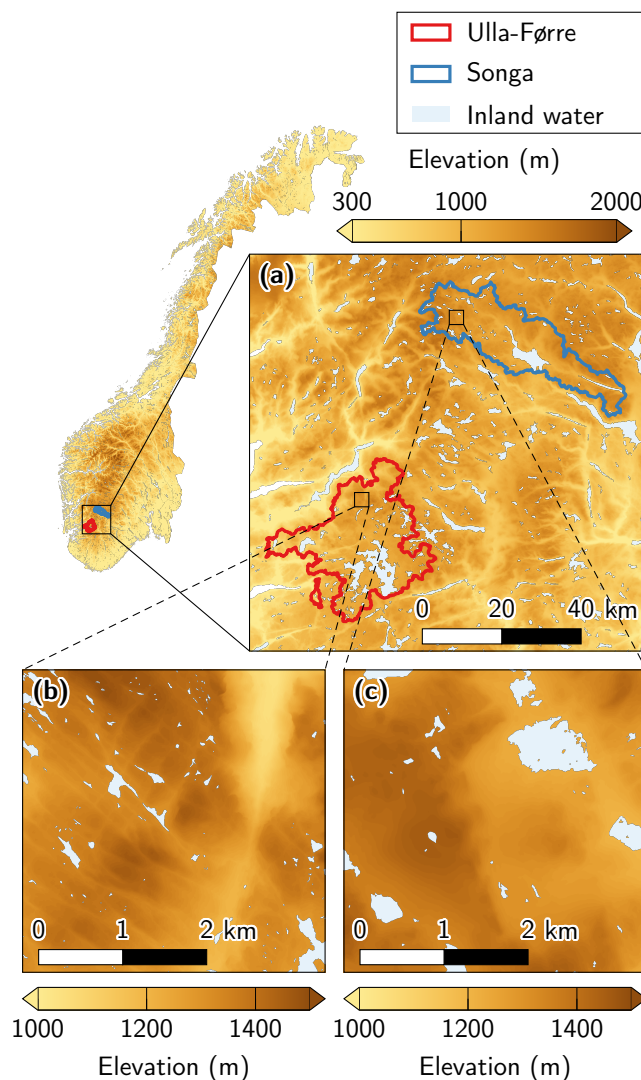


Figure 1. Overview of the study regions and their location in South Norway. Panel a shows the boundaries of the Ulla-Førre and Songa regions on a map of elevation and inland water. Panels b and c show subregions of Ulla-Førre and Songa, respectively, in greater detail. These insets use separate colour scales, shown beneath each panel, to highlight the small-scale topographic features. The elevation data on the country scale were provided by the Norwegian Mapping Authority (2026a). Panels b and c show elevation data from the snow-free lidar scans.

on 2024-04-22 and 2025-03-19. To our knowledge, these are the most spatially extensive airborne lidar snow depth scans available in Europe to date.

The maps in Fig. 2 show the scans acquired on 22 April 2024. The spatial accumulation patterns they show are representative of those found in other scans. In Ulla-Førre, the largest snow depths are recorded in the high-elevation northern part of the



region, whereas less snow is present in the lower-elevation areas (Fig. 2a). On the high mountain plateau in the north, snow predominantly accumulates in the elongated trenches, whilst the ridges show a shallower snow cover (Fig. 2b). In Songa, the deepest snowpack is found in the wetter, higher-elevation western part of the catchment, whilst less snow accumulates in the drier, lower-elevation eastern part (Fig. 2d). At small scales, the spatial variability in Songa differs markedly from that in Ulla-Førre, with lower overall variation, likely owing to the smoother terrain in this region (Fig. 2e). These snow depth maps show that accumulation patterns differ substantially between the two regions, largely reflecting differences in terrain.

When considering the remaining lidar scans obtained, the small-scale snow accumulation patterns remain qualitatively similar (not shown). For instance, in Ulla-Førre, large snow depths are recorded in the elongated trenches every year. Quantitatively, however, the probability density functions of measured snow depth differ between years and regions (Fig. 2c and 2f). For example, in Ulla-Førre, snow depths above approximately 3 m have similar probability densities in 2022 and 2024. In 2024, however, shallow snow depths (< 0.80 m) appear to be more common than in 2022, potentially because the scan was conducted later in the season and was thus more affected by snowmelt. Nevertheless, both scans have almost the same mean snow depth. In 2025, snow was scarcer than in 2024 in both Ulla-Førre and Songa. Snow depths do not just vary by year, but also by region. For example, snow depths above 5 m are less common in Songa, occurring in 5.5 % of the area in 2024, compared with 9.7 % in Ulla-Førre in the same year.

2.3 Model setup and forcing data

In this study, we use the energy-balance snow model FlexibleSnowModelOSHD.jl (abbreviated as FSMOSHD; Magnusson et al., 2026), which is based on the intermediate-complexity FSM model developed by Essery (2015). FSM and variants of this model have been used in numerous studies, for example, to investigate parameter sensitivity (Günther et al., 2019), the impact of model resolution on simulation results (Magnusson et al., 2019), the effect of data assimilation methods (Deng et al., 2024; Oberrauch et al., 2024), and the importance of snow redistribution processes (Marsh et al., 2024; Quéno et al., 2024). The model resolves mass and energy balances across a set of snow layers (three in our case). The surface energy-balance equation is solved analytically, whilst heat conduction through the snow and soil is computed using the Crank–Nicolson method (Crank and Nicolson, 1996). Here, the model is set up without snow redistribution, allowing grid cells to be run independently from each other. Earlier studies indicate that FSM-like models can provide snow-hydrological simulations of similar quality to those from more complex models, but at a lower computational cost (Magnusson et al., 2015). For this reason, the model is used for operational snow-hydrological forecasting in Switzerland (Mott et al., 2023) and is an appropriate choice for our study.

As meteorological input, we use the MET Nordic dataset (Neuville et al., 2025; Norwegian Meteorological Institute, 2023, 2026), which provides hourly fields of air temperature, relative humidity, longwave radiation, shortwave radiation, wind speed, precipitation, and air pressure at 1 km horizontal resolution. It is based on the 2.5 km numerical weather prediction model MEPS (MetCoOp Ensemble Prediction System; Eresmaa et al., 2026) combined with statistical interpolation of weather-station data (Lussana et al., 2019). We downscale the MET Nordic meteorological fields to grids with 90, 300, and 900 m horizontal resolution following the procedures summarised by Magnusson et al. (2025) and Mott et al. (2023). All variables are bilinearly interpolated from the coarse-scale native grid of the forcing data to the finer-scale grid of the model.

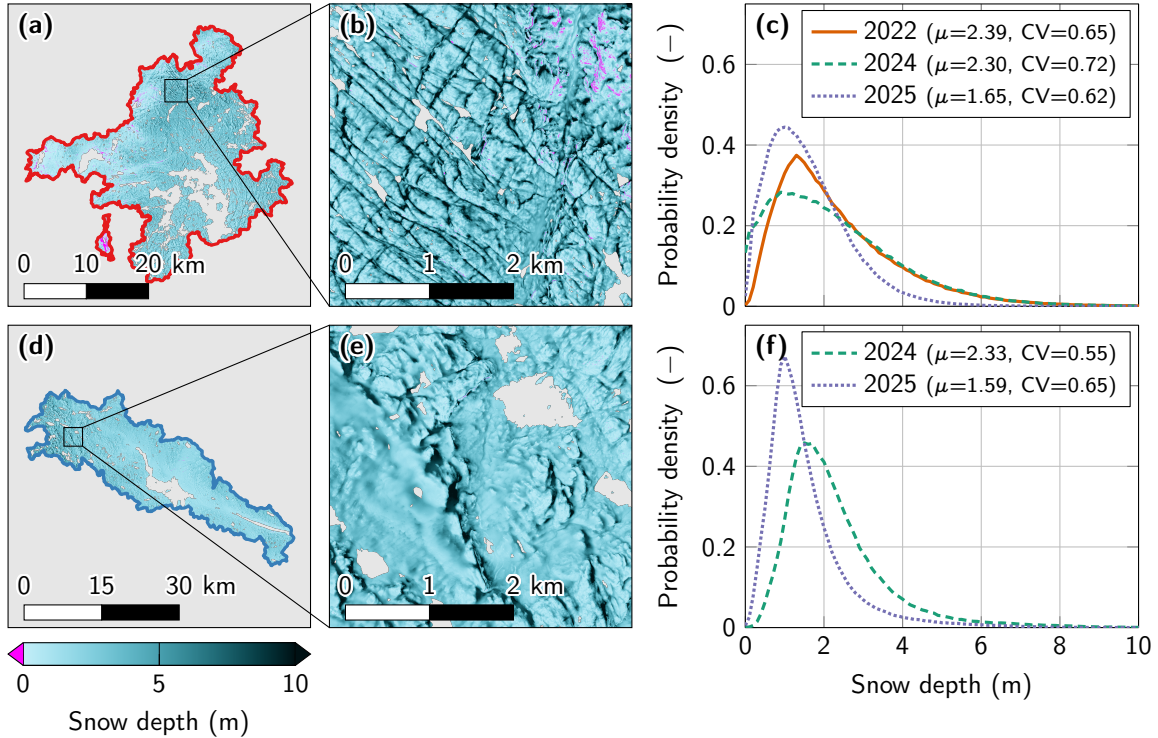


Figure 2. Snow depth measured with lidar scanning in Ulla-Førre (Panels a–c) and Songa (Panels d–f). The maps in Panels a and d show the full-area lidar scans of the catchments acquired on 22 April 2024, whilst the insets in Panels b and e show small-scale snow depth patterns for two subregions. The red and blue lines indicate the catchment outlines of Ulla-Førre and Songa, respectively, whilst grey areas depict areas without lidar data (e.g. inland water and areas outside the study regions) and magenta indicates snow depth measurements of 0 (and below). Panels c and f show the probability density functions of the full-area snow depth for the scans obtained during the late accumulation season in different years (2022 not available for Songa). The mean (μ) and coefficient of variation (CV) of the distributions are indicated in the legends.

Air temperature and relative humidity are additionally corrected for elevation. For air temperature, a constant lapse-rate correction of 6.0 K km^{-1} is applied. Relative humidity is downscaled via dew-point temperature, computed from the coarse fields, interpolated with the same lapse rate, and converted back to relative humidity using the downscaled air temperature. To obtain rain and snowfall from the total precipitation, we apply a temperature-based phase-separation method following Mazzotti et al. (2020), with a smooth transition from snowfall to rainfall around a threshold temperature;

$$S(i, t) := P(i, t) \frac{1}{1 + \exp\left(\frac{T(i, t) - T_{\text{thr}}}{m_P}\right)}, \quad R(i, t) := P(i, t) - S(i, t), \quad \forall i, t, \quad (1)$$

where P , S , and R denote total precipitation, snowfall, and rainfall as functions of the grid cell index i and time t . The parameters $T_{\text{thr}} = 1.04^\circ\text{C}$ and $m_P = 0.15^\circ\text{C}$ that control the centre and width of the logistic function were calibrated by Mazzotti et al. (2020). Although simple, this method yields a more physically realistic partitioning into precipitation phases,



including mixed precipitation, than a step-function approach, where all precipitation falls either as snow or rain based on temperature.

2.4 Snowfall scaling methods

170 2.4.1 Direct scaling method of Vögeli et al. (2016)

Vögeli et al. (2016) proposed the direct scaling method, which scales total precipitation in each grid cell based solely on measured snow depths. To primarily target snowfall rather than rainfall, the precipitation scaling is not applied to grid cells where the air temperature exceeds the threshold temperature (T_{thr}). Precipitation also remains unaltered in regions without lidar data and where the lidar data is below 0.05 m. In grid cells with valid lidar data and cold temperatures, precipitation is scaled
 175 by the ratio of the measured snow depth in that grid cell to the spatial mean of the measured snow depth over the area with valid lidar data:

$$c_{\text{DS}}(i) := \frac{h_{\text{lidar}}(i)}{h_{\text{lidar, mean}}}, \quad \forall i \text{ such that } h_{\text{lidar}}(i) \geq 0.05 \text{ m}, \quad (2)$$

$$P_{\text{DS}}(i, t) := \begin{cases} c_{\text{DS}}(i)P_{\text{mean}}(t), & \text{if } T(i, t) \leq T_{\text{thr}} \wedge h_{\text{lidar}}(i) \geq 0.05 \text{ m}, \\ P(i, t), & \text{otherwise.} \end{cases} \quad (3)$$

Here, P and T denote the unscaled precipitation and air temperature forcing, h_{lidar} denotes the measured snow depth, and
 180 $h_{\text{lidar, mean}}$ and $P_{\text{mean}}(t)$ denote the spatial means of measured snow depth and precipitation at time t , both over the area with valid lidar data. The resulting scaled precipitation is denoted P_{DS} , where the subscript ‘DS’ stands for direct scaling method. Because the direct scaling method adjusts total precipitation rather than snowfall, the scaled precipitation P_{DS} is partitioned into snowfall and rainfall only after scaling, using the scheme described in Section 2.3. This reversed ordering introduces an inconsistency: the partitioning in Eq. (1) treats T_{thr} as the centre of a smooth snow-to-rain transition, whereas Eq. (3) applies it
 185 as a step function. Therefore, at temperatures close to T_{thr} , part of the snowfall may go unscaled whilst part of the rainfall may be scaled.

Although the direct scaling method has the advantage that lidar-derived spatial snow depth patterns can be translated directly into spatial snowfall patterns without requiring a model run, it has two important disadvantages. First, any basin-average biases in the precipitation data remain uncorrected, because the scaling factor is determined without knowledge of that bias. Second,
 190 the spatial patterns present in the original precipitation forcing are discarded. For small areas, it may be reasonable to assume that the meteorological forcing does not contain meaningful information on the spatial variability of snowfall. In larger regions, however, the spatial variability in the forcing data represents potentially valuable information.

2.4.2 Iterative scaling method of Vögeli et al. (2016)

To improve upon the direct scaling method, Vögeli et al. (2016) also propose an iterative scaling method that, unlike the
 195 direct scaling method, adjusts the snowfall fields obtained after phase-separating the total precipitation (Eq. (1)). The method



determines the snowfall scaling factors as the ratio of observed to simulated snow depth and is only applied in regions with valid lidar data that are substantially above 0 (denoted $\mathcal{V}$):

$$c_{\text{IS},1}(i) := \frac{h_{\text{lidar}}(i)}{h_{\text{model},0}(i)}, \quad \forall i \in \mathcal{V} := \{i \mid h_{\text{lidar}}(i) \geq 0.05 \text{ m}\}. \quad (4)$$

Here, $h_{\text{model},0}$ denotes the initial simulated snow depth at the time of the lidar observation, $c_{\text{IS},1}$ the resulting scaling factor after this first iteration, and the subscript ‘IS’ denotes the iterative scaling method. The computed scaling factor is then applied to scale snowfall as

$$S_{\text{IS},1}(i, t) := \begin{cases} c_{\text{IS},1}(i) \cdot S(i, t), & \text{if } i \in \mathcal{V}, \\ S(i, t), & \text{otherwise,} \end{cases} \quad \forall i, t, \quad (5)$$

where S denotes the unscaled snowfall forcing defined in Section 2.3, and $S_{\text{IS},1}$ the corrected snowfall after the first iteration of this method. Running the model with this scaled snowfall should result in an improved match between observed and simulated snow depth. This scaling method assumes that uniformly scaling all snowfall inputs in a grid cell by a factor c results in a proportional change in the simulated snow depth at the time of the lidar observation, as illustrated by the solid blue line in Fig. 3:

$$h_{\text{model}} = c \cdot h_{\text{model},0}, \quad (6)$$

The scaling factor $c_{\text{IS},1}$ in Eq. (4) can be obtained from this approximation by substituting the desired result $h_{\text{model}} = h_{\text{lidar}}$ into Eq. (6). However, this approximation neglects that changing snowfall throughout the simulation also alters snowpack processes affecting snowpack evolution, such as densification, melt, and sublimation. Consequently, the final snow depth at the time of the lidar observation does not generally reproduce the observed snow depth after the first application of the scaling method. To reduce such errors, additional iterations (indexed by j) can be performed by generalising Eq. (4) and (5) to:

$$c_{\text{IS},j}(i) := \frac{h_{\text{lidar}}(i)}{h_{\text{model,IS},j-1}(i)} c_{\text{IS},j-1}(i), \quad \forall j \in \{2, 3, \dots\}, \quad \forall i \in \mathcal{V}, \quad (7)$$

$$S_{\text{IS},j}(i, t) := \begin{cases} c_{\text{IS},j}(i) \cdot S(i, t), & \text{if } i \in \mathcal{V}, \\ S(i, t), & \text{otherwise,} \end{cases} \quad \forall j \in \{1, 2, \dots\}, \quad \forall i, t. \quad (8)$$

Here, the scaling factors at iteration j , $c_{\text{IS},j}$, are determined from the lidar scan and the modelled snow depths $h_{\text{model,IS},j-1}$ at the time of the lidar observation resulting from the model run with the snowfall input $S_{\text{IS},j-1}$ obtained in the previous iteration.

To reduce computation time for this iterative scaling method, we compute the scaling factors as follows: For each grid cell, we first load the forcing data for one entire simulation period into memory. Next, we run the model until the lidar observation to obtain $h_{\text{model,IS},1}$ and calculate the scaling factor following Eq. (4). We continue iterating by running the snow model with this scaling and calculating new scaling factors as in Eq. (7). We stop iterating once the absolute difference in snow depth between the observation and the latest simulation falls below 0.01 m, or upon reaching a maximum of 100 iterations. This threshold of 0.01 m is smaller than the uncertainty in the lidar data (e.g. Painter et al. (2016) and Currier et al. (2019) report an accuracy of

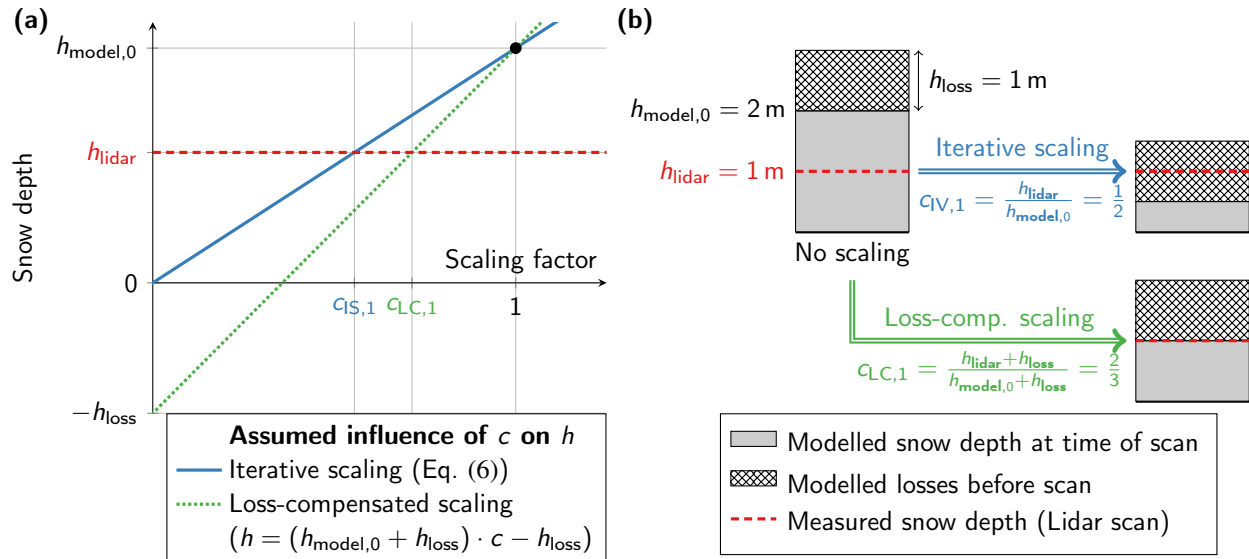


Figure 3. Panel a shows how the assumed relation between snow depth and snowfall scaling factor differs between the iterative scaling method of Vögeli et al. (2016) (blue solid line) and the loss-compensated method (green dotted line). The snowfall scaling in these methods is determined from the intersection of these relations with the line $h = h_{\text{lidar}}$. Panel b illustrates how substantial melt influences the model outcome after applying one iteration of the two scaling methods.

< 10 cm for similar lidar observations). The goal of scaling methods, however, is to match the model to the lidar observations including their errors. This is beneficial when the accuracy of the model without scaling is much lower than the accuracy of the measurement, which is the case in our study areas. We run the iterative process in parallel across grid cells, which are independent of one another.

The main disadvantage of the iterative scaling method is that it does not account for melt and sublimation losses occurring before the snow depth measurement. If losses are substantial, the snowfall scaling overcompensates for the discrepancy between modelled and observed snow depth, as illustrated by the example in Fig. 3b. In this example, the modelled peak snow depth without scaling is 3 m. Before the lidar flight, 1 m of snow depth is lost due to melt and sublimation in the simulation, so that 2 m of snow is predicted at the time of the lidar observation. Because this is twice the measured snow depth of 1 m, the scaling factor becomes 1/2 (see Eq. (4)). Rerunning the model with half the original snowfall yields around 1.5 m of accumulated snow, whilst 1 m is still lost to melt and sublimation, resulting in a simulated depth of 0.5 m at the time of the observation. Thus, after snowfall scaling, the modelled snow depth is half that of the observation (see Fig. 3). Additional iterations cannot improve this result: the second scaling factor is $c_{S,2} = 2 \cdot 0.5 = 1$ (Eq. (7)), which is identical to the unscaled case, and subsequent iterations oscillate between these values without converging. In situations with more severe melt, this overcompensation can even lead to divergence.



2.4.3 New loss-compensated scaling method

240 To prevent overcompensation in the presence of substantial melt and sublimation losses, as occurs with the iterative scaling method, we propose using a model estimate of snow losses to approximate both the observed and the modelled peak snow depth. That is, with the estimates of the losses, we reformulate Eq. (7) and Eq. (8) as follows:

$$c_{LC,j}(i) := \frac{h_{\text{lidar}}(i) + h_{\text{loss},j-1}(i)}{h_{\text{model},LC,j-1}(i) + h_{\text{loss},j-1}(i)} c_{LC,j-1}(i), \quad \forall j \in \{1, 2, \dots\}, \forall i \in \mathcal{V}, \quad (9)$$

$$S_{LC,j}(i, t) := \begin{cases} c_{LC,j}(i) \cdot S(i, t), & \text{if } i \in \mathcal{V}, \\ S(i, t), & \text{otherwise,} \end{cases} \quad \forall j \in \{0, 1, 2, \dots\}, \forall i, t. \quad (10)$$

245 Here, the subscript ‘LC’ stands for the loss-compensated method, $h_{\text{model},LC,j}(i)$ denotes the modelled snow depth obtained with snowfall input $S_{LC,j}(i)$ at the time of the lidar observation, and $h_{\text{loss},j}$ denotes the estimated cumulative losses in this model run until the time of the lidar scan. Thus, the numerator in Eq. (9) represents an estimate of the snow depth at the lidar observation if the losses prior to the observation had not occurred, whilst the denominator represents the modelled peak snow depth. An initial model run ($j = 0$) without scaling ($c_{LC,0}(i) := 1, \forall i$) is used to start the iterative process. The losses are estimated from
 250 the simulation results as

$$h_{\text{loss},j}(i) := \frac{m_j(i) + s_j(i)}{\rho_j(i)}, \quad \forall i, j, \quad (11)$$

where m_j denotes the cumulative snowmelt, s_j the cumulative sublimation, and ρ_j the bulk snow density at the time of the lidar scan in simulation j (with snowfall input $S_{LC,j}$).

Like the iterative scaling method, the loss-compensated method assumes that the accumulated snow depth scales linearly
 255 with the scaling factor. However, it additionally assumes that the losses are independent of the scaling factor. This underlying assumption about the relation between the scaling factor and snow depth is illustrated by the dotted green line in Fig. 3a. Nevertheless, owing to these simplifying assumptions, a small discrepancy can remain between modelled and observed snow depth after the first iteration. Additional iterations can then be applied to further reduce this discrepancy, following Eq. (9) and (10) again. By comparing Eq. (7) and (9), we can see that the multiplicative changes in scaling factor are closer to one
 260 when using the loss-compensated scaling method than when using the iterative scaling method. The loss-compensated method might therefore converge more slowly, but with less risk of overshooting or diverging in further iterations. For this method, we use the same computation procedure and stopping criteria as for the iterative scaling method (see Section 2.4.2).

To illustrate the behaviour of the loss-compensated method, we reconsider the example in Fig. 3b. In this example, the modelled peak snow depth without scaling is 3 m and 1 m of snow is lost to melt and sublimation before the lidar observation,
 265 so that the unscaled model predicts 2 m at the time of the scan. In the loss-compensated method, the true peak snow depth is estimated to be 2 m, derived from the sum of the observed snow depth of 1 m and the model estimated losses of 1 m. The scaling factor from Eq. (9) is the ratio of our best estimate of the true peak snow depth to the model’s prediction; $c_{LC,1} = 2/3$, rather than $1/2$ obtained with the iterative scaling method. Rerunning the model with this factor scales the peak snow depth in the model to approximately $2/3 \cdot 3 \text{ m} = 2 \text{ m}$, aligning it with our best estimate of the peak snow depth based on the observation.



270 Because the modelled losses (1 m) were used to determine this scaling, the model using loss-compensated scaling closely matches the observation of 1 m at the time of the lidar observation.

2.5 Description of simulations and evaluation

To cover all available lidar scans, we perform simulations for Ulla-Førre over the periods 2021-09-01 to 2022-07-01, 2023-09-01 to 2024-07-01, and 2024-09-01 to 2025-07-01, whilst for Songa only the latter two periods are used. We initialise each
 275 simulation with default parameter values and soil states, and assume snow-free conditions at the start of each period. For each period and region, we run simulations without snowfall scaling, and with snowfall adjusted by the direct scaling method, the iterative scaling method, and the loss-compensated scaling method (see Section 2.4). We conduct all experiments using the forcing data presented in Section 2.3, at model resolutions of 90, 300, and 900 m. Finally, we assess whether scaling factors derived from one scan can improve simulation results in other years. For this analysis of transferability across years, we apply
 280 the scaling factors obtained using the methods outlined above to perform simulations in all years for which lidar data are available for evaluation. To be able to compare the results of the different resolutions during evaluation, bias and RMSE are weighted by the area of the lidar data that contributed to each grid cell:

$$\text{Bias}(h_{\text{model}}) := \frac{\sum_{i \in \mathcal{E}} a(i)(h_{\text{model}}(i) - h_{\text{lidar}}(i))}{\sum_{i \in \mathcal{E}} a(i)}, \quad (12)$$

$$\text{RMSE}(h_{\text{model}}) := \sqrt{\frac{\sum_{i \in \mathcal{E}} a(i)(h_{\text{model}}(i) - h_{\text{lidar}}(i))^2}{\sum_{i \in \mathcal{E}} a(i)}}. \quad (13)$$

285 Here, $h_{\text{model}}(i)$ and $h_{\text{lidar}}(i)$ denote the snow height in grid cell i , $a(i)$ denotes the area of the lidar data at 3 m resolution that was used to find $h_{\text{lidar}}(i)$ at the model resolution and $\mathcal{E} := \{i \mid h_{\text{lidar}}(i) \geq 0 \text{ m}\}$ denotes the set of grid cell indices used for evaluation.

3 Results

3.1 Effectiveness of different scaling methods

290 Figure 4 illustrates snow simulation performance without snowfall scaling and with the different snowfall scaling methods for the Ulla-Førre region using one of the lidar scans (22 April 2024) at one spatial resolution (90 m). Without applying scaling, the model qualitatively reproduces the large-scale patterns in snow accumulation (compare Fig. 4a and Fig. 2a). Both the model and observations show the greatest snow depths in the high-elevation regions in the northern part of the area and the shallowest snowpacks in the low-lying western part (see also the site map in Fig. 1). Indeed, elevation alone explains 85.1 % of the variance
 295 in simulated snow depth, whereas it explains only 34.5 % of the variance in observed snow depth. The small-scale patterns, which are not primarily driven by elevation, are not well captured by the model without scaling; the snow depth pattern in the model is much smoother than in the observations (compare Fig. 4b and Fig. 2b). This leads to a low correlation of 0.59 ($r^2 = 0.35$) between modelled and observed snow depth and an RMSE of 1.35 m. Furthermore, the model simulation without



scaling overestimates the observed snow depth substantially (Fig. 4c); the mean snow depth in the simulation is 36.6 % higher
 300 than in the observations (3.15 m and 2.30 m, respectively).

The model with direct scaling produces higher small-scale variability in snow depth than the unscaled simulation (compare
 Fig. 4b and Fig. 4e), capturing small-scale features such as large snow accumulation in trenches and thin snowpack in wind-
 exposed areas that are also visible in the observations (Fig. 2b). However, on the large scale, errors are even greater than those
 of the unscaled run, with substantial underestimation of snow accumulation in the low-lying western region and overestimation
 305 on the eastern high plateau (see Fig. 4f). In fact, in the western region the direct scaling method produces snow-free conditions
 that contradict the thin but persistent snow cover observed there. In addition to this overestimation of the large-scale spatial
 variability of snow depth, the total snow amount is still overestimated. The direct scaling method performs worse than the
 model without scaling in terms of both bias (1.06 m or 46.1 %) and RMSE (1.58 m).

The iterative scaling method of Vögeli et al. (2016) reduces the bias to -0.05 m (-2.0 %), which is substantially lower than
 310 the bias without scaling and with direct scaling. Furthermore, unlike the unscaled simulation and the simulation with direct
 scaling, we find strong visual agreement between this model simulation and the observations when comparing the variability
 of snow depth on both small and large spatial scales (compare Fig. 4g and Fig. 4h to Fig. 2a and Fig. 2b). In most grid cells,
 the errors are much lower than for the two previous methods (see Fig. 4i). Errors fall below 0.10 m in 72.5 % of the grid cells
 after only one model iteration. However, in 9.2 % of the grid cells, the errors exceed 0.25 m, mainly in the low-lying western
 315 part. As shown in Section 3.2, these high errors cannot always be reduced even when increasing the number of iterations.

In contrast to the iterative scaling method, the loss-compensated method provides low errors across the entire domain, also
 in regions where substantial melt occurred (see Fig. 4l). Overall, the bias is 1.0 % after the first iteration, lower than the bias
 resulting from both the direct and iterative scaling method. The large-scale spatial patterns, small-scale patterns, and extent
 of snow free areas closely resemble those of the observations (compare Fig. 4j and Fig. 4k to Fig. 2a and Fig. 2b). With the
 320 loss-compensated method, errors are below 0.10 m in 89.0 % of the grid cells and exceed 0.25 m in only 0.3 % of grid cells
 after one iteration. Furthermore, increasing the number of iterations for the loss-compensated method reduces the errors to
 arbitrarily low values in most grid cells, as shown below in Section 3.2.

3.2 Comparison of the iterative and loss-compensated scaling methods in regions with substantial melt

As outlined above, the difference in performance between the iterative and loss-compensated scaling methods is largest in the
 325 low-lying regions, such as the western part of Ulla-Førre. As shown in Fig. 5a, substantial melt occurred in this region up to
 the time of the lidar scan. We attempted to reduce the errors of the iterative scaling method in these regions by applying more
 iterations (as in Eq. (7)). However, as Fig. 5b shows, the errors cannot be reduced below 0.01 m for many grid cells. In total,
 the method fails to converge to a solution with an error below 0.01 m for 11.4 % of the grid cells even after 100 iterations. The
 grid cells that do not converge to an observation-consistent solution coincide with areas that experienced substantial snowmelt
 330 before the lidar observation and with the grid cells that had the largest errors after the first iteration (compare Fig. 5b to Fig. 4i).
 The average modelled melt that occurred before the lidar scan is 312 mm in the region where the method does not converge,

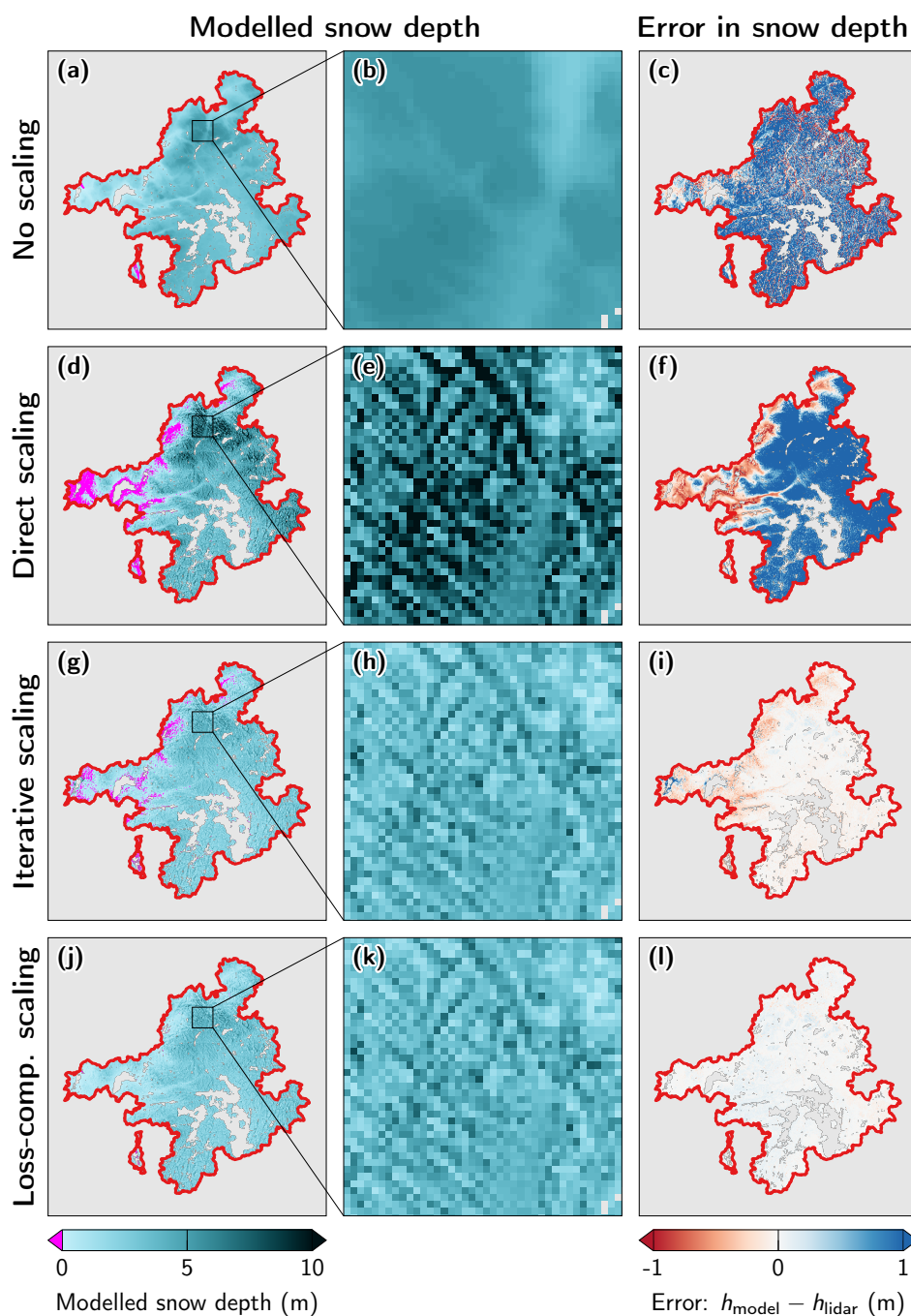


Figure 4. Maps of simulated snow depth and errors with and without snowfall scaling in the Ulla-Førre region on 22 April 2024, at a model resolution of 90 m. Panel a shows the modelled snow depth without snowfall scaling, Panel b shows an inset, indicated by the black square, and Panel c shows the difference in snow depth between the model and the lidar scan. The red outline indicates the catchment boundary. Similarly, Panels d–f and g–i show the results for the direct and iterative scaling methods of Vögeli et al. (2016), respectively. Panels j–l show results from the newly developed loss-compensated method. Here, we show the results after one iteration step for the iterative scaling method and the loss-compensated approach.

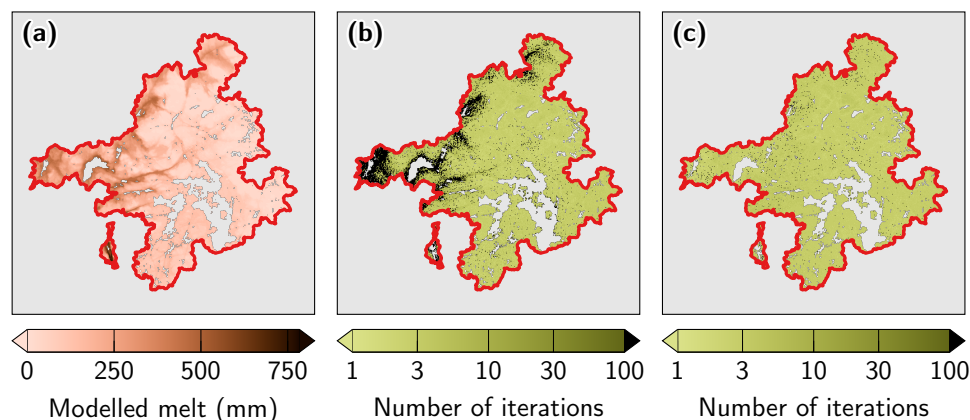


Figure 5. Simulated snow melt accumulated until the date of the lidar scan from 22 April 2024 (Panel a) and the number of iterations required to achieve an error of less than 0.01 m for the iterative scaling method of Vögeli et al. (2016) (Panel b) and the loss-compensated method (Panel c).

whereas the average over the entire domain is 144 mm. At higher elevations, the iterative scaling method appears to be more successful and converges quickly to an observation-consistent solution in most grid cells (see Fig. 5b).

In contrast to the iterative scaling method, the loss-compensated scaling method produces low errors even in regions where substantial melt occurred up to the time of the lidar scan (compare Fig. 4l and 5a). As a quantitative example, in regions with more than 200 mm of melt, the bias of the loss-compensated method is 0.03 m (2.8 %) compared with -0.09 m (-8.0 %) for the iterative method. By contrast, in regions with less than 100 mm of melt, the loss-compensated method performs similarly to the iterative scaling method (with biases of 0.01 m and -0.02 m, respectively). Furthermore, by applying more iterations, the loss-compensated method reduces the error to below 0.01 m in 98.5 % of the grid cells. On average, over all grid cells, 2.9 iterations are required to achieve these low errors. However, in 1.5 % of the grid cells, the loss-compensated method also fails to converge to a solution with an error below 0.01 m. As Fig. 5c shows, the grid cells where the method fails to converge to an observation-consistent solution are roughly evenly scattered over most of the domain, in high-elevation regions, the method is successful in almost all grid cells.

3.3 Performance of scaling methods across regions, years, and resolutions

To study whether the performance of the different scaling methods generalises across regions, years, and resolutions, Fig. 6 summarises the results of all model experiments. Without scaling, the model substantially overestimates snow depth in Ulla-Førre in all simulated years, with biases between 0.66 m and 0.84 m. In Songa, the model is closer to the observations, with biases below 0.10 m. This large systematic overestimation in Ulla-Førre also leads to a substantially higher RMSE than in Songa. For both regions, the bias is similar across resolutions, whilst the RMSE is higher at finer resolutions, indicating that small-scale variability in observed snow depth is not well captured by the unscaled simulation.



In the Ulla-Førre region, the model performance with direct scaling is worse than that of the unscaled simulation in terms of bias and RMSE across almost all years and resolutions (see Fig. 6). The bias increases by up to 0.33 m (41.1 %) and the RMSE by up to 0.68 m compared with the unscaled simulation. In the Songa region, however, the bias with direct scaling is similar to that of the model without scaling (up to 0.10 m and 0.14 m, respectively), whilst the method improves RMSE, especially at the
355 finest resolution of 90 m, where RMSE is reduced by up to 0.51 m. When direct scaling is applied, RMSE does not strongly depend on resolution, in contrast to the unscaled simulations. This shows that the model captures small-scale variability in snow depth well with direct scaling.

The iterative scaling method of Vögeli et al. (2016) already produces very low errors after the first iteration, particularly in the Songa region, with biases below 0.01 m and RMSE below 0.08 m. In the Ulla-Førre region, the results are more mixed;
360 biases are lower than when no scaling is applied, ranging from -0.05 m in 2024 to 0.14 m at the highest resolution in 2025. The RMSE after the first iteration is usually lower than that without scaling, except in 2025, for which the RMSE at the two highest resolutions greatly exceeds the unscaled case. This high RMSE is caused by extremely high errors in individual grid cells, arising from completely unrealistic scaling factors. Furthermore, errors persist in 2024 and 2025 in the Ulla-Førre region when more than one iteration is applied (as in Eq. (7)): convergence to an error below 0.01 m is not achieved in up to 11.6 %
365 of the grid cells. In the Songa region, applying more iterations is more successful than in the Ulla-Førre region, with at least 98.2 % of grid cells converging to an error below 0.01 m.

Unlike the iterative scaling method, the loss-compensated method produces low errors for all simulations even after the first iteration. With this method, unrealistic scaling factors are avoided and the simulations consistently appear robust, even in the Ulla-Førre region in 2025. Occasionally, such as in Ulla-Førre in 2022, the loss-compensated method produces slightly
370 higher errors in terms of bias and RMSE than the iterative scaling method after the first iteration (see Fig. 6). Nevertheless, when multiple iterations are applied, the loss-compensated method clearly outperforms the iterative method, converging to an observation-consistent solution in nearly all grid cells for all simulations. Where both methods converge, however, the loss-compensated method often requires slightly more iterations than the iterative scaling method (on average 3.3 and 3.0 iterations, respectively). For both methods, the number of iterations until an observation-consistent solution is found tends to increase
375 slightly with model resolution.

3.4 Transferability of scaling factors across years

Figure 7 shows the modelling results for 2024 using scaling factors obtained using the loss-compensated method from the lidar scans in this and other years. As shown before, the simulations nearly replicate the lidar scan when the scaling factors are derived from measurements in the simulated year (see Fig. 7c and d). When the scaling factors from 2022 or 2025 are
380 used instead, the large-scale snow distribution resembles that of the observation in 2024 (compare Fig. 7e and 7i with Fig. 7a). However, simulations relying on the 2022 scaling factors overestimate snow depth in lower-elevation regions whilst slightly underestimating it at higher elevations (Fig. 7g). When using the 2025 scaling factors, the simulations tend to underestimate the observations across most elevations (Fig. 7k). Still, when using scaling factors from 2022 or 2025, the correlation between simulation and observations is 0.88 and 0.91, respectively, which is considerably larger than the correlation of 0.59 for the un-



Region	Scan date	Res. (m)	No scaling		Direct scaling		Iterative scaling				Loss-comp. scaling			
			Bias (m)	RMSE (m)	Bias (m)	RMSE (m)	Bias ₁ (m)	RMSE ₁ (m)	% converg.	Mean n_{iter}	Bias ₁ (m)	RMSE ₁ (m)	% converg.	Mean n_{iter}
Ulla-Førre	2022-3-19	900	0.74	0.87	0.91	1.18	0.03	0.05	99	2.82	0.13	0.15	99	3.97
		300	0.75	0.97	0.93	1.21	0.03	0.05	99	2.86	0.13	0.16	99	3.99
		90	0.75	1.29	0.87	1.19	0.02	0.07	98	2.90	0.12	0.21	99	4.06
Ulla-Førre	2024-4-22	900	0.82	0.93	1.16	1.61	-0.05	0.13	88	3.32	0.03	0.05	99	2.94
		300	0.84	1.04	1.15	1.61	-0.05	0.22	89	3.39	0.03	0.06	99	2.93
		90	0.84	1.35	1.06	1.58	-0.05	0.57	89	3.52	0.02	0.08	99	2.94
Ulla-Førre	2025-3-14	900	0.66	0.75	0.86	1.20	0.01	0.27	89	3.36	0.11	0.12	100	3.79
		300	0.67	0.81	0.87	1.22	0.09	3.37	89	3.32	0.11	0.13	99	3.78
		90	0.68	0.94	0.83	1.21	0.14	10.64	89	3.31	0.10	0.15	99	3.74
Songa	2024-4-22	900	0.09	0.35	0.12	0.34	0.01	0.04	98	2.75	0.01	0.03	99	2.64
		300	0.10	0.50	0.14	0.35	0.01	0.05	98	2.83	0.01	0.04	99	2.71
		90	0.10	0.84	0.11	0.33	0.00	0.08	98	2.98	0.00	0.07	99	2.83
Songa	2025-3-19	900	0.10	0.29	0.10	0.29	0.00	0.02	100	2.48	0.01	0.05	100	3.03
		300	0.10	0.42	0.12	0.32	0.00	0.02	99	2.54	0.01	0.08	100	3.14
		90	0.10	0.68	0.11	0.33	0.00	0.04	99	2.65	0.01	0.12	100	3.35

Figure 6. Bias and RMSE of the different snowfall scaling methods evaluated at the time of the assimilated scan. For the iterative scaling method of Vögeli et al. (2016) and the loss-compensated method, these statistics are reported after the first iteration. The effectiveness of further iterations is evaluated by the percentage of grid cells that converge to an error below 0.01 m within 100 iterations and by the average number of iterations required to achieve convergence in these grid cells.

385 scaled simulation. In terms of bias and RMSE, the performance is also substantially better than that of the unscaled simulation, with a reduction in bias of at least 82 % and in RMSE of at least 53 % in 2024.

At small scales, the simulation results using the scaling factors from 2022 or 2025 also largely reproduce the observed patterns in 2024 (compare Fig. 7f and j with Fig. 7b). However, as with the large-scale snow distribution, differences between the simulations also exist here: in particular, the results relying on the 2025 scaling factors tend to underestimate the snow
390 depth variability, and simulation errors can be locally large (see Fig. 7h and l). Nevertheless, when considering the probability density function (pdf) of simulated snow depths over the complete domain, scaling factors from other years provide a substantial improvement compared with the unscaled simulation (see Fig. 7n), even though the observed pdf of snow depths varies considerably between years (see Fig. 7m). Thus, scaling factors derived from the loss-compensated method capture important

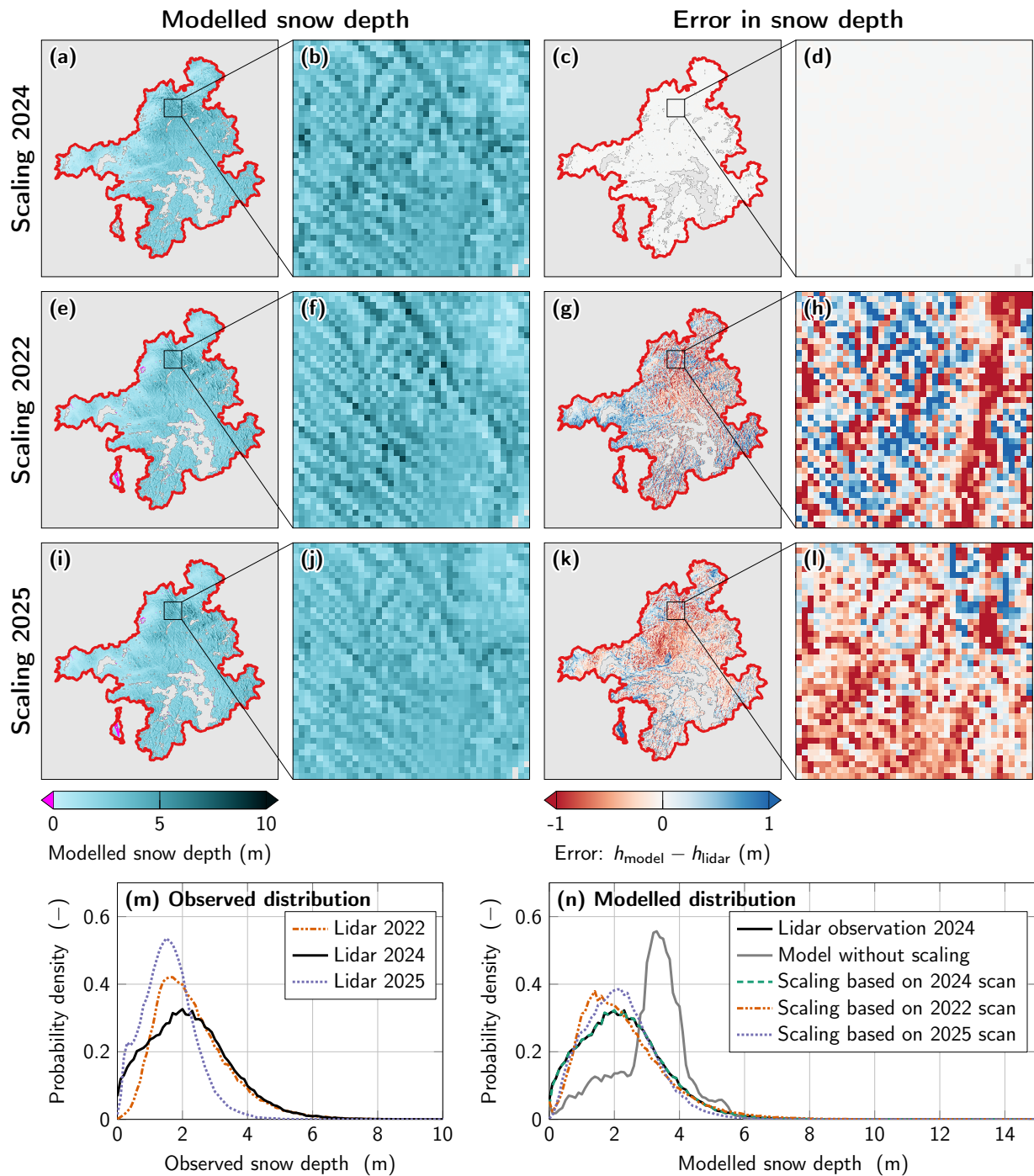


Figure 7. Maps of simulated snow depth and related errors on 22 April 2024 using scaling factors derived from the lidar scan from 22 April 2024 (Panels a–d), 19 March 2022 (Panels e–h), and 14 March 2025 (Panels i–l) and probability density functions of lidar-observed snow depth (Panel m) and simulated snow depth with and without scaling, compared with the 2024 lidar observation (Panel n). All results are shown for the 90 m grid resolution and were obtained using the scaling factors from the final iteration of the loss-compensated method.

patterns from a lidar scan that can provide substantial improvement in other years compared with applying no scaling, although the best simulation results are obtained using scaling factors derived from scans in the simulated year itself.

Figure 8 summarises the simulation performance when transferring the scaling factors across years for all methods. Overall, only the loss-compensated method shows a consistent improvement in performance compared to the unscaled simulation and the best performance is achieved using the scaling factors derived from lidar scans in the simulated year. When considering the simulation results that rely on scaling factors from other years, the loss-compensated method systematically reduces the bias by at least 74 % in Ulla-Førre and 67 % in Songa, while the RMSE is reduced by at least 50 % in all evaluations. For the direct scaling method, the performance in Ulla-Førre is worse than when no scaling is applied. For Songa, on the other hand, direct scaling shows improved RMSE at this resolution (90 m), whilst the bias remains essentially unchanged. Finally, for the iterative scaling method of Vögeli et al. (2016), simulation performance clearly deteriorates when using the scaling factors obtained for 2025 in the Ulla-Førre region, both in 2025 and in other years. When basing the simulations on the scaling factors derived from the remaining years, simulation performance becomes broadly similar to that of the loss-compensated method. For the Songa region, the simulations relying on the iterative scaling method perform as well as the loss-compensated method.

			Assimilation										
			No scaling	Direct scaling			Iterative scaling			Loss-comp. scaling			
				2022	2024	2025	2022	2024	2025	2022	2024	2025	
Evaluation	Ulla-Førre	Bias	2022-03-19	0.75	0.87	0.91	0.91	0.02	-0.06	0.18	0.00	-0.01	-0.09
			2022-06-05	0.66	0.85	1.00	0.98	-0.03	-0.02	0.18	-0.04	-0.02	-0.14
			2024-04-22	0.84	0.99	1.06	1.06	0.05	-0.05	0.14	0.03	0.00	-0.08
			2025-03-14	0.68	0.78	0.83	0.83	0.10	0.03	0.14	0.08	0.07	0.00
	RMSE	2022-03-19	1.29	1.19	1.57	1.50	0.07	0.78	19.93	0.01	0.57	0.59	
		2022-06-05	1.12	1.21	1.48	1.40	0.23	0.66	20.94	0.25	0.47	0.47	
		2024-04-22	1.35	1.50	1.58	1.59	0.61	0.57	21.29	0.63	0.04	0.50	
		2025-03-14	0.94	1.19	1.31	1.21	0.45	0.47	10.64	0.47	0.37	0.01	
	Songa	Bias	2024-04-22	0.10		0.11	0.10		0.00	-0.04		0.00	-0.03
			2025-03-19	0.10		0.10	0.11		0.03	0.00		0.03	0.00
RMSE		2024-04-22	0.84		0.33	0.58		0.08	0.36		0.01	0.38	
		2025-03-19	0.68		0.34	0.33		0.26	0.04		0.26	0.01	

Figure 8. Bias and RMSE (both in m) of simulated snow depth for the model run without scaling and with scaling factors obtained from the three scaling methods (see Section 2.4). Columns indicate the year of the lidar scan near peak accumulation from which the scaling factors were derived; rows indicate the lidar scan against which the simulations are evaluated, for both study regions. All results are shown for the 90 m grid resolution. For the iterative scaling method of Vögeli et al. (2016), the scaling factors obtained after the first iteration are used since divergence in further iterations deteriorates the results. For the loss-compensated method, the scaling factors after the last iteration are used. There was no scan in Songa in 2022, therefore there are empty cells for this region-year combination.



4 Discussion

In our study areas in South Norway, the FSMOSHD model with unadjusted MET Nordic input data is unable to accurately capture the snow amounts and their spatial variability. This model without scaling has two major problems. First, there is a large positive bias in the Ulla-Førre region. The bias is smaller in the Songa region. This difference likely reflects spatial variability in the quality of the forcing data, particularly precipitation. We used a precipitation product with assimilated observations (see Section 2.3), the accuracy of which can vary spatially due to differences in station density and measurement quality. If no scaling is applied, these inaccuracies result in biased snow estimates. Second, the unadjusted model also fails to capture the small-scale variability of snow depth, as processes such as preferential deposition and redistribution are not represented in the FSMOSHD model version used in this study.

We tested different snowfall scaling methods to reduce biases and introduce spatial variability in the snowpack. However, currently available methods are not always capable of deriving scaling factors that result in a good match between the modelled and observed snow depth at the time of the lidar observation. The direct scaling method by Vögeli et al. (2016) cannot correct biases and overestimates spatial variability on large scales. This overestimation is more severe in our study regions than in previous work (e.g. Brauchli et al., 2017; Kiewiet et al., 2022; Pflug et al., 2021; Vögeli et al., 2016) because the large spatial extent of our study regions makes it common for a single precipitation event to fall as snow at high elevations but as rain in lower-lying areas. As a result, high elevations experience more snowfall events, which are scaled by a factor greater than one, whilst the fewer snowfall events at lower elevations are scaled by a factor smaller than one (as a consequence of Eq. (3) and a positive trend in snow depth with elevation). This produces an unrealistic amplification of the large-scale spatial variability of snow depth, which is more pronounced in Ulla-Førre than in Songa. The iterative scaling method corrects for biases and improves the representation of spatial variability on both small and large scales. However, systematic errors occur when using this method in regions where melt occurred before the lidar observation, due to the overcompensation in the snowfall scaling factors described in Section 2.4.2. Although Vögeli et al. (2016) applied the iterative scaling method only for a single iteration, we examined whether multiple iterations could reduce the remaining discrepancies between simulation and observation. This was not possible in regions with substantial melt because the method did not converge to an observation-consistent scaling factor in these areas. In summary, the iterative scaling method outperforms the direct scaling method but fails in regions where substantial melt has occurred before the lidar acquisition. Therefore, neither method is appropriate for applications where reliable results are essential, especially if part of the region is affected by melt before the lidar observation.

The loss-compensated method we introduce in this study effectively integrates lidar scans into the model, even if melt occurred, by considering a model estimate of the losses that took place before the lidar observation. Specifically, snowfall is scaled by the ratio $(h_{\text{lidar}} + h_{\text{loss}})/(h_{\text{model}} + h_{\text{loss}})$. Since losses are non-negative, this ratio is closer to one than the ratio $h_{\text{lidar}}/h_{\text{model}}$ used in the iterative scaling method. Therefore, overcompensation is avoided in regions where melt occurred. Furthermore, the smaller changes to the scaling factor between iterations help the method to converge to an observation-consistent solution (< 0.01 m error) in nearly all grid cells, although the number of iterations required is often slightly higher than for the iterative scaling method (see Fig. 6). These additional iterations do not strongly impact the total computation

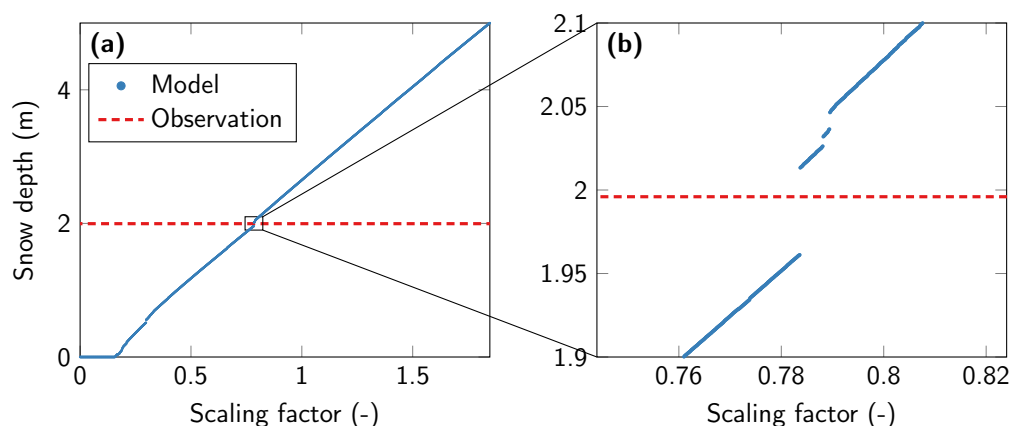


Figure 9. Relationship between scaling factor and simulated snow depth in a grid cell where a discontinuity in the simulation occurs at the lidar-observed snow depth. Here, a step size of 0.0001 in the scaling factor is used between simulations. The maximum difference in simulated snow depth between two consecutive snowfall scaling factors is 0.05 m (a jump from 1.96 m to 2.01 m). Panel b shows a zoom of Panel a around this largest discontinuity.

time of the method since our technical implementation is more efficient than previous implementations of scaling methods; all forcing data for a grid cell is loaded before inferring the scaling factor, so further iterations do not require loading the meteorological data again (see Section 2.4.2). Overall, the loss-compensated method produces substantially better results than both the direct and iterative scaling methods, with a computation time comparable to that of the iterative method.

Our results show that the loss-compensated method still fails to converge to an error below 0.01 m in approximately 1 % of the grid cells (see Fig. 6). To understand why this non-convergence occurs, Fig. 9 shows the relationship between snowfall scaling factor and modelled snow depth for an example grid cell. Simulated snow depth increases with the snowfall scaling factor, but not continuously. Discontinuities in this relationship prevent any scaling factor from reducing the error between model and observation to less than 0.01 m (see Fig. 9b). Consequently, no scaling method can find a factor that reproduces the observed snow depth at the time of the lidar scan. Such discontinuities occur in physics-based snow models because their layering schemes, constraints, and parametrisations impose thresholds on snow depth or snow water equivalent that determine which process descriptions are applied. A slight change in input can trigger a different layering or activate a different parametrisation at some point during the simulation, resulting in a step change in modelled snow depth after this point in time. In principle, these discontinuities could be smoothed, as Kavetski and Kuczera (2007) showed for a simple temperature index model. For a physics-based model, the large number of discontinuous descriptions makes it difficult to apply such smoothing. Nevertheless, the effect of the discontinuities on the resulting scaling factors and snow depth usually remains limited; in our study regions, the maximum deviation between measured and modelled snow depth after scaling with the loss-compensated method was at most 0.07 m.



Overall, with the loss-compensated method, simulated snow depth can be matched to the observations with very low errors
460 at the time of the lidar scan. We tested the method with the FSMOSHD snow model, but it can also be applied to reduce
biases and introduce spatial variability in other distributed snow models that describe snow height, density, and snow losses,
as long as they do not already account for snow redistribution processes between grid cells. Nevertheless, some limitations
of snowfall scaling methods remain. First, errors in simulated variables other than snow depth can still be large. For snow
water equivalent, for instance, errors are now dominated by errors in modelled snow density instead of snow depth (Raleigh
465 and Small, 2017). Furthermore, snowfall scaling methods implicitly assume that errors in modelled losses are much smaller
than errors in snow accumulation. However, when estimating peak snow water equivalent, errors in modelled snow losses can
become significant if the lidar scan is acquired well after the onset of snowmelt. Conversely, if the lidar scan is performed too
early, the scaling factors may represent individual snowfall events instead of the snow distribution at peak accumulation. This
is undesirable since the distribution at peak of winter is most important for snow-hydrological forecasting and is often more
470 consistent between years than individual events (Pflug et al., 2021; Schirmer et al., 2011). Therefore, acquiring the lidar scans
near peak accumulation is optimal to increase generalisability and reduce errors in peak snow water equivalent estimation
arising from, for example, inaccurate simulation of snow losses. Still, it is highly valuable to be able to effectively use lidar
scans to improve model performance even if these scans have been substantially affected by snow losses. This is particularly
true in our study regions, where adverse weather conditions and persistent cloud cover can prevent airborne laser scanning
475 for weeks, and the peak of winter is later in the high mountains than in the lower elevated regions. For these reasons, the
development of the loss-compensated method, which produces accurate results even in the presence of melt, is relevant to
many practical applications.

When applying scaling factors obtained using the loss-compensated method to a different year, the simulated snow depth
patterns resemble, but do not exactly match, the observed patterns (Fig. 7). These differences likely arise from year-to-year
480 variations in dominant wind direction during and after snowfall events, as well as in avalanche activity, both of which alter
small-scale deposition patterns. Crucially for snow-hydrological modelling, the probability density functions of modelled snow
depth closely match those observed when using the loss-compensated approach (Fig. 7n). This modelled distribution largely
governs the timing of snowmelt contributions to runoff and must therefore be accurately simulated in a hydrological forecasting
setting, such as hydropower production planning. The close match between the modelled and observed probability density
485 functions is non-trivial, given that the observed snow depth distributions differ markedly between years (Fig. 7m). Furthermore,
the cross-year transferability analysis for the loss-compensated method shows a reduction of at least 67 % in bias and 50 %
in RMSE for snow depth in our study regions compared to the unscaled simulation (see Fig. 8). The iterative and direct
scaling methods perform worse than the loss-compensated method. The iterative method occasionally fails to infer correct
scaling factors, resulting in large errors in every year. The direct scaling method performs more similarly between years when
490 transferring scaling factors, as shown in earlier studies (Pflug et al., 2021; Vögeli et al., 2016). However, as the performance of
this method is generally much lower than that of the other scaling methods, we cannot recommend its use, especially in large
study regions. Overall, our analysis shows that scaling factors provide valuable information for correcting spatial snowfall



patterns, even in years other than that of the lidar acquisition, provided they are correctly inferred by an appropriate method, such as the loss-compensated approach.

495 In this study, we have shown that the loss-compensated method is an effective and operationally viable approach for integrating lidar data into physics-based snow modelling for large basins. However, lidar scans are expensive and not always possible to conduct due to adverse weather conditions. Thus, developing methods to improve snow simulations that rely on only a few past scans, or that can take advantage of scans covering only a small part of a basin, is an important goal for operational applications. Our study contributes to this goal: we show that improved snow simulations can be achieved by transferring scaling factors across years. Furthermore, with a library of lidar scans from past years, different sets of snowfall scaling factors can be computed to create an ensemble of snow simulations for the current year. Such an ensemble can support operational inflow forecasts with a more realistic uncertainty in initial snow states. It would be possible to select the best ensemble members using measurements from manual ground observations or automatic weather stations to further increase the performance, but the best simulation results will still be obtained using a lidar scan from the current year. To reduce costs, future work could focus on correcting snowfall scaling factors derived from past years using limited-area scans from the current year. A related question is what percentage of the catchment should be covered by such partial scans, what the optimal location for such measurements is, and whether it is possible to use drone-based acquisition to reduce costs and increase weather independence through lower flight altitudes. Another potential direction for future research is to use terrain characteristics to infer scaling factors and predict snowfall in regions lacking lidar data. This would make it possible to determine suitable scaling factors in small areas where there is no lidar data, for example, due to localised cloud cover. Finally, current snowfall scaling methods neglect additional types of observations, in particular, ground measurements of snow density. Such data should be incorporated to reduce the residual errors in estimated peak snow water equivalent, the most important variable for inflow forecasts in snow-dominated regions.

5 Conclusions

515 In this study, we corrected the snowfall input to physics-based snow models using existing and newly developed snowfall scaling methods. The study covers two large basins, Ulla-Førre and Songa, in South Norway. Simulations using snowfall fields obtained with the direct scaling method of Vögeli et al. (2016) reproduced observed small-scale snow depth patterns, but could not reliably integrate the assimilated lidar observations at larger scales. Simulations using snowfall fields obtained with the iterative scaling method of Vögeli et al. (2016) provided more accurate results than those relying on the direct scaling method. The simulation captured the basin-averaged snow depth well, and small-scale snow depth patterns mostly matched those of observations. However, in regions where substantial melt occurred before the lidar observation, this method produced unrealistic snowfall scaling factors that resulted in erroneous snow simulations. Thus, for large mountainous basins where snow often starts to melt in the low-lying areas before peak accumulation is reached in higher elevated areas, neither the direct nor the iterative scaling method can be used to obtain reliable snowfall corrections from lidar observations.



525 Based on the strengths and weaknesses of existing methods, we developed a new scaling method inspired by the iterative scaling method of Vögeli et al. (2016). This newly developed method finds the optimal scaling factor by comparing the model results with the observations whilst compensating for losses due to snowmelt and sublimation that occurred before the lidar survey. With this loss-compensated scaling method, we consistently obtained accurate results for all regions, years, and model resolutions, converging to arbitrarily small errors in almost all grid cells. In a few grid cells, errors between the observation and
 530 the simulation remained due to discontinuities in the physics-based snow model, but these were limited to a few centimetres. Overall, the loss-compensated method produced optimal snowfall corrections, and we consider it a clear advance over existing snowfall scaling methods.

Finally, we evaluated the transferability of scaling factors across years by using one year's scan to correct snowfall in other years. Although model performance was best when a scan from the same year was used, applying scaling factors from
 535 a different year still reduced the bias in snow depth by at least 67 % and the RMSE by 50 % relative to simulations with uncorrected forcing data. After such corrections to the snowfall input, the simulated probability density of snow depth, which is relevant for predicting the timing of runoff, also matched the observations well. Consistent improvements in simulation performance were achievable only when the snowfall corrections were derived with the loss-compensated method.

To conclude, the newly developed method is suitable for determining snowfall corrections from lidar observations over large
 540 mountainous basins. The snowfall scaling factors obtained with this method can be used to improve the snowfall input to models, both for the same winter in which the lidar scan was performed and for other winters. This is promising, for example, for hydropower operations, where the improved description of the snowpack can enhance forecasts of runoff volume and timing during multiple melt seasons. Future research could further improve the quality of these runoff forecasts by incorporating additional measurement types, such as snow density and runoff observations. Another potential future research direction would
 545 be to develop snowfall correction methods that combine full-area scans from past years with limited-area scans from the current winter to achieve optimal model performance at minimal cost.

Code and data availability. The code for the analyses in this paper, including the implementations of the described scaling methods, is available here: <https://www.envidat.ch/#/review/0e20436c-7359-417e-b5de-68484ab80c0b> (Oprel et al., 2026). For snow modelling, the code uses the snow model FlexibleSnowModelOSHD.jl, version 0.1.0 (Magnusson et al., 2026).

550 For three subareas, each covering 13 km², we also provide all processed data required to run the scaling experiments, including meteorological input data and lidar snow depth measurements. These data allow the results shown in Figures 2b, 2e, 4b, 4e, 4h, 4k, 7b, 7f and 7j to be reproduced and provide equivalent plots for a lower-elevated area in Ulla-Forre. From these results, the same qualitative conclusions can be drawn as in this paper.

Full-area meteorological input data are available from the Norwegian Meteorological Institute. We used the rerun archive version 3
 555 for data before 2023-01-31 (<https://thredds.met.no/thredds/catalog/metpparchivev3/catalog.html>; Last accessed: 2026-10-06); Norwegian Meteorological Institute, 2023) and the operational archive for later dates (<https://thredds.met.no/thredds/catalog/metpparchive/catalog.html>; Last accessed: 2026-10-06; Norwegian Meteorological Institute, 2026). Elevation data and N50 Map Data used for the outlines of natural



and regulated lakes are available for the entire study areas from the Norwegian Mapping Authority (<https://data.geonorge.no/sosi/hoydedata> and <https://data.geonorge.no/sosi/kartdata>; Last accessed: 2026-10-06; Norwegian Mapping Authority, 2026a, b).

560 Lidar data for the entire study areas will be available from Statkraft AS in 2028 (contact: gaute.lappegard@statkraft.com). The full-area lidar data are commercially sensitive, as recent snow observations can provide an advantage in energy-market operations. Statkraft AS, which owns the data, has therefore placed them under embargo until 2028.

Author contributions. JO and JM conceptualised the study. The new scaling method was developed and implemented by JO, meteorological downscaling methods were developed by TJ, and post-processing of the lidar scans was carried out by FM. All authors contributed to the
565 discussion and interpretation of the results. AS, FM, and JM acquired the funding, and the project was administered by GL. MB and JM supervised the work. JO wrote the manuscript with help from JM, and all other authors reviewed and edited the manuscript.

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

Acknowledgements. This work was supported by the Research Council of Norway and Statkraft AS under the SnowInflow project (NFR 346308). The authors would like to thank Geir Nævdal and Fabio Zeiser for inspiring discussions and suggestions that improved the
570 manuscript.



References

- Aalstad, K., Westermann, S., Schuler, T. V., Boike, J., and Bertino, L.: Ensemble-Based Assimilation of Fractional Snow-Covered Area Satellite Retrievals to Estimate the Snow Distribution at Arctic Sites, *The Cryosphere*, 12, 247–270, <https://doi.org/10.5194/tc-12-247-2018>, 2018.
- 575 Andreadis, K. M. and Lettenmaier, D. P.: Assimilating Remotely Sensed Snow Observations into a Macroscale Hydrology Model, *Advances in Water Resources*, 29, 872–886, <https://doi.org/10.1016/j.advwatres.2005.08.004>, 2006.
- Anghileri, D., Voisin, N., Castelletti, A., Pianosi, F., Nijssen, B., and Lettenmaier, D. P.: Value of Long-term Streamflow Forecasts to Reservoir Operations for Water Supply in Snow-dominated River Catchments, *Water Resources Research*, 52, 4209–4225, <https://doi.org/10.1002/2015WR017864>, 2016.
- 580 Baba, M. W., Gascoin, S., and Hanich, L.: Assimilation of Sentinel-2 Data into a Snowpack Model in the High Atlas of Morocco, *Remote Sensing*, 10, 1982, <https://doi.org/10.3390/rs10121982>, 2018.
- Berg, J., Reynolds, D., Quéno, L., Jonas, T., Lehning, M., and Mott, R.: A Seasonal Snowpack Model Forced with Dynamically Downscaled Forcing Data Resolves Hydrologically Relevant Accumulation Patterns, *Frontiers in Earth Science*, 12, 1393260, <https://doi.org/10.3389/feart.2024.1393260>, 2024.
- 585 Brauchli, T., Trujillo, E., Huwald, H., and Lehning, M.: Influence of Slope-Scale Snowmelt on Catchment Response Simulated With the *Alpine3D* Model, *Water Resources Research*, 53, 10723–10739, <https://doi.org/10.1002/2017WR021278>, 2017.
- Bruland, O.: Present and Future Hydropower Scheduling in Statkraft, in: *AGU Fall Meeting Abstracts*, vol. 2012, pp. H43J–05, <https://ui.adsabs.harvard.edu/abs/2012AGUFM.H43J..05B>, 2012.
- Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Örn Hreinsson, E., and Woods, R. A.: Representing Spatial Variability of Snow Water Equivalent in Hydrologic and Land-Surface Models: A Review, *Water Resources Research*, 47, <https://doi.org/10.1029/2011WR010745>, 2011.
- 590 Crank, J. and Nicolson, P.: A Practical Method for Numerical Evaluation of Solutions of Partial Differential Equations of the Heat-Conduction Type, *Advances in Computational Mathematics*, 6, 207–226, <https://doi.org/10.1007/BF02127704>, 1996.
- Crumley, R. L., Bachand, C. L., and Bennett, K. E.: Snow Distribution Patterns Revisited: A Physics-Based and Machine Learning Hybrid Approach to Snow Distribution Mapping in the Sub-Arctic, *Water Resources Research*, 60, e2023WR036180, <https://doi.org/10.1029/2023WR036180>, 2024.
- 595 Currier, W. R., Pflug, J., Mazzotti, G., Jonas, T., Deems, J. S., Bormann, K. J., Painter, T. H., Hiemstra, C. A., Gelvin, A., Uhlmann, Z., Spaete, L., Glenn, N. F., and Lundquist, J. D.: Comparing Aerial Lidar Observations With Terrestrial Lidar and Snow-Probe Transects From NASA’s 2017 SnowEx Campaign, *Water Resources Research*, 55, 6285–6294, <https://doi.org/10.1029/2018WR024533>, 2019.
- 600 Davidson-Chaput, J., Arsenault, R., Martel, J.-L., and Troin, M.: Value of Various Elements of the Hydrological Forecasting Chain: Is There a Successful Pathway for Improving the Overall Performance?, *Journal of Hydrology*, 658, 133186, <https://doi.org/10.1016/j.jhydrol.2025.133186>, 2025.
- Deems, J. S., Fassnacht, S. R., and Elder, K. J.: Interannual Consistency in Fractal Snow Depth Patterns at Two Colorado Mountain Sites, *Journal of Hydrometeorology*, 9, 977–988, <https://doi.org/10.1175/2008JHM901.1>, 2008.
- 605 Deng, G., Liu, X., Shen, Q., Zhang, T., Chen, Q., and Tang, Z.: Remote Sensing Data Assimilation to Improve the Seasonal Snow Cover Simulations Over the Heihe River Basin, Northwest China, *International Journal of Climatology*, 44, 5621–5640, <https://doi.org/10.1002/joc.8656>, 2024.



- Egli, L., Jonas, T., Grünwald, T., Schirmer, M., and Burlando, P.: Dynamics of Snow Ablation in a Small Alpine Catchment Observed by Repeated Terrestrial Laser Scans, *Hydrological Processes*, 26, 1574–1585, <https://doi.org/10.1002/hyp.8244>, 2012.
- 610 Eresmaa, R., Andrae, U., Berggren, L., Bremnes, J. B., Fortelius, C., Frogner, I.-L., Gregow, E., Grote, R., Ridal, M., Vignes, O., Ansper, I., Azad, R., Begens, J., Bergholt, L., Blyverket, J., Daniel, L., Engdahl, B. J., Erlandsen, H., Eshagh, M., Hasu, M., Homleid, M., Idestrom, P., Ivarsson, K.-I., Lerner-Vilu, A., Mets, A., Moldengauere, S., Noer, G., Nurmi, E., Partio, M., Prieditis, G., Saarikalle, E., Schönach, D., Shapkalijevski, M., Sild, K., Sitčihins, V., Sokka, N., Spjelkavik, S., Tack, A., Vaštšenko, A., Wettergren, A., Ylinen, K., and Zandovskis, U.: The Operational Forecast Process at MetCoOp, *Bulletin of the American Meteorological Society*, 107, E946–E963, <https://doi.org/10.1175/BAMS-D-25-0225.1>, 2026.
- 615 Essery, R.: A Factorial Snowpack Model (FSM 1.0), *Geoscientific Model Development*, 8, 3867–3876, <https://doi.org/10.5194/gmd-8-3867-2015>, 2015.
- Freudiger, D., Kohn, I., Seibert, J., Stahl, K., and Weiler, M.: Snow Redistribution for the Hydrological Modeling of Alpine Catchments, *WIREs Water*, 4, e1232, <https://doi.org/10.1002/wat2.1232>, 2017.
- 620 Gerber, F., Mott, R., and Lehning, M.: The Importance of Near-Surface Winter Precipitation Processes in Complex Alpine Terrain, *Journal of Hydrometeorology*, 20, 177–196, <https://doi.org/10.1175/JHM-D-18-0055.1>, 2019.
- Griessinger, N., Schirmer, M., Helbig, N., Winstral, A., Michel, A., and Jonas, T.: Implications of Observation-Enhanced Energy-Balance Snowmelt Simulations for Runoff Modeling of Alpine Catchments, *Advances in Water Resources*, 133, 103410, <https://doi.org/10.1016/j.advwatres.2019.103410>, 2019.
- 625 Günther, D., Marke, T., Essery, R., and Strasser, U.: Uncertainties in Snowpack Simulations—Assessing the Impact of Model Structure, Parameter Choice, and Forcing Data Error on Point-Scale Energy Balance Snow Model Performance, *Water Resources Research*, 55, 2779–2800, <https://doi.org/10.1029/2018WR023403>, 2019.
- Haberkorn, A., Wever, N., Hoelzle, M., Phillips, M., Kenner, R., Bavay, M., and Lehning, M.: Distributed Snow and Rock Temperature Modelling in Steep Rock Walls Using Alpine3D, *The Cryosphere*, 11, 585–607, <https://doi.org/10.5194/tc-11-585-2017>, 2017.
- 630 Hale, K., Kiewiet, L., Trujillo, E., Krohe, C., Hedrick, A., Marks, D., Kormos, P., Havens, S., McNamara, J., Link, T., and Godsey, S.: Drivers of Spatiotemporal Patterns of Surface Water Inputs in a Catchment at the Rain-Snow Transition Zone of the Water-Limited Western United States, *Journal of Hydrology*, 616, 128699, <https://doi.org/10.1016/j.jhydrol.2022.128699>, 2023.
- Hedrick, A. R., Marks, D., Havens, S., Robertson, M., Johnson, M., Sandusky, M., Marshall, H.-P., Kormos, P. R., Bormann, K. J., and Painter, T. H.: Direct Insertion of NASA Airborne Snow Observatory-Derived Snow Depth Time Series Into the *iSnobal* Energy Balance Snow Model, *Water Resources Research*, 54, 8045–8063, <https://doi.org/10.1029/2018WR023190>, 2018.
- 635 IEA: Energy System of Norway, <https://www.iea.org/countries/norway>, 2025.
- Kavetski, D. and Kuczera, G.: Model Smoothing Strategies to Remove Microscale Discontinuities and Spurious Secondary Optima in Objective Functions in Hydrological Calibration, *Water Resources Research*, 43, 2006WR005195, <https://doi.org/10.1029/2006WR005195>, 2007.
- 640 Kiewiet, L., Trujillo, E., Hedrick, A., Havens, S., Hale, K., Seyfried, M., Kampf, S., and Godsey, S. E.: Effects of Spatial and Temporal Variability in Surface Water Inputs on Streamflow Generation and Cessation in the Rain–Snow Transition Zone, *Hydrology and Earth System Sciences*, 26, 2779–2796, <https://doi.org/10.5194/hess-26-2779-2022>, 2022.
- Koch, F., Gascoin, S., Achmüller, K., Schattan, P., Wetzel, K.-F., Deschamps-Berger, C., Lehning, M., Rehm, T., Schulz, K., and Voigt, C.: Superconducting Gravimeter Observations Show That a Satellite-Derived Snow Depth Image Improves the Simulation of the Snow Water



- 645 Equivalent Evolution in a High Alpine Site, *Geophysical Research Letters*, 51, e2024GL112483, <https://doi.org/10.1029/2024GL112483>, 2024.
- Leisenring, M. and Moradkhani, H.: Snow Water Equivalent Prediction Using Bayesian Data Assimilation Methods, *Stochastic Environmental Research and Risk Assessment*, 25, 253–270, <https://doi.org/10.1007/s00477-010-0445-5>, 2011.
- Lundquist, J. D. and Dettinger, M. D.: How Snowpack Heterogeneity Affects Diurnal Streamflow Timing, *Water Resources Research*, 41, 2004WR003649, <https://doi.org/10.1029/2004WR003649>, 2005.
- 650 Lussana, C., Seierstad, I. A., Nipen, T. N., and Cantarello, L.: Spatial Interpolation of Two-metre Temperature over Norway Based on the Combination of Numerical Weather Prediction Ensembles and *in Situ* Observations, *Quarterly Journal of the Royal Meteorological Society*, 145, 3626–3643, <https://doi.org/10.1002/qj.3646>, 2019.
- Magnusson, J., Wever, N., Essery, R., Helbig, N., Winstral, A., and Jonas, T.: Evaluating Snow Models with Varying Process Representations for Hydrological Applications, *Water Resources Research*, 51, 2707–2723, <https://doi.org/10.1002/2014WR016498>, 2015.
- 655 Magnusson, J., Winstral, A., Stordal, A. S., Essery, R., and Jonas, T.: Improving Physically Based Snow Simulations by Assimilating Snow Depths Using the Particle Filter, *Water Resources Research*, 53, 1125–1143, <https://doi.org/10.1002/2016WR019092>, 2017.
- Magnusson, J., Eisner, S., Huang, S., Lussana, C., Mazzotti, G., Essery, R., Saloranta, T., and Beldring, S.: Influence of Spatial Resolution on Snow Cover Dynamics for a Coastal and Mountainous Region at High Latitudes (Norway), *Water Resources Research*, 55, 5612–5630, <https://doi.org/10.1029/2019WR024925>, 2019.
- 660 Magnusson, J., Bühler, Y., Quéno, L., Cluzet, B., Mazzotti, G., Webster, C., Mott, R., and Jonas, T.: High-Resolution Hydrometeorological and Snow Data for the Dischma Catchment in Switzerland, *Earth System Science Data*, 17, 703–717, <https://doi.org/10.5194/essd-17-703-2025>, 2025.
- Magnusson, J., Haugeneder, M., Mazzotti, G., Oprel, J., and Oberrauch, M.: FlexibleSnowModelOSHD.Jl: V0.1.0, Zenodo, <https://doi.org/10.5281/ZENODO.21322825>, 2026.
- 665 Marsh, C. B., Lv, Z., Vionnet, V., Harder, P., Spiteri, R. J., and Pomeroy, J. W.: Snowdrift-Permitting Simulations of Seasonal Snowpack Processes Over Large Mountain Extents, *Water Resources Research*, 60, e2023WR036948, <https://doi.org/10.1029/2023WR036948>, 2024.
- Mazzotti, G., Essery, R., Moeser, C. D., and Jonas, T.: Resolving Small-Scale Forest Snow Patterns Using an Energy Balance Snow Model With a One-Layer Canopy, *Water Resources Research*, 56, e2019WR026129, <https://doi.org/10.1029/2019WR026129>, 2020.
- 670 Mott, R., Winstral, A., Cluzet, B., Helbig, N., Magnusson, J., Mazzotti, G., Quéno, L., Schirmer, M., Webster, C., and Jonas, T.: Operational Snow-Hydrological Modeling for Switzerland, *Frontiers in Earth Science*, 11, 1228158, <https://doi.org/10.3389/feart.2023.1228158>, 2023.
- Nester, T., Kirnbauer, R., Parajka, J., and Blöschl, G.: Evaluating the Snow Component of a Flood Forecasting Model, *Hydrology Research*, 43, 762–779, <https://doi.org/10.2166/nh.2012.041>, 2012.
- 675 Neuville, A., Båserud, L., Nipen, T. N., Seierstad, I. A., and Lussana, C.: MET Nordic Analysis of Hourly Precipitation over Scandinavia, in: EMS Annual Meeting Abstracts 2025, vol. 22, pp. EMS2025–202, Ljubljana, Slovenia, <https://doi.org/10.5194/ems2025-202>, 2025.
- Norwegian Mapping Authority: Height DTM 50, <https://data.geonorge.no/sosi/hoydedata>, (Last accessed: 2026-10-06), 2026a.
- Norwegian Mapping Authority: N50 Map Data, <https://data.geonorge.no/sosi/kartdata>, (Last accessed: 2026-10-06), 2026b.
- Norwegian Meteorological Institute: MET Nordic Reanalysis Version 3, <https://thredds.met.no/thredds/catalog/metpparchive3/catalog.html>, (Last accessed: 2026-10-06), 2023.
- 680 Norwegian Meteorological Institute: MET Nordic Operational Archive, <https://thredds.met.no/thredds/catalog/metpparchive/catalog.html>, (Last accessed: 2026-10-06), 2026.



- Oberrauch, M., Cluzet, B., Magnusson, J., and Jonas, T.: Improving Fully Distributed Snowpack Simulations by Mapping Perturbations of Meteorological Forcings Inferred From Particle Filter Assimilation of Snow Monitoring Data, *Water Resources Research*, 60, e2023WR036994, <https://doi.org/10.1029/2023WR036994>, 2024.
- Opel, J., Jonas, T., Stordal, A. S., Matt, F. N., Lappegard, G., Brunner, M. I., and Magnusson, J.: Code and Data to Demonstrate Loss-Compensated Snowfall Scaling Using Large-Area Lidar Observations in South Norway, <https://doi.org/10.16904/envidat.793>, During the review process, code and data are temporarily available at: <https://www.envidat.ch/#/review/0e20436c-7359-417e-b5de-68484ab80c0b>, 2026.
- Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R., Marks, D., Mattmann, C., McGurk, B., Ramirez, P., Richardson, M., Skiles, S. M., Seidel, F. C., and Winstral, A.: The Airborne Snow Observatory: Fusion of Scanning Lidar, Imaging Spectrometer, and Physically-Based Modeling for Mapping Snow Water Equivalent and Snow Albedo, *Remote Sensing of Environment*, 184, 139–152, <https://doi.org/10.1016/j.rse.2016.06.018>, 2016.
- Pflug, J. M., Hughes, M., and Lundquist, J. D.: Downscaling Snow Deposition Using Historic Snow Depth Patterns: Diagnosing Limitations From Snowfall Biases, Winter Snow Losses, and Interannual Snow Pattern Repeatability, *Water Resources Research*, 57, e2021WR029999, <https://doi.org/10.1029/2021WR029999>, 2021.
- Pfohl, A. K. D., Currier, W. R., Boardman, E. N., Pflug, J. M., and Abel, M. R.: Investigating the Effects of Spatial Snow Water Equivalent Heterogeneity on Runoff Timing and Magnitude, *Hydrological Processes*, 40, e70659, <https://doi.org/10.1002/hyp.70659>, 2026.
- Pulka, T., Hermegger, M., Ehrendorfer, C., Lüking, S., Avanzi, F., Formayer, H., Schulz, K., and Koch, F.: Evaluating Precipitation Corrections to Enhance High-Alpine Hydrological Modeling, *Journal of Hydrology*, 645, 132202, <https://doi.org/10.1016/j.jhydrol.2024.132202>, 2024.
- Quéno, L., Mott, R., Morin, P., Cluzet, B., Mazzotti, G., and Jonas, T.: Snow Redistribution in an Intermediate-Complexity Snow Hydrology Modelling Framework, *The Cryosphere*, 18, 3533–3557, <https://doi.org/10.5194/tc-18-3533-2024>, 2024.
- Raleigh, M. S. and Small, E. E.: Snowpack Density Modeling Is the Primary Source of Uncertainty When Mapping Basin-wide SWE with Lidar, *Geophysical Research Letters*, 44, 3700–3709, <https://doi.org/10.1002/2016GL071999>, 2017.
- Raleigh, M. S., Lundquist, J. D., and Clark, M. P.: Exploring the Impact of Forcing Error Characteristics on Physically Based Snow Simulations within a Global Sensitivity Analysis Framework, *Hydrology and Earth System Sciences*, 19, 3153–3179, <https://doi.org/10.5194/hess-19-3153-2015>, 2015.
- Revuelto, J., Vionnet, V., López-Moreno, J.-I., Lafaysse, M., and Morin, S.: Combining Snowpack Modeling and Terrestrial Laser Scanner Observations Improves the Simulation of Small Scale Snow Dynamics, *Journal of Hydrology*, 533, 291–307, <https://doi.org/10.1016/j.jhydrol.2015.12.015>, 2016.
- Richter, B., Schweizer, J., Rotach, M. W., and Van Herwijnen, A.: Modeling Spatially Distributed Snow Instability at a Regional Scale Using Alpine3D, *Journal of Glaciology*, 67, 1147–1162, <https://doi.org/10.1017/jog.2021.61>, 2021.
- RIEGL Laser Measurement Systems GmbH: RIEGL VQ-1560 II-S, https://newsroom.riegl.international/wp-content/uploads/2020/11/RIEGL_VQ-1560II-S_Preliminary-Datasheet_2020-10-05_01.pdf, 2020.
- Rudisill, W. J., Flores, A. N., Marshall, H. P., Siirila-Woodburn, E., Feldman, D. R., Rhoades, A. M., Xu, Z., and Morales, A.: Cold-Season Precipitation Sensitivity to Microphysical Parameterizations: Hydrologic Evaluations Leveraging Snow Lidar Datasets, *Journal of Hydrometeorology*, 25, 143–160, <https://doi.org/10.1175/JHM-D-22-0217.1>, 2024.
- Schirmer, M., Wirz, V., Clifton, A., and Lehning, M.: Persistence in Intra-annual Snow Depth Distribution: 1. Measurements and Topographic Control, *Water Resources Research*, 47, 2010WR009426, <https://doi.org/10.1029/2010WR009426>, 2011.



- Sexstone, G. A., Driscoll, J. M., Hay, L. E., Hammond, J. C., and Barnhart, T. B.: Runoff Sensitivity to Snow Depletion Curve Representation within a Continental Scale Hydrologic Model, *Hydrological Processes*, 34, 2365–2380, <https://doi.org/10.1002/hyp.13735>, 2020.
- Statkraft AS: Kvittdal Power Plant, <https://www.statkraft.com/about-statkraft/where-we-operate/norway/saurdal-hydropower-plant/>, 2025.
- Vögeli, C., Lehning, M., Wever, N., and Bavay, M.: Scaling Precipitation Input to Spatially Distributed Hydrological Models by Measured
725 Snow Distribution, *Frontiers in Earth Science*, 4, <https://doi.org/10.3389/feart.2016.00108>, 2016.
- Warscher, M., Strasser, U., Kraller, G., Marke, T., Franz, H., and Kunstmann, H.: Performance of Complex Snow Cover Descriptions in a Distributed Hydrological Model System: A Case Study for the High Alpine Terrain of the Berchtesgaden Alps, *Water Resources Research*, 49, 2619–2637, <https://doi.org/10.1002/wrcr.20219>, 2013.
- Wolff, M. A., Isaksen, K., Petersen-Øverleir, A., Ødemark, K., Reitan, T., and Brækkan, R.: Derivation of a New Continuous Adjustment
730 Function for Correcting Wind-Induced Loss of Solid Precipitation: Results of a Norwegian Field Study, *Hydrology and Earth System Sciences*, 19, 951–967, <https://doi.org/10.5194/hess-19-951-2015>, 2015.