Using snow depth ~~observation~~ observations to provide insight into the quality of ~~regional-scale~~ snowpack simulations for regional-scale avalanche forecasting

Simon Horton[1,2] and Pascal Haegeli[1]

[1]Simon Fraser University, Burnaby, BC, Canada
[2]Avalanche Canada, Revelstoke, BC, Canada

Correspondence: Simon Horton (shorton@avalanche.ca)

Abstract. The combination of numerical weather prediction and snowpack models has potential to provide valuable information about snow avalanche conditions in remote areas. However, the output of snowpack models is sensitive to precipitation inputs, which can be difficult to verify in mountainous regions. To examine how existing observation networks can help interpret the accuracy of snowpack models, we compared snow depths predicted by a weather-snowpack model chain with data from automated weather stations and manual observations. Data from the 2020-21 winter were compiled for 21 avalanche forecast regions across western Canada covering a range of climates and observation networks. To perform regional-scale comparisons, ~~snowpack~~ SNOWPACK model simulations were run at select grid points from the HRDPS numerical weather prediction model to represent conditions at treeline elevations and observed snow depths were ~~interpolated~~ upscaled to the same locations. Snow depths in the Coast Mountain range were systematically overpredicted by the model, while snow depths in many parts of the interior Rocky Mountain range were underpredicted. ~~The impact of these biases~~ These discrepancies had a greater impact on ~~the simulated avalanche~~ simulated snowpack conditions in the interior ranges, where faceting was more sensitive to snow depth. To put the comparisons in context, the quality of the upscaled observations were assessed with ~~uncertainties in the interpolations and~~ by checking whether snow depth ~~increases~~ changes during stormy periods were consistent with the forecast avalanche hazard. While some regions had high quality observations, ~~many regions had large uncertainties~~other regions were poorly represented by available observations, suggesting in some situations ~~the~~ modelled snow depths could be more reliable than ~~the~~ observations. The analysis provides insights into the potential for validating weather and snowpack models with readily available observations, and for how avalanche forecasters can better interpret the accuracy of snowpack simulations.

1 Introduction

Numerical weather prediction (NWP) models provide valuable information to avalanche forecasters, as avalanche conditions are heavily influenced by the evolution of weather patterns. With years of operational experience, forecasters develop a grounded understanding of how well specific NWP models predict weather in their local mountains. Predicting snowpack conditions in a similar way may be possible by forcing snowpack evolution models such as SNOWPACK or Crocus with

~~weather data~~ output from NWP models (Morin et al., 2020). However, developing operational trust and understanding in snowpack models is difficult due to the complexity and spatial variability of ~~the snowpack~~mountain snowpacks.

Efforts to verify snowpack models have faced challenges due to ~~uncertainties with verification datasets and a lack of objective verification frameworks (Morin et al., 2020). Approaches in academic literature have included verifying model output with regional-scale~~ various sources of uncertainty in the models, the verification data, and their spatial presentiveness (Morin et al., 2020) . Snowpack models have been evaluated with different types of verification data including avalanche hazard assessments (Bellaire et al., 2017; Giraud et al., 1987), ~~satellite-retrieved optical properties (Charrois et al., 2016; Cluzet et al., 2020)~~snowpack stability tests (Reuter et al., 2015), snow profile stratigraphy observations (Bellaire and Jamieson, 2013; Brun et al., 1992; Durand et al., 1999; Lehning et al., 2002; Viallon-Galinier et al., 2020), ~~snowpack stability tests (Reuter et al., 2015), and with snow height~~ satellite-retrieved optical properties (Charrois et al., 2016; Cluzet et al., 2020), and snow depth and snow water equivalent observations ~~(Bellaire et al., 2011; Durand et al., 2009; Lafaysse et al., 2013; ?; Winstral et al., 2018). These studies primarily represent case studies at specific locations for specific periods when detailed verification data were collected. While these types of studies are valuable for assessing the general skill of snowpack models, they offer limited operational insight about when snowpack models can be trusted for avalanche forecasting. Producing such operational insight requires a~~ (Bellaire et al., 2011; Durand et al., 2009; Lafaysse et al., 2013). Many of these datasets have limited application for real-time ~~verification framework based on continuous data streams.~~ monitoring because they are point scale observations and required dedicated data collection campaigns. Quéno et al. (2016) and Vionnet et al. (2019) performed regional-scale comparisons of modelled snow depths with spatially distributed point observations and Winstral et al. (2018) and Cluzet et al. (2022) have recently developed systems to assimilate snow depth observations into snowpack models. Real-time model verification systems would be valuable for avalanche forecasters learning how to interpret snowpack model output.

Several types of ~~observation data~~ data streams are available to avalanche forecasters, including manual observations of weather, snowpack, and avalanches from field observers and continuous data ~~streams~~ from automated weather stations and avalanche detection networks. While each of these data streams could ~~provide verification data for~~ be used to verify aspects of snowpack models, all observation networks are limited in their spatial-temporal coverage and their representativeness of regional-scale avalanche conditions. Lundquist et al. (2020) identify similar challenges with observation networks for ~~related~~ applications such as mountain hydrology and ecology, and therefore conclude that in many contexts the skill in modelling mountain precipitation is now comparable ~~or superior~~ to observation networks~~in many contexts. Accordingly, in some contexts,~~ . The quality and density of avalanche-related observations vary by region, but it could be misleading to assume available field observations are more representative ~~than model simulations~~of regional-scale conditions than models in contexts where observations are sparse or irregular.

~~Sensitivity studies consistently find precipitation input~~ While snowpack models are sensitive to all weather input variables, precipitation is consistently identified as the main driver of uncertainty in ~~snowpack models~~ the simulated stratigraphy (Raleigh et al., 2015; Richter et al., 2020). Observations of winter precipitation are available from different types of measurements including cumulative precipitation from rain gauges, snow water equivalent from snow pillows, and snow depth from acoustic sensors or manual probing (Wang et al., 2017). Of these, snow depth observations are typically the most abundant and representative type of precipitation observation available in most avalanche ~~forecasting~~ forecast regions, and accordingly could be a relatively simple data stream to perform operational ~~snowpack model~~ verification.

In this study, snow ~~depth observations were~~ depths from a NWP and snowpack model chain are compared to snow depth observations compiled from existing networks ~~used by avalanche forecasters~~ across western Canada~~to assess the accuracy of snowpack models forced with NWP inputs~~. The objective was to investigate ~~the potential of verifying snowpack models in real-time with snow depth observations~~how these observations could help avalanche forecasters interpret the accuracy of snowpack models at regional scales. This included assessing the reliability of the observations, comparing modelled and observed snow depths across space and time~~at a regional-scale~~, and investigating the impacts of incorrect snow depths on the resulting snowpack stratigraphy. The results provide insights into how avalanche forecasters can better interpret the accuracy of snowpack simulations in data-sparse regions and highlight the potential for snow depth observations to verify and improve NWP and snowpack models in mountainous terrain.

2 Data

The three key datasets for this study were snowpack simulations, snow depth observations, and avalanche hazard assessments. This section outlines how each of these datasets were compiled to evaluate regional-scale snowpack simulations across western Canada for the 2020-21 winter season. The ~~study period~~ analysis was restricted to ~~four winter months~~ a four-month period when regular manual observations and hazard assessments were available (~~2020-12-01 to 2021-03-31)~~December 1 to March 31), which limits the scope of the study to the snow accumulation period when avalanche forecasters are primarily interested in dry snow avalanches.

2.1 Study area

Data were compiled for 21 public avalanche forecast regions in British Columbia, Alberta, and Yukon (Fig. 1a). The regions were grouped into three mountain ranges based on their predominant snow climate characteristics – the maritime Coast Mountains, the transitional Columbia Mountains, and the continental Rocky Mountains (Shandro and Haegeli, 2018). To collect data that was representative of avalanche conditions in these regions, we focused the analysis on conditions ~~at~~ near treeline elevations. Avalanche forecasters classify terrain into alpine, treeline, and below treeline bands based on terrain and vegetation characteristics. The snowpack at treeline tends to be sheltered from the wind and therefore more homogenous than in the alpine.

Since there are no strict definitions for the boundaries of the vegetation bands, ~~and there is some subjectivity in how they are used operationally, we had to derive~~ we derived an objective reference for ~~the treeline elevation for the purpose of this study. Publicly available~~ treeline elevations from a digital elevation model and land cover classification data~~were used to search for the upper extent of forested terrain throughout the study area~~. First, pixels with any type of forested land cover were ~~extracted~~ idenitifed from the 30 m resolution 2015 Land Cover of North America (North American Land Change Monitoring System, 2015) ~~. The~~ and the elevation of these pixels was extracted from a 30 arc-second resolution global digital elevation model

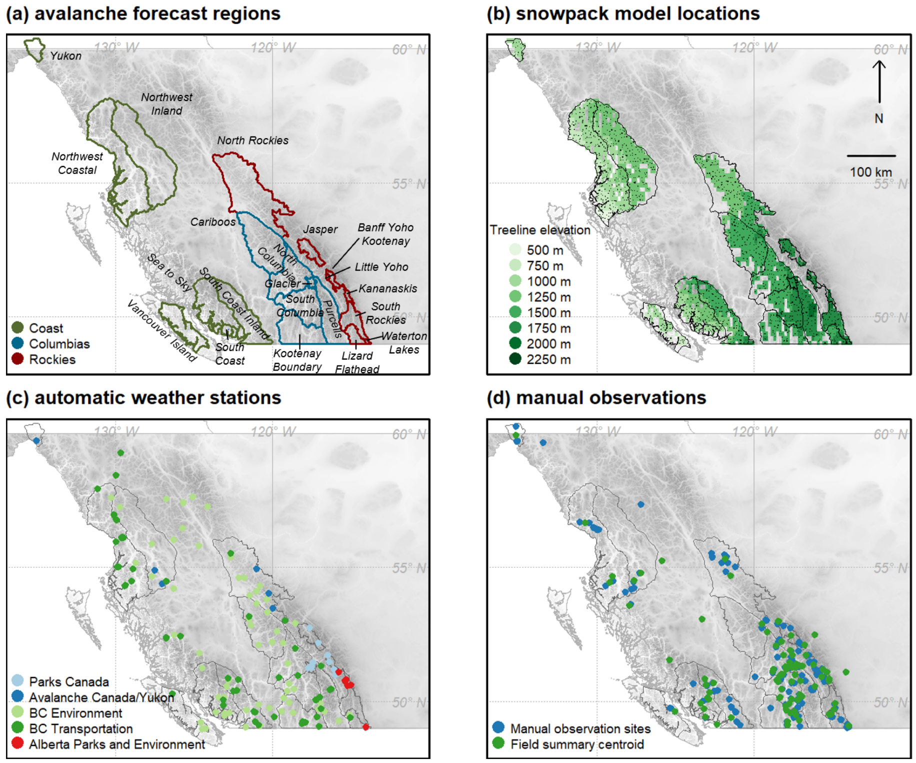

Figure 1. Location of (a) avalanche forecast regions, (b) snowpack model ~~simulation~~ grid points for each 10 x 10 km grid cell colour-coded by local treeline elevation~~at that location~~, (c) automatic weather stations, and (d) manual observations and field summaries.

(Danielson and Gesch, 2011)~~with 30 arc-second resolution (approximately 145 m). The treeline elevation at each pixel was derived by searching for~~ . The elevations were aggregated on a 10 x 10 km grid to determine the maximum elevation of forested terrain within each grid cell, as well as the maximum elevation of ~~forested terrain within a 10 x 10 km moving window across the study~~all terrain within the cell. Local estimates of treeline elevation were taken from any cell where the maximum elevation was at least 250 m above the maximum forest elevation, as these cells likely contained high alpine terrain where trees would not grow. The local treeline estimates were interpolated across the entire study area with a thin spline regression to estimate the upper extent of forested terrain. The resulting elevations ranged from ~~1100~~ 545 m in the ~~northwest to 2200~~ northwestern regions to 2025 m in the ~~southeast~~ southeastern regions (Fig. 1b). Feedback from avalanche forecasters suggested ~~this method produced elevations that~~ the elevations derived from this method roughly aligned with their understanding of vegetation bands in western Canada.

2.2 Snowpack simulations

Snowpack simulations were produced by forcing SNOWPACK (Lehning et al., 1999) with output from the High-Resolution Deterministic Prediction System (HRDPS), an operational NWP model run by the Canadian Meteorological Centre on a 2.5

km horizontal grid (Milbrandt et al., 2016). Simulations were configured to represent treeline snowpack conditions across the 21 avalanche forecast regions.

~~Rather than running simulations at all NWP grid pointswithin the regions, 856 grid points were sampled to represent conditions at treeline (Fig. 1b). The~~ A subset of grid points were sampled to represent regional-scale conditions rather than running SNOWPACK at all available grid points. First, only grid points within 100 vertical metres of the local treeline elevation were considered. Then the forecast regions were split ~~on a~~ into 10 x 10 km grid ~~and then the average treeline elevation was computed in each grid cell. The HRDPS grid point with the closest match to the computed treeline elevation~~ cells and one grid point was selected for each ~~grid cell . Grid cells were discarded if no grid points were within 100 vertical metres of the treeline elevation, which occurred 24% of the time (either because the NWP model topography was smoother than the real terrain or because the grid cell did not actually contain high elevation terrain). Selecting a single grid point in~~ cell by choosing the point with the median accumulated precipitation over the study period. Selecting one point for each 10 x 10 km grid cell was found to effectively reduce redundant information, ~~as there were minimal variations between HRPDS grid points at similar elevations within the same grid cell.~~ while selecting the point with median accumulated precipitation smoothed out small-scale orographic effects resolved by the model. Grid cells without treeline elevation terrain were omitted, resulting in a total of 1260 NWP grid points covering the study area (Fig. 1b).

Snowpack simulations were produced for these 1260 grid points by concatenating forecasts from each operational run of the HRDPS, starting on 1 September 2020. The HRDPS is initialized every six hours, and SNOWPACK was forced with data from the 6 to 12 predictive hours from each run. ~~SNOWPACK~~ Inputs to SNOWPACK included hourly accumulated precipitation, air temperature and relative humidity at 2 m above ground, wind speed at 10 m above ground, and incoming longwave and shortwave radiation fluxes at the surface. SNOWPACK version 3.6 was configured with default settings for flat-field profiles and with wind transport disabled. The height of snowpack (~~HS~~HS), hereafter referred to as snow depth, ~~and the height of new snow over 24 h periods (HN) were output for each day of the study period~~ was output each day at 0 UTC (16:00/17:00 local time).

2.3 Snow depth observations

Snow depth observations were compiled from three sources: automated weather stations, manual observations, and field summaries. The following sections describe how ~~each of these sources were compiled to obtain~~ daily values of snow depth (~~HS) and height of new snow (HN) for each day of the study period at 0 UTC~~HS) were obtained from each source.

2.3.1 Automated weather stations

Weather data was queried from an Avalanche Canada database that aggregates automated weather stations (AWS) deemed relevant for operational avalanche forecasting (Fig. 1c). ~~146 of~~ Of the 238 ~~AWS~~ stations in this database~~are~~, 110 were equipped with acoustic snow depth sensors and were within 500 vertical meters of treeline elevation. These stations ~~are~~ were operated by various agencies including Parks Canada (14 stations), British Columbia Ministry of Transportation and Infrastructure (~~63~~ 25 stations), British Columbia Ministry of Environment (~~56~~ 59 stations), Alberta Parks and Environment (6 stations), Avalanche

Canada (~~5~~ 4 stations), and the Yukon Avalanche Association (2 stations). While some networks already ~~apply~~ applied quality control to the data, an additional spike removal filter was applied to remove observations where ~~HS~~ HS increased or decreased more than 10 cm in one hour. ~~HS recordings~~ HS values were extracted at 0 UTC each day. ~~Calculating height of new snow (HN) from AWS data is challenging due to settlement and instrument noise (Wang et al., 2017). For this study, HN was calculated by summing hourly HS changes over 24 h periods. HN was set to zero when this sum was negative. Summing both positive and negative changes was found to be more reliable than summing only positive changes, as instrument noise frequently resulted in HS increases that would appear as several centimetres of new snow on days when precipitation did not occur. The disadvantage of this method is that HN is underestimated on days with significant snowpack settlement. The~~ The number of daily observations fluctuated ~~over the study period~~ due to sensor and transmission errors, resulting in a median of ~~129 HS observations and 106 HN observations~~ 100 observations per day.

2.3.2 Manual observations

Avalanche safety operations report manual weather observations on the Canadian Avalanche Association's Information Exchange (InfoEx). Weather observations consist of manual measurements taken at fixed instrumented study plots following standards published by the Canadian Avalanche Association (2016) ~~. HS~~ where snow depth is measured on a permanent stake~~and HN is measured with a ruler on a snow board that is cleared daily~~. While observations are typically made twice per day, only afternoon observations were included in this study to be consistent with the timing of the other data sources. Over the entire study period, 94 operations reported observations from ~~197~~ 194 different study plots with known geographic coordinates and elevations (Fig. 1d). The number of ~~daily observations~~ reports submitted fluctuated over the study period, with a median of ~~75 HS observations and 58 HN observations~~ 71 observations per day.

2.3.3 Field summaries

Avalanche operations also report field weather summaries to the InfoEx, which are distinct from weather observations made at fixed study plots. Field weather summaries summarize the range of conditions encountered in a broad geographic area while travelling in the mountains (Canadian Avalanche Association, 2016). Field summaries include subjective and spatially broader estimates of ~~HS and HN~~ HS that are relevant for assessing avalanche hazard in that area, often based on several measurements made throughout the day. The geographic extent of these observations is less precise than manual observations and are defined by polygons covering their operating areas and with elevation ranges that typically span treeline elevation. Fig. 1d shows the polygon centroids for ~~99~~ 98 operations that reported field summaries over the study period. Fewer operations ~~submitted~~ submit field summaries than manual observations, ~~with~~ resulting in a median of ~~19 HS observations and 16 HN observations~~ 18 observations per day.

2.4 Avalanche hazard assessments

Avalanche hazard assessments were compiled to assess how well modelled and observed ~~snow~~ data captured impactful snowfall events. Daily forecasts were compiled for all 21 forecast regions over the study period. Forecasts ~~are~~ were published following the workflow described by the conceptual model of avalanche hazard (Statham et al., 2018), and include a nowcast assessment of current danger ratings and avalanche problems. To quantify ~~the impact of~~ regional-scale snowfall events, days with storm slab avalanche problems at treeline were identified. Storm slab problems have the most direct link with new snow under the North American avalanche problem definitions (Statham et al., 2018), and Horton et al. (2020b) show a strong statistical relationship between storm slab problems and ~~height of new snow~~ new snow depths. Two components from the hazard assessments were extracted to characterize ~~the impacts of these~~ snowfall events: the presence of a storm slab avalanche problem at treeline (a binary value of absent or present), and the danger rating at treeline on days with storm slab problems (ordinal values of 1-Low, 2-Moderate, 3-Considerable, 4-High, and 5-Extreme). The number of days with storm slab problems over the study period ranged from six for the Kananaskis region to 73 for the Northwest Coastal region.

3 Methods

This section describes how the datasets were prepared and analyzed to answer ~~the following~~ three questions:

1. How reliable were the observations for regional-scale assessments?
2. How did modelled and observed snow depths compare across space and time?
3. How do ~~incorrect snow depths impact the modelled~~ snow depth discrepancies impact the interpretation of snowpack stratigraphy?

Fig. 2 summarizes the methods used to answer these questions. First, the observed snow depths were ~~interpolated onto the same grid as the snowpack models to produce daily maps of modelled and observed HS and HN to allow consistent analysis across space and time~~ upscaled to a regional-scale to match the scale and extent of the model data (Sect. ~~??). The~~ 3.1). Then daily snow depth changes were calculated for both datasets (Sect. 3.2) and the reliability of the ~~observations was~~ upscaled observations were assessed with two methods; first, from the uncertainty in the ~~interpolations~~upscaling method, and second, by fitting regression models to see whether avalanche conditions were better explained by ~~height of new snow~~ snow depth changes from the modelled or ~~observed~~ observation data (Sect. 3.3). ~~Basic statistical measures were used to compare the snow depths~~ Snow depths were compared across space and time with basic statistical measures (Sect. 3.4). Finally, to investigate the impact on snowpack stratigraphy, two sets of simulated snow profiles were compared~~; one~~ : one produced with the original NWP model ~~inputs~~ data and the second with ~~bias-corrected~~ adjusted precipitation inputs (Sect. 3.5).

3.1 Upscaling observations to regional-scale

Observed snow depths were ~~aggregated and interpolated to the same spatial grid as the modelled data. The interpolations were produced~~ upscaled to a regional scale in two steps: first a lapse rate adjustment was applied to estimate the observed value at the

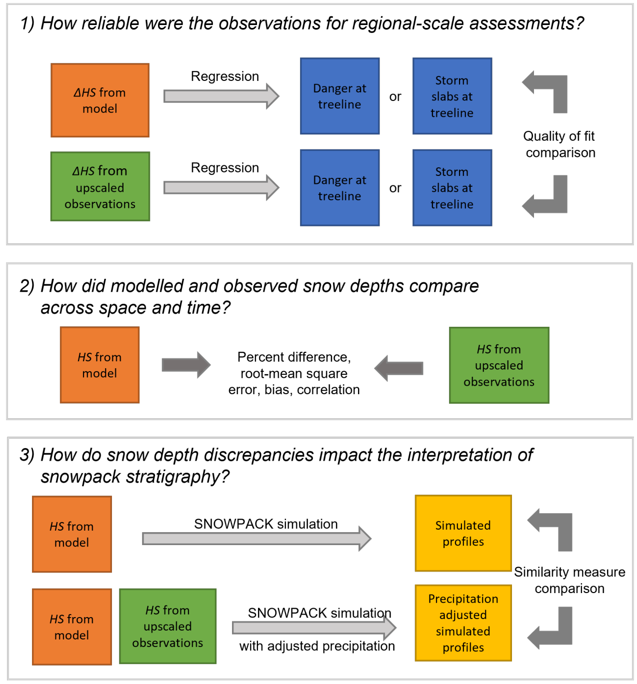

Figure 2. Methods used to analyze the reliability of the observations at regional-scales, compare modelled and observed snow depths, and measure the impact of ~~incorrect~~ snow ~~depths~~ depth discrepancies on snowpack simulations.

local treeline elevation, then a spatial interpolation was applied to predict ~~snow depths at treeline elevations~~ treeline elevation snow depths across the entire study area. The interpolation was not designed to capture fine-scale patterns across terrain, but rather aggregate available observations ~~into a consistent format for regional-scale comparisons~~in an analogous way to how human forecasters interpret point observations in regional hazard assessments.

Elevation ~~corrections~~ adjustments were applied to the AWS and manual observations to account for differences between the observation elevation (z_{obs}) and the ~~treeline elevation at that location~~ local treeline elevation ($z_{treeline}$). Field summaries were not adjusted as their elevations were unknown, but ~~are intended to~~ typically already represent conditions at treeline elevations (Canadian Avalanche Association, 2016). ~~HS~~ HS was lapse rate adjusted with the exponential precipitation adjustment factor proposed by Thornton et al. (1997)~~which has been~~, which is commonly applied in hydrological models (e.g., Liston and Elder, 2006; Schirmer and Jamieson, 2015):

$$HS_{treeline} = HS_{obs}\frac{(1+0.35(z_{treeline}-z_{obs}))}{(1-0.35(z_{treeline}-z_{obs}))} \quad (1)$$

where $HS_{treeline}$ is the adjusted snow depth at treeline derived from ~~an~~ the original observed depth ~~$HS_{treeline}$~~HS_{obs}. A lapse rate factor of 0.35 was suggested for winter months in western USA by Thornton et al. (1997). Since ~~only observations~~ the observations were restricted to those within 500 vertical metres of treeline~~were included,~~, the original snow depths were multiplied by ~~a factor of~~ factors ranging from 0.70 to 1.42. ~~For the entire set of observations, the average snow depth increased~~

~~by a factor of 1.18 after applying treeline adjustments. The same lapse rate adjustment was applied to height of new snow where HS in Eq. 1 was replaced with HN.~~

Spatial interpolations were applied to the treeline corrected ~~HS and HN~~ HS observations with simple kriging. The gstat package for R was used to fit a unique variogram model for ~~HS and HN~~ HS on each day of the study period to describe the structure of the spatial correlation (Pebesma, 2004). Given the sparsity of observations, the best fitting of three possible variogram models were chosen (spherical, exponential, and pure nugget~~models)~~), where the best fit was chosen according gstat's default method of weighting residuals (i.e., number of pairs in a bin divided by the square of the bin's distance). Simple kriging was applied with the best variogram model to predict ~~HS and HN at all 856~~ HS at all 1260 model grid points for each day of the study period. Each prediction combined all available observations from AWS, manual observations, and field summaries.

The kriging predictions also provided an estimate of prediction variance (σ^2_{HS}), which was used to estimate the uncertainty in ~~observed snow depth. The relative~~ the upscaled observations. Relative kriging standard deviation ($RKSD$) was defined at the square root of the prediction variance divided by the predicted height:

$$RKSD = \frac{\sqrt{\sigma^2_{HS}}}{HS} \tag{2}$$

The $RKSD$ is a relative measure ~~of uncertainty~~ where small values suggest the interpolation error was small relative to the predicted snow depth and large values suggest the error was large relative to the predicted snow depth.

3.2 Calculation daily snow depth change

To identify snowfall events, daily snow depth changes (ΔHS) were calculated by subtracting HS on consecutive days (e.g., Quéno et al., 2016; Vionnet et al., 2019). The objective was to identify snowfall events large enough to impact avalanche conditions, and so negative values of ΔHS were set to zero. For the model data ΔHS was calculated at each grid point for every day of the study period. For the observation data ΔHS was calculated from the original observations for all cases where HS was observed at the same location on consecutive days, and then the same upscaling methods were applied to estimate upscaled ΔHS values at treeline elevations on the same spatial grid as the model data.

3.3 Regression models relating snow depth change to avalanche conditions

To assess the reliability of the modelled and observed data ~~sets~~to represent regional-scale conditions, regression models were fit to predict ~~the regional-scale~~ forecast avalanche conditions from ~~height of new snow . Height of new snow~~ snow depth changes. Snowfall has been shown to be a strong predictor of avalanche ~~release, as numerous studies have highlighted the link between new snow with avalanche danger (?) and~~ hazard, as Schirmer et al. (2009) found three-day accumulations of new snow to be a strong predictor of avalanche danger and Horton et al. (2020b) found three-day snow accumulations to have a strong influence on the presence of storm slab avalanche problems~~(Horton et al., 2020b)~~. Therefore, for this analysis we ~~assume that~~ assumed increases in snow depth should be proportional to the likelihood of storm slab avalanche problems and the likelihood of increased avalanche danger.

To test this assumption, regression models were fit to predict (1) the presence of storm slab problems, and (2) the danger rating using ~~height of new snow over three days (HN(3d))~~ the three-day snow depth change ($\Delta HS(3d)$) as the single predictor. ~~Height of new snow over three days~~ Three-day snow depth change was calculated by ~~adding the current day's HN with the HN from the previous two days.~~ summing ΔHS with a three day moving window over the study period.

For each forecast region, a univariate logistic regression model ~~with a single predictor~~ was fit to predict the probability of a storm slab avalanche problem being present based on ~~the modelled HN(3d)~~ $\Delta HS(3d)$ from model data at grid points in the region and an analogous logistic regression model was fit using ~~the observed HN(3d)~~ $\Delta HS(3d)$ from upscaled observation data at the same ~~grid points~~locations. The goodness of fit of the two models were compared with McFadden's pseudo R-squared measure (an analogous value to the coefficient of determination that can be applied to general linear models) to assess whether the modelled or ~~observed HN(3d)~~ upscaled observations explained storm slab presence better in each forecast region.

A similar approach was used to fit ordinal regression models in each forecast region to predict the probability of the danger rating increasing with ~~HN(3d)~~$\Delta HS(3d)$. These models were restricted to days when storm slab avalanche problems were present, as the danger would more likely be driven by other factors on days without storm slab problems. The ordinal regression models were fit using the cumulative link model from the ordinal package for R (Christensen, 2019). Again, McFadden's pseudo R-squared measure was used to determine which dataset ~~explained danger ratings better~~ provided a better explanation of danger ratings in each region.

3.4 Comparing snow depths

The continuous ~~grids of modelled and observed snow depths~~ snow depth data from the model and upscaled observations allowed plotting spatial comparison on maps and temporal comparisons on timeseries. ~~Modelled and observed snow~~ Snow depths were compared with basic quantitative metrics, including the percent difference to provide a relative measure of the differences, root mean square error to assess the magnitude of the differences, bias to assess the prevailing direction of the differences, and Spearman rank-order correlation coefficient to determine whether a set of modelled and ~~observed snow depths~~ upscaled observation data increased at the same locations or times. These statistics are defined as follows:

$$\text{Percent difference (\%)} = \frac{(HS_{mod} - HS_{obs})}{HS_{obs}} * 100\% \tag{3}$$

$$\text{Root mean square error (m)} = \sqrt{\frac{1}{n}\sum(HS_{mod} - HS_{obs})^2} \tag{4}$$

$$\text{Bias (m)} = \frac{1}{n}\sum(HS_{mod} - HS_{obs}) \tag{5}$$

where HS_{mod} is the modelled snow depth, HS_{obs} is the ~~observed snow depth~~ snow depth from upscaled observations, and n is the number of ~~observations~~observation pairs. The Spearman rank-ordered correlation coefficient is defined as the Pearson correlation between the rank values of HS_{mod} and HS_{obs}, rather than the correlation between the values themselves. ~~The same statistics were calculated for height of new snow HN.~~

3.5 Adjusting precipitation inputs of the simulated snow profiles

To illustrate the impact ~~incorrect snow depths~~ snow depth discrepancies could have on how forecasters interpret the simulated snowpack structure, a sample of locations were chosen to ~~test a simple bias correction method~~compare the snowpack simulations with the original NWP model inputs and profiles with adjusted precipitation inputs. One representative ~~profile~~ location was chosen for each region by selecting the grid point that most frequently had the median modelled snow depth over the study period. A ~~bias correction~~ precipitation adjustment factor (k) was calculated for each location according to

$$k = \overline{\frac{HS_{obs}}{HS_{mod}}} \tag{6}$$

based on the modelled and ~~observed snow depths~~upscaled HS at that location over the study period. Averaging the factor ~~over each day of the study period smoothed out variations that appeared on individual days~~provided a rough approximation of the potential magnitude and direction of precipitation biases in at that location. The hourly precipitation inputs at each grid point location were multiplied by ~~this correction factor, and then~~ k and SNOWPACK was re-run ~~with the bias-corrected precipitation .~~ to produce a second set of profiles. This method assumes the main driver of snow depth differences was biased precipitation inputs. While precipitation biases play an important role in these difference, errors in modelled snow depth are influenced by many processes in snowpack models including the parameterization of snow density (Helfricht et al., 2018), handling of precipitation type, and simulating settlement processes. While this simple precipitation adjustment method is not the state-of-the art in data assimilation (Cluzet et al., 2022; Winstral et al., 2018), it provides a simple approach to analyze the impact of precipitation errors in different regions.

The original and ~~bias-corrected~~ precipitation-adjusted profile simulations were compared numerically by applying the similarity measure introduced by Herla et al. (2021). The similarity measure performs a pairwise comparison of two profiles from a perspective of stratigraphy features relevant to avalanche hazard. To calculate the similarity measure, the algorithm first aligns each layer with dynamic time warping, where the deposition date, grain type, and hardness are used to match layers in the two profiles. After alignment, the algorithm compares the similarity of the grain type and hardness of aligned layers, while putting more weight on weak layers, melt-freeze crusts, and new snow layers. The similarity measure ~~values range~~ ranges from 0 to 1, where 0 corresponds to highly dissimilar profiles and 1 corresponds to identical profiles. The similarity value of the original and ~~bias-corrected~~ precipitation-adjusted profiles were computed for each day of the study period to quantify the impact adjusting precipitation inputs had on the interpretation of avalanche hazard conditions.

4 Results

4.1 Reliability of snow depth observations for regional-scale assessments

Snow depths were ~~interpolated~~ upscaled with greater confidence in areas with dense observation networks~~(Fig. 3)~~, such as the southern end of the Coast range and the central ~~Columbias.~~ Columbia range (Fig. 3). In these areas, the average ~~standard deviation of interpolated~~ $RKSD$ of the snow depths was ~~roughly 30-40~~20-30% of the total estimated snow depth, while

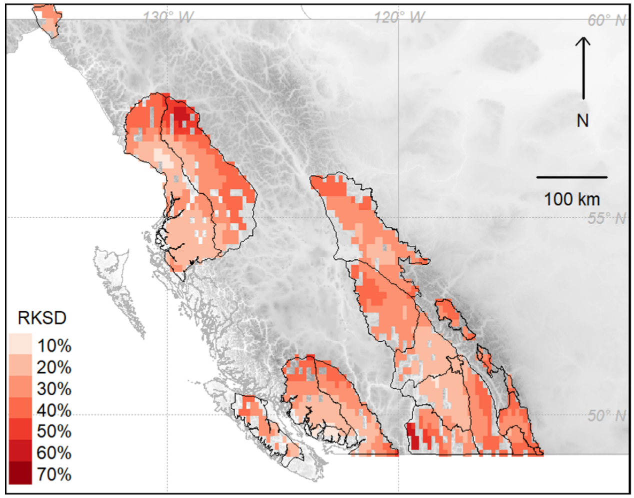

Figure 3. The average relative kriging standard deviation ($RKSD$) over the study period shows areas ~~where there was~~ with relatively high or low confidence in ~~the~~ upscaling observed snow depths ~~interpolated from observations~~to regional-scales.

areas with ~~the greatest uncertainty had average RKSD~~ lowest confidence had $RKSD$ values up to ~~50-60% . Areas with low confidence in interpolated snow depth included northern parts within each coastal region and the southern regions of the Columbias and Rockies~~40-60% of the snow depth. The median ~~RKSD for all 856~~ $RKSD$ for all grid points over the study period was ~~46~~34%, suggesting the aggregation of snow depths from available observations had considerable uncertainty across many forecast regions.

~~Predicting avalanche conditions with height of new snow over three days yielded comparable results with modelled and observed data. Logistic regression models predicting the presence or absence of storm slab avalanche problems performed better with modelled data in 10 of 21 regions (Table 1). This included most coastal regions and several of the southern interior regions. Ordinal regression models predicting danger ratings performed better with modelled data in 11 of 21 regions. This included the same coastal regions, fewer of the southern interior regions, and the addition of some of the northern interior regions.~~

~~In the~~ Predicting avalanche conditions with snow depth changes over three-day periods yielded relatively comparable results with model and upscaled observation data (Table 1). In the Columbia regions the observation data consistently explained the presence of storm slab avalanche problems and danger ratings better than modelled data. However, in several of the Coast and Rockies regions the modelled data explained avalanche conditions better than the observation data, and in many regions the McFadden's pseudo R-squared values were comparable in magnitude.

In the interior, the regions where avalanche conditions were predicted better with ~~observation data~~ upscaled observations were typically the regions ~~relatively smaller RKSD~~ with smaller $RKSD$ values (e.g., Glacier and ~~North~~ South Columbia), suggesting the ~~weather observation networks are relatively~~ observations were relatively consistent and representative of avalanche conditions~~in these regions~~. However, the ~~southern coastal regions~~ South Coast region also had areas with relatively small

Table 1. Comparison of regression models predicting the presence of storm slab avalanche problems and danger ratings with modelled data versus ~~observed snow~~ upscaled observation data. McFadden's pseudo R-squared values are listed for each model to compare their goodness of fit~~with asterisks identifying the~~. The model with ~~a~~ the better fit for each region is identified with an asterisks.

Range	Region	Storm slab problem		Danger rating	
		~~Observed~~ Observations	~~Modelled~~ Model	~~Observed~~ Observations	~~Modelled~~ Model
Coast	Yukon	~~0.05~~0.01	~~0.2~~0.12*	~~0~~0.00	~~0.03~~0.04*
	Northwest Coastal	~~0.29~~0.21	~~0.32~~0.22*	~~0.02~~0.01	~~0.04~~0.05*
	Northwest Inland	~~0.24~~0.12*	~~0.08~~0.07	~~0.04~~0.09*	~~0.04~~ 0.06
	Vancouver Island	~~0.030.07~~0.12*	0.04	~~0.08~~0.07*	0.04
	South Coast	~~0.28~~0.31	0.32*	~~0.01~~0.04	~~0.07~~0.04*
	Sea To Sky	~~0.270.37~~0.34*	~~0.04~~0.33	~~0.07~~0.08*	0.03
	South Coast Inland	~~0.32*~~0.30	~~0.24~~0.18	~~0.06~~0.09*	0.06
Columbias	Cariboos	~~0.36~~0.29*	~~0.29~~0.23	~~0.120.17~~0.21*	0.15
	North Columbia	~~0.38~~0.25*	~~0.34~~0.24	~~0.2~~0.13*	~~0.15~~ 0.11
	Glacier	~~0.4~~0.35*	~~0.32~~0.29	~~0.21~~0.18*	~~0.14~~ 0.08
	South Columbia	~~0.310.32~~0.27*	~~0.2*~~0.23	0.12*	0.08
	Kootenay Boundary	~~0.38~~0.31*	~~0.32~~0.24	~~0.13~~0.07*	~~0.05~~ 0.03
	Purcells	~~0.220.23~~0.24*	~~0.27~~0.21	0.29*	~~0.18~~ 0.14
Rockies	North Rockies	~~0.26~~0.16*	~~0.21~~0.13	~~0.060.06~~0.09*	0.05
	Jasper	~~0.15~~0.16*	~~0.1~~0.12	~~0.06*~~0.02	~~0.05~~ 0.08*
	Banff Yoho Kootenay	~~0.17~~0.07	0.15*	~~0.140.20.23~~0.33*	0.21
	Little Yoho	~~0.26~~0.17	0.29*	~~0.21~~0.33	~~0.26~~0.24*
	Kananaskis	~~0.17*~~0.20	0.14	~~0.31~~1.00*	~~0.3~~ 0.33
	Lizard Flathead	~~0.350.37~~0.46*	0.35	0.04	~~0.09~~0.08*
	South Rockies	~~0.32~~0.33	~~0.36~~0.38*	~~0.030.08~~0.10*	0.09
	Waterton Lakes	~~0.15~~0.13*	~~0.07~~0.10	~~0.22~~0.21*	~~0.09~~ 0.13

~~RKSD~~$RKSD$, yet avalanche conditions were predicted better with modelled ~~snowfall~~snow depth changes, suggesting that even though the ~~weather~~ observation networks were relatively consistent they may ~~be~~ have been less representative of conditions in avalanche terrain.

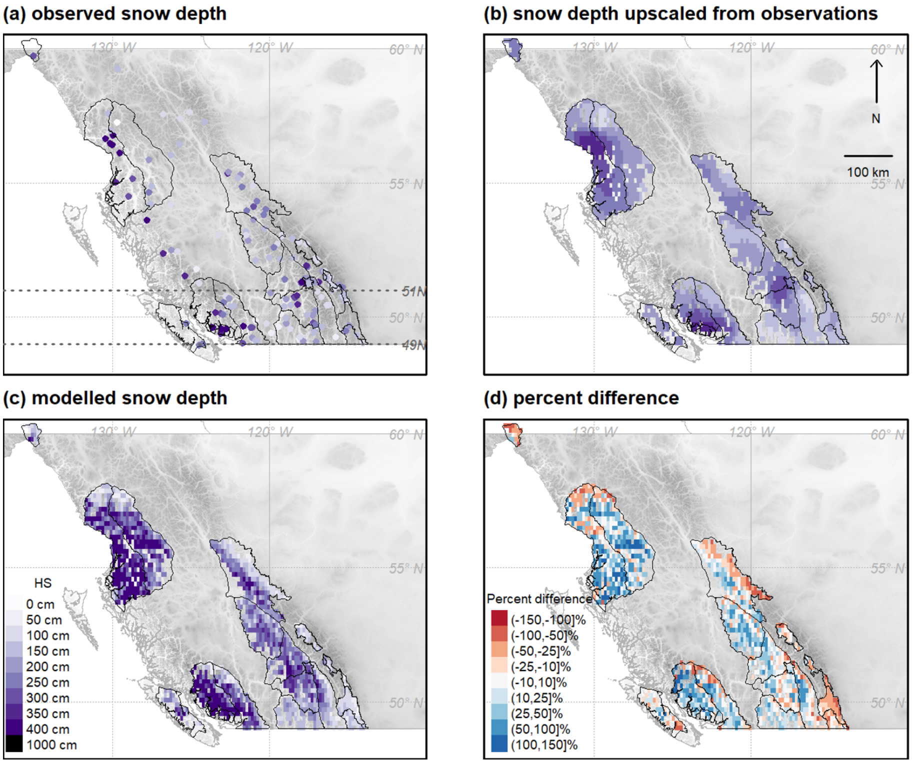

Figure 4. Comparison of observed and modelled snow depths on 31 March 2021 with (a) original snow depths from 176 observations, (b) the observed snow depths upscaled to treeline elevations at 1260 model grid points, (c) the modelled snow depth at the same grid points, and (d) the percent difference between the snow depths from model and upscaled observation data (positive values indicate modelled depths were greater than upscaled observations). Panel (a) shows the latitudinal transect between 49° and 51° N used in Fig. 5.

4.2 Comparing modelled and observed snow depths

4.2.1 Regional-scale spatial patterns

Snow depths on 31 March 2021 are mapped in Fig. 4 to illustrate regional-scale patterns by the end of the winter season. Observed snow depths were relatively deep ~~(i.e., greater than 300 cm)~~ along western parts of the Coast range as well as the central parts of the Columbias ~~.~~ (i.e., greater than 250 cm). Shallower snow depths were observed on the eastern side of the Coast range, the perimeters of the Columbias, and throughout most of the Rockies (Fig. 4a-b). Modelled snow depths had similar regional-scale patterns, but with more extreme differences between the deep and shallow snowpack areas within each region (Fig. 4c). For example, ~~observed snow depths ranged from 115 to 491~~ the snow depths from upscaled observations ranged from 114 to 399 cm across the entire study area, while modelled snow depths ranged from ~~39 to 917~~ 50 to 939 cm. Despite the different ranges, the median snow depths were similar with values of ~~257 and 285~~ 227 and 259 cm for the observed and modelled data, respectively.

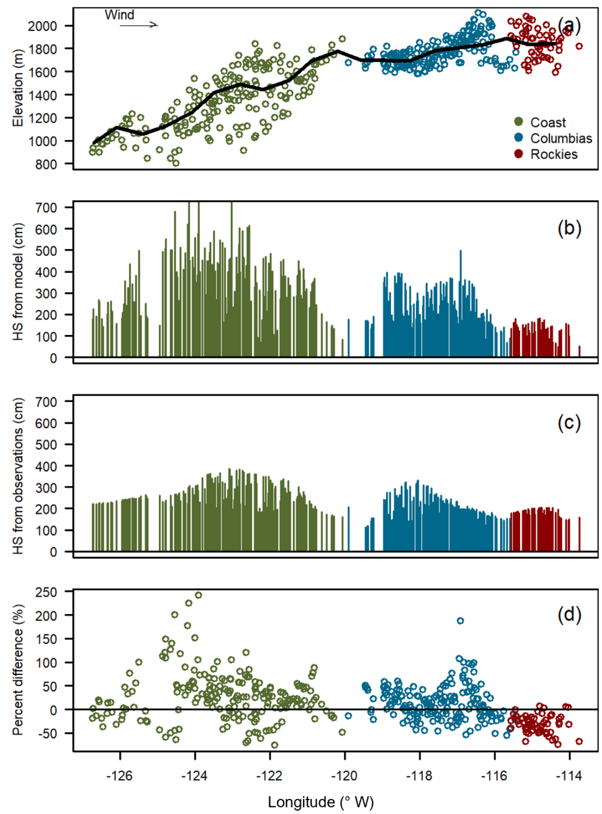

Figure 5. Latitudinal cross-section between 49° and 51° N of (a) the elevation of treeline grid points with the black line showing the moving average, (b) modelled snow depth at treeline grid points, (c) snow depths from upscaled observations at the same grid points, and (d) percent difference between model and upscaled observation snow depths.

~~Comparison of observed and modelled snow depths on 31 March 2021 with (a) original snow depths from 166 observations, (b) the observed snow depths interpolated to treeline elevations at 856 NWP model grid point locations, (c) the modelled snow depth at the same grid points, and (d) the percent difference between the modelled and observed snow depths (positive values indiciate modelled depths are greater than observed).~~

~~The percent difference between modelled and observed snow depths show where regional-scale discrepancies were most pronounced (Fig. 4d). Modelled snow depths were substantially greater than local observations in the Coast range, especially on the western (upslope) side of the range. With large snow depths in the Coast range, the root mean square errors were relatively large and exceeded 100 cm in many regions (Table 2). Snow depth biases were positive in all Coast regions, except for the Yukon. Correlations between modelled and observed snow depths within each Coast region ranged from 0.25 to 0.75, suggesting there was moderate to strong agreement on the location of the relatively deeper and shallower areas within each region.~~

Table 2. Statistics comparing ~~modelled and observed~~ snow depths from model and upscaled observation data within each forecast region on 31 March 2021. (Footnote: [a] Number of point observations ~~within~~ in forecast region ~~boundaries~~ prior to ~~interpolation. The interpolated~~ upscaling, however the upscaling includes observations ~~also consider observations~~ from neighbouring areas.)

Range	Region	Observations[a]	Number of grid points	Root mean square error (cm)	Bias (cm)	Correlation
Coast	Yukon	1	~~6~~ 18	~~113~~ 125	~~-74~~ -91	~~0.75~~ 0.79
	Northwest Coastal	~~18~~ 10	~~149~~ 211	~~205~~ 184	~~127~~ 96	0.40
	Northwest Inland	6	193	123	75	0.61
	Vancouver Island	~~2~~ 4	~~17~~ 34	~~101~~ 100	~~9~~ -18	~~0.25~~ -0.26
	South Coast	5	~~6~~ 12	~~66~~ 116	~~2~~ 80	~~0.66~~ 0.16
	Sea To Sky	~~6~~ 5	~~62~~ 86	~~194~~ 190	~~131~~ 126	~~0.33~~ 0.41
	South Coast Inland	~~10~~ 12	~~62~~ 92	~~88~~ 95	~~-1~~ 33	~~0.59~~ 0.72
Columbias	Cariboos	~~10~~ 11	~~72~~ 109	79	~~21~~ 51	~~0.08~~ 0.38
	North Columbia	~~8~~ 5	~~52~~ 70	~~74~~ 75	~~-15~~ 26	~~0.59~~ 0.45
	Glacier	7	~~3~~ 4	~~19~~ 52	~~-4~~ 2	~~0.50~~ 0.20
	South Columbia	~~18~~ 13	~~50~~ 71	~~81~~ 69	~~-10~~ 11	~~0.46~~ 0.44
	Kootenay Boundary	~~13~~ 14	~~43~~ 55	~~70~~ 44	~~-51~~ 22	~~0.54~~ 0.24
	Purcells	~~13~~ 11	~~45~~ 67	~~97~~ 89	~~30~~ 27	0.45
Rockies	North Rockies	~~14~~ 13	~~104~~ 152	~~91~~ 82	~~-32~~ -6	~~0.42~~ 0.46
	Jasper	~~9~~ 7	~~13~~ 19	~~56~~ 43	~~-14~~ -28	~~0.45~~ 0.49
	Banff Yoho Kootenay	11	~~3~~ 6	~~29~~ 52	~~-9~~ -19	~~0.50~~ 0.26
	Little Yoho	0	2	~~48~~ 26	~~46~~ 26	-1.00
	Kananaskis	5	~~2~~ 3	~~147~~ 99	~~-144~~ -96	~~-1.00~~ -0.50
	Lizard Flathead	~~10~~ 5	~~10~~ 19	~~33~~ 46	~~-23~~ -34	-0.15
	South Rockies	~~23~~ 2	~~23~~ 34	~~66~~ 76	~~-53~~ -68	~~-0.55~~ -0.21
	Waterton Lakes	~~1~~ 5	~~1~~ 3	~~95~~ 69	~~-95~~ -51	~~NA~~ -1.00

The percent difference between snow depths from model and upscaled observation data show where regional-scale discrepancies were most pronounced (Fig. 4d). Modelled snow depths were substantially greater than observations in most parts of the Coast range, especially on the western (upslope) side of the range. With large snow depths in the Coast range, the root mean square errors were relatively large and exceeded 100 cm in many regions (Table 2) and the average snow depth biases were positive, except for the Yukon and Vancouver Island regions. Positive snow depth correlations for most regions in the Coast range suggest there was moderate agreement on the location of the relatively deeper and shallower areas at a sub-regional scale, except for the Vancouver Island and South Coast regions, where weak or negative correlations suggest the model and observation data disagreed on the location of deeper areas.

In the Columbias, ~~modelled and observed snow depths~~ snow depths from the model and upscaled observation data were relatively similar (Fig. 4d), although the modelled depths were substantially ~~lower in the southwestern parts of the range and substantially higher~~ higher than observations in the western Purcell and central Cariboo regions. The root mean square error ~~in~~ for regions in the Columbias ranged from ~~19 to 97~~ 44 to 89 cm (Table 2), with relatively smaller biases than the other ranges. Positive correlations between ~~0.45 and 0.59~~ 0.20 and 0.45 suggest the model and observation data had moderate agreement in where the relatively deeper and shallower snowpacks existed ~~, except for the Cariboo region where the correlation was only 0.08~~throughout the range.

In the Rockies, there was reasonable agreement between snow depths in the central regions of Banff Yoho Kootenay and Jasper, but modelled depths were substantially lower than ~~observed~~ upscaled observations in the southern and northern parts of the range (Fig. 4d). Negative biases existed in all regions except Little Yoho (Table 2), highlighting how modelled depths were systematically ~~less than observed~~ lower than observed depths in the Rockies. While the regions in the southern Rockies had fewer grid points to compute correlations, their negative values mean the model and observation data disagreed on where the relatively deeper and shallower snowpacks existed.

To further illustrate the predominant spatial patterns across the three major mountain ranges, snow depths are plotted along a latitudinal cross-section between ~~49 and 51 10001000~~49° and 5110001000 N ~~in~~ (Fig. 5). The western (windward) side of the Coast range had substantially larger modelled snow depths than ~~observed~~upscaled observations, while there was reasonable agreement along the highest terrain and the leeward side of the Coast range. In the Columbias, the windward side had the greatest observed snow depths, which were slightly larger than modelled, however, modelled depths were greater than observed over the highest terrain in the eastern side of the Columbias. In the Rockies, there was minimal variation between the windward and leeward sides for both the modelled and observed snow depths.

4.2.2 Temporal patterns

4.2.3 ~~Temporal patterns~~

Timeseries of snow depths show the agreement between ~~modelled and observed~~ model and observation data over the course of the 2020-21 winter (Fig. 6). Some regions such as the ~~Cariboos~~South Coast Inland, Glacier, and Banff Yoho Kootenay had relatively consistent agreement over the season, while other regions ~~, such as Sea to Sky and Waterton Lakes,~~ had major discrepancies. Discrepancies arose from both systematic biases that persisted for the duration of the season and from differences originating from specific events. For example, systematic biases are evident in the Sea to Sky region where modelled depths were consistently greater than ~~observed~~ upscaled observations and in the Kananaskis region where modelled depths were consistently ~~less than observed~~lower. Examples of specific events causing discrepancies include Vancouver Island where the observed depths increased in early January without a corresponding increase in modelled depths, and in the ~~Northwest Coastal~~ Yukon region where changes in modelled and observed depths occurred at different times throughout the season. ~~The interquartile ranges of modelled snow depths were wider than the observations, especially in regions with fewer observations.~~

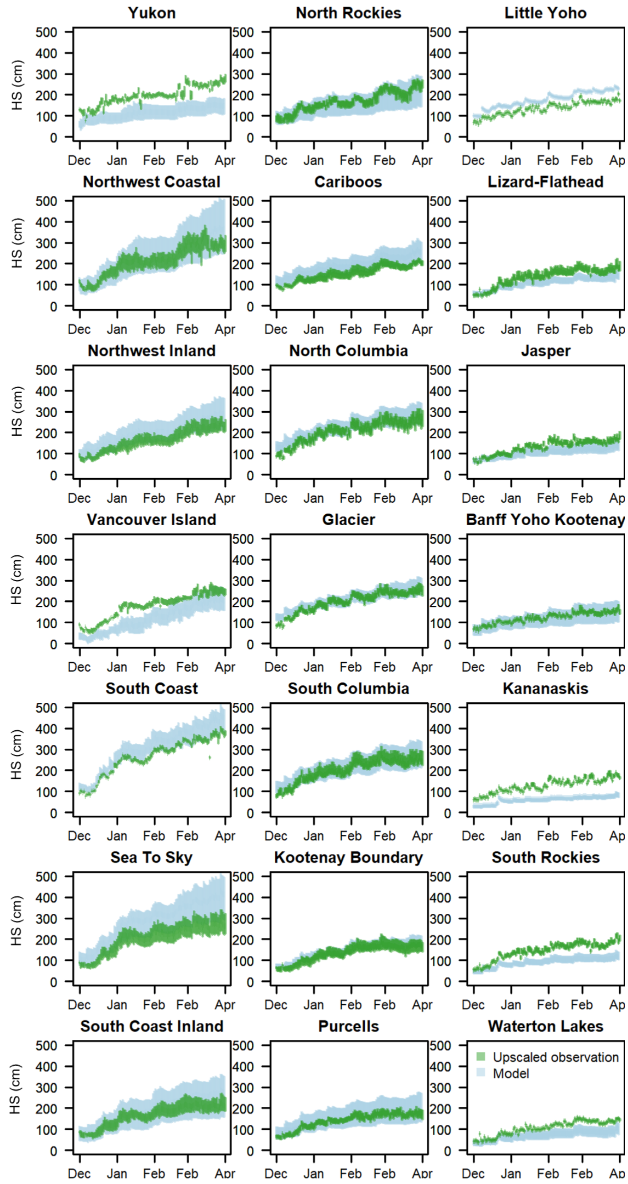

Figure 6. ~~Latitudinal cross-section between 49 and 51 °N~~ Interquartile range of ~~(a) the elevation of NWP grid points, (b) modelled~~ snow ~~depth at treeline grid points, (c) observed snow~~ depths ~~interpolated to~~ for each avalanche forecast region throughout the ~~same grid points, and~~ study period from model (~~d~~blue) ~~percent difference between modelled~~ and ~~observed snow depths~~upscaled observation (green) data.

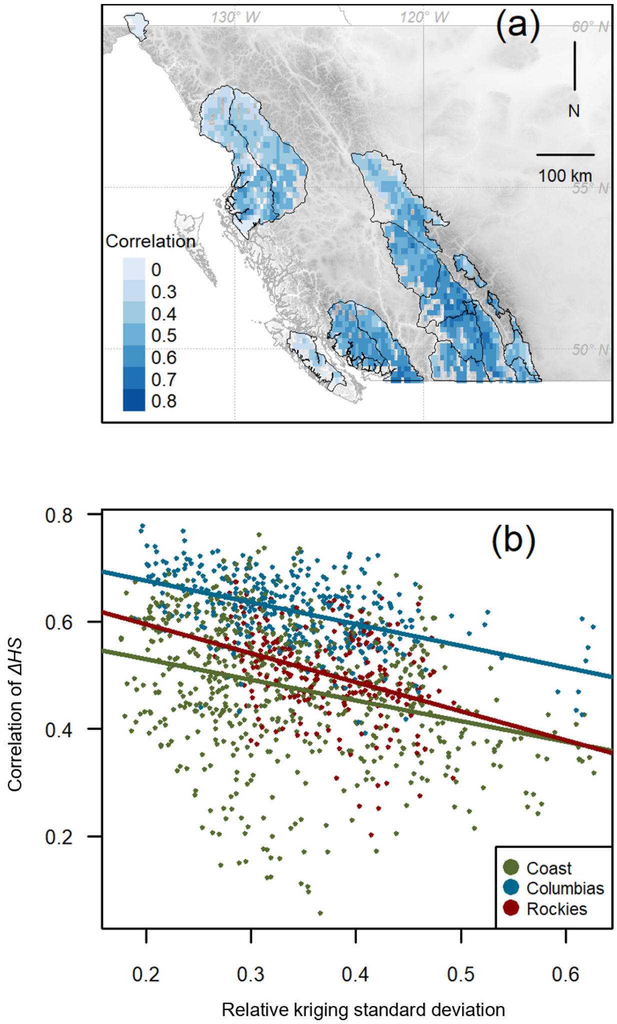

Figure 7. ~~Interquartile range of~~ Correlation between daily change in snow ~~depths for each avalanche forecast region throughout~~ depth ΔHS from model and upscaled observation data over the study period ~~from modelled~~ (~~blue~~a) mapped across the study area and ~~observed~~ (~~green~~b) ~~data~~plotted against the average relative kriging standard deviation representing the uncertainty in the upscaled observations. Simple regression lines are fitted for each mountain range to show poorer ΔHS correlations when upscaled observations had greater uncertainty.

Agreement in the timing of new snow was measured with the correlation between ~~modelled and observed height of new snow (HN) at each location (~~daily snow height changes ΔHS from the model and upscaled observation data (Fig. 7a). ~~Across all 856 grid points , the median correlation was 0.69 and ranged from 0.26 to 0.91~~The median correlation for all grid points was 0.54, but ranged from 0.06 to 0.78. Correlations were strongest in the southern Coast range, the central Columbias, and the central Rockies and relatively weaker in the northern Coast range and southern Rockies. The correlation was generally ~~lower~~ weaker in areas with larger ~~RKSD~~ $RKSD$ (Fig. 7b), suggesting disagreements were not only influenced by model errors but also by uncertainty in the upscaled observations.

~~Correlation between modelled and observed height of new snow (HN) over the study period (a) mapped across the study area and (b) plotted against the uncertainty in the observations from the average relative kriging standard deviation (RKSD). Simple regression lines are fitted for each mountain range to show poorer HN correlations when observations had greater uncertainty.~~

4.3 Impact on modelled snowpack structure

4.3.1 Regional-scale snowpack patterns with NWP forcings

Snow profile simulations from each forecast region illustrate typical climatic patterns across western Canada (Fig. 8). Grain types are colour-coded following the suggestions of Horton et al. (2020a) rather than Fierz et al. (2009) to better highlight features relevant to avalanche conditions (i.e., new snow and weak layers). The maritime climate in the Coast range resulted in thick layers of new snow, rounded grains, and melt forms. Profiles in the Columbias had typical transitional snowpacks with thick layers of rounded grains interspersed with thin layers of persistent grain types. The continental Rockies had the thinnest snowpacks and were largely composed of faceted grain types. ~~Grain types in Fig. 8 are colour-coded following the suggestions of Horton et al. (2020a) to highlight features relevant to avalanche conditions.~~

~~Simulated snowpack structure at a representative location in each forecast region. The snow depth interpolated from available observations at the same locations are shown with black squares to show the agreement between modelled and observed snow depths. Snowpack layers are colour-coded with surface hoar in red (SH), depth hoar in dark blue (DH), precipitation particles in bright yellow (PP), decomposing and fragmented precipitation particles in light yellow (DF), rounded grains in pink (RG), rounding faceted particles in mauve (FCxr), facet crystals in light blue (FC), melt-freeze crusts in bright green (MFcr), melt forms in light green (MF), and ice formations in blue-green (IF).~~

~~Plotting the snow depths interpolated from available observations at each profile location in Fig. 8 shows the potential accuracy of the simulations~~Overlaying the snow depth from upscaled observations on each simulated profile shows how well the snow depths agreed over time. The modelled stratigraphy was likely simulated better at locations where the ~~observed and modelled snow depths closely agreed~~ modelled snow depth closely agreed with upscaled observations (e.g., South ~~Coast Inland, Glacier, and Banff Yoho Kootenay regions). However, many locations~~ Columbia), although matching snow depths is not sufficient to guarantee accurate stratigraphy. Many simulated profiles had large discrepancies ~~between modelled and observed snow depths~~with the upscaled observations, in which case it is less clear whether the model was erroneous or the upscaled observations were not representative for that location.

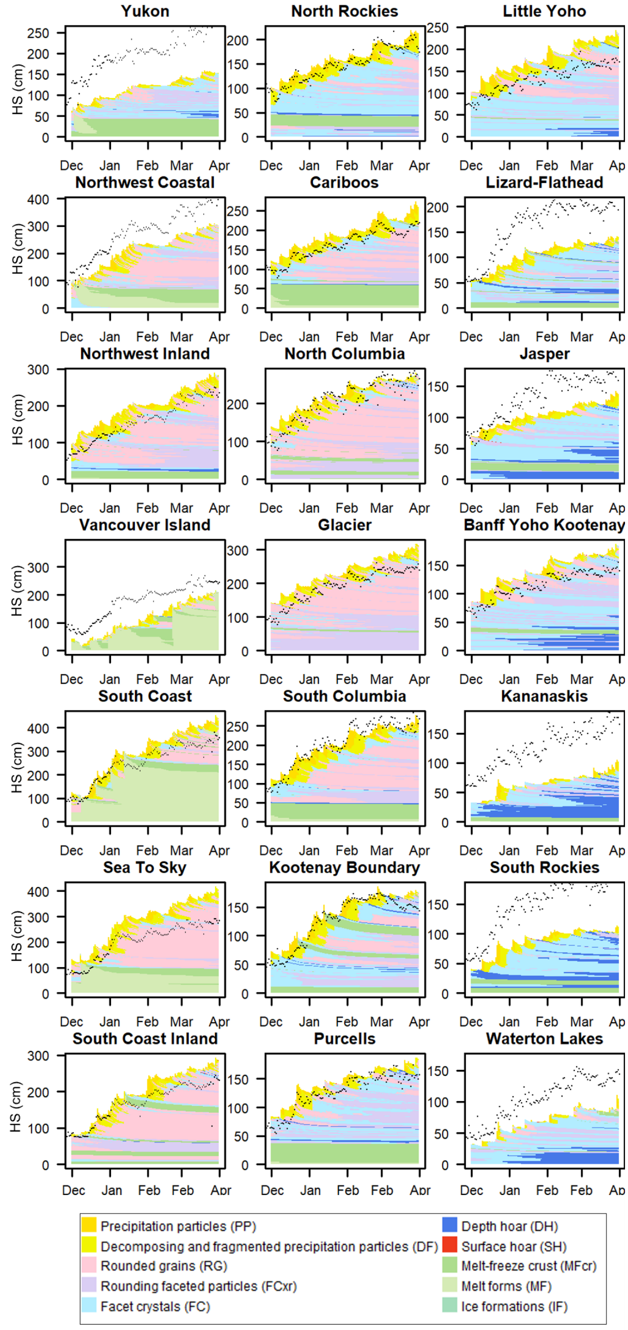

Figure 8. Simulated snowpack structure at a representative location in each forecast region with the snow depths from upscaled observations shown with black squares. Snowpack layers are colour-coded by grain type following Horton et al. (2020a).

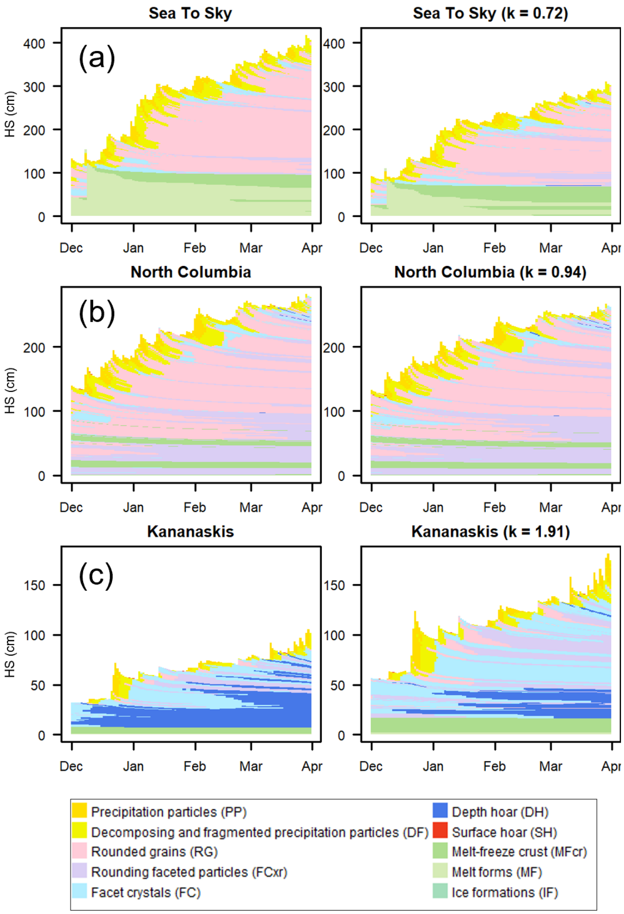

Figure 9. Timelines of simulated snowpack structure before and after precipitation adjustments at locations in (a) in the Sea to Sky region where modelled snow depths were much larger than upscaled observations, (b) the North Columbia region where the modelled and upscaled observation snow depths closely agreed, and (c) in the Kananaskis region where modelled snow depths were much smaller than upscaled observations. Snowpack layers are colour-coded by grain type following Horton et al. (2020a).

~~The bias-correction factors~~

4.3.2 Impact of adjusting precipitation inputs

The precipitation-adjustment factors (k) calculated by comparing the modelled ~~and observed depths at these locations with Eq. ?? ranged from 0.71~~ snow depths to upscaled observations at the locations in Fig. 8 ranged from 0.72 for the Sea to Sky location (where observed depths were much smaller than modelled ~~to 2.43~~ depths) to 1.91 for the Kananaskis location (where observed depths were much greater~~. The interquartile range of correction factors for these locations ranged from 0.84 to 1.22. The interpolated snow depths at a single location were not consistent over time because the availability of observations varied~~

~~each day, hence averaging the bias-correction over the entire study period was deemed more appropriate than applying bias corrections at finer time scales.~~

4.3.3 ~~Impact of bias-corrected precipitation inputs~~

~~Applying bias-corrections to the precipitation inputs~~). Adjusting precipitation inputs by these factors resulted in a variety of impacts on the simulated snowpack structure. Profiles are compared for ~~three specific cases in~~ one location in each mountain range in Fig. 9, including an example with a large reduction of precipitation for the Sea to Sky profile in the Coast range (k = ~~0.71~~0.72), a small reduction in precipitation for the ~~Glacier profille~~ North Columbia profile in the Columbias (k = ~~0.97~~0.94), and a large increase of precipitation for the Kananaskis profile in the Rockies (k = ~~2.43~~1.91). Profiles for the remaining regions are shown in Appendix A.

~~Timelines of simulated snowpack structure before and after bias-corrected precipitation inputs at locations in (a) in the Sea to Sky region where modelled snow depths were much larger than observed, (b) the Glacier region where the modelled and observed snow depths closely agreed, and (c) in the Kananaskis region where modelled snow depths were much smaller than observed. Snowpack layers are colour-coded according to grain type (see Fig. 8).~~

~~The bias-corrected~~ The adjusted Sea to Sky profile contained similar snowpack features to the original profile (Fig. 9a), despite a dramatic reduction in the amount of precipitation. Similar layers of rounded grains and the same prominent weak layers existed in both profiles. The average similarity value of the two sets of profiles was 0.79, suggesting the profiles were relatively similar in terms of features important to avalanche hazard. The similarly measure ~~did not have~~ had a statistically significant ~~change~~ decrease over the course of the season, meaning differences ~~did not compound~~ compounded over time. The ~~bias-corrected profile was effectively a thinner version of the original profile.~~

~~The bias-corrected Glacier profile~~ adjusted North Columbia profile had minimal differences from the original ~~profiles~~ profile after adjusting the precipitation by a factor of ~~0.97~~ 0.94 (Fig. 9b), with an average similarity value of ~~0.97~~ 0.91 that did not change significantly over time.

~~The bias-corrected~~ The adjusted Kananaskis profile, however, had a dramatically different snowpack structure after increasing precipitation by a factor of ~~2.43~~ 1.91 (Fig. 9c). The original profile was primarily composed of a thick depth hoar layer and small amounts of new snow above it, which after adjustment, was stretched into thicker layers of facets, rounding facets, and a few thin well-defined depth hoar layers. The average similarity value of ~~0.64 suggested~~ 0.50 suggests the profiles were substantially different in terms of avalanche hazard considerations~~. Unlike the Sea to Sky and Glacier profiles, the similarity value had~~ , and the similarity values a statistically significant decrease over the course of the season~~, suggesting the differences compounded over time~~.

The similarity between the original and ~~bias-corrected~~ precipitation-adjusted profiles are shown for all 21 regions in Fig. 10. The similarity values ranged from ~~0.60 for the Waterton profiles to 1.00~~ 0.50 for the Kananaskis profiles to 0.94 for the South ~~Coast Inland~~ Columbia profiles, with a median value of ~~0.86. Small bias-corrections~~ 0.81. Small precipitation adjustments generally resulted in profiles that were similar to the original profiles from an avalanche hazard perspective, as was the case for the ~~Glacier profile. Large adjustments in precipitation in either direction (i.e., increase or decrease)~~ North Columbia

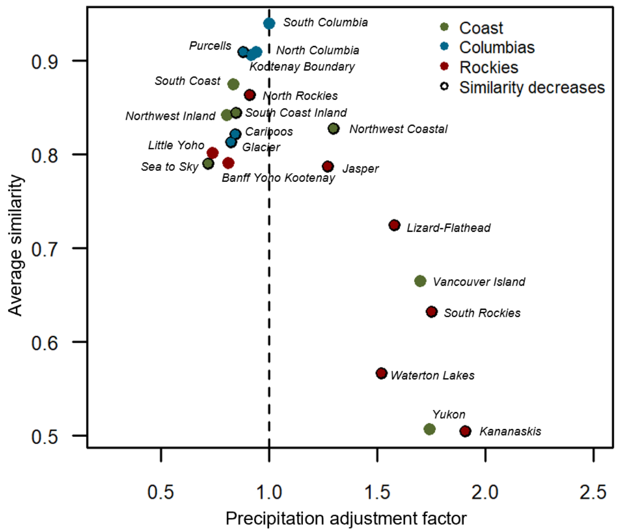

Figure 10. Average similarity in snowpack structure between original and ~~bias-corrected~~ precipitation-adjusted snow ~~profile simulations~~ profiles at locations in each forecast region. The similarly is plotted against the magnitude of the ~~bias correction~~ precipitation adjustment factor~~applied to~~. Locations where the ~~precipitation inputs~~similarity has a statistically significant decrease over time are outlined with black circles.

profile, while large changes in precipitation resulted in more substantial structural changes. The impacts of ~~bias-corrections~~ precipitation adjustments were most dramatic for profiles in the Rockies, as ~~the four~~ seven of the ten locations with the lowest similarity values ~~(all less than 0.75)were all in the Rockies. Profiles in the Coast and~~ were in the Rockies (all with values less than 0.8). All profiles in the Columbias had similarity values greater than 0.8~~(except for the Sea to Sky profile), suggesting bias-corrections~~, suggesting precipitation adjustments had relatively smaller impacts on simulated avalanche conditions. ~~When comparing two locations with relatively small bias-corrections, a 0.97 correction in Glacier resulted in a similarity of 0.95 compared to a similar correction of 0.98 in Banff Yoho Kootenay that resulted in a similarity of 0.92. In this case, the small adjustment in precipitation in Banff Yoho Kootenay resulted in more substantial changes to the snowpack structure. The similarity value~~ Profiles in the Coast range had a wide range of similarity values. The similarity values became statistically lower over the course of the season in ~~1~~ 3 of 7 Coast profiles, ~~4 of 7~~ 3 of 6 Columbias profiles, and 6 of 8 Rockies profiles, suggesting the impact of ~~bias-corrections become~~ precipitation adjustments became more compounded over time in colder continental climates.

5 Discussion

5.1 Implications for weather forecasting

Lundquist et al. (2020) argue the increasing skill in modelling mountain precipitation relative to traditional ~~observations~~ observation networks warrants new multi-disciplinary methods for evaluating models. Avalanche safety operations spend considerable effort collecting weather observations at high elevations, which could be a valuable addition to ~~verifying weather forecasts . Avalanche weather observations are rarely included in meteorological data assimilation or analysis products (Roy et al., 2018; ?) , as they rarely meet meteorological standards (e.g., World Meteorological Organization). One challenge is that most meteorological products measure precipitation in millimetres of water equivalent, but the coverage and quality of precipitation gauges in mountainous terrain is relatively much poorer than snow depth observations~~verify weather forecasts and improve analysis products which primarily rely on low elevation precipitation measurements (e.g., Roy et al., 2018). Our analysis of ~~available datasets reveals notable patterns in the performance of the HRPDS weather model in~~ upscaled snow depth observations revealed patterns in how the HRDPS weather model performed in remote regions of western Canada, ~~including the~~ such as the systematic overprediction of precipitation on the windward side of the Coast range (Fig. 5). This specific bias likely originated from the NWP model's handling of hydrometeor drift and spillover in complex terrain (Mo et al., 2019), which ~~is a difficult to calibrate parameterization. Additional~~ could be better calibrated with avalanche weather observations. An additional analyses (not shown) using snow water equivalent and precipitation measurements from the same observation networks ~~did not provide as clearresults. Therefore, it would be worthwhile to incorporate more snow observations into weather products, as done with the SNODAS product produced in the contiguous United States (?)~~were not as clear, likely because fewer sites measure these variables.

This study highlights several shortcomings of resolving regional-scale conditions with the mountain observation networks ~~across western Canada, as suggested by Lundquist et al. (2020)~~in western Canada. For example, many ~~regions~~ areas had sparse and intermittent observations that ~~result~~ resulted in large uncertainties when ~~interpolating snow depths~~ upscaling snow depths to regional scales (Fig. 3~~a). Model verification metrics often performed poorer in areas with greater observation uncertainty (Fig. 7), suggesting NWP models may not actually perform as poorly as verification metrics suggest and that observations should not be treated as absolute ground truth.~~). Evaluating model performance in these areas should be done with caution, as the classic interpretations of the verification metrics may be less meaningful (e.g., Table 2).

The regression models ~~highlighted~~ highlight that in many regions ~~the expert assessed avalanche conditions were explained better with modelled snow depth changes than with observed snow depth changes~~ snow depth changes from the model data could explain regional avalanche conditions with comparable skill to the upscaled observations (Table 1)~~. In particular~~, with the main exceptions being small data-rich regions like Glacier. However, avalanche conditions in ~~the~~ some of the large heavy snowfall regions in the Coast range were better explained by modelled ~~snowfall than observations (Fig. 4b-c). While this could be partly explained by the fact avalanche forecasters put~~ snow depth changes than by upscaled observations. This could be caused by forecasters putting more weight on ~~NWP model forecasts in their assessments in~~ weather forecasts in these regions, ~~it~~ but could also be ~~explained~~ caused by shortcomings of the observation networks in these harsh coastal environments. ~~Except~~

~~for small data-rich regions like Glacier, regression models with modelled snowfall were relatively similar in performance to models with observed snowfall.~~ While these regression models were limited by human assessment errors and by over-simplifying the factors influencing avalanche hazard, they provide a simple example of how model ~~information is~~ weather data can be comparable to observation ~~information~~data in some regional-scale contexts.

5.2 Implications for snowpack modelling

Given precipitation has been shown to be a primary source of uncertainty in snowpack simulations (Raleigh et al., 2015; Richter et al., 2020), meaningful methods to ~~verify~~ identify and correct erroneous precipitation inputs could dramatically improve the quality of snowpack models. However, this study highlights large gaps and uncertainties in many observation networks that warrant careful approaches when evaluating snowpack models, especially at regional scales.

~~The comparisons of modelled and observed snow depths across western Canada presented in this study suggest a large-scale precipitation bias existed with too much precipitation in Coast range and too little precipitation in the Rockies (Fig. 4 and Fig. 5). Correcting these biases with a constant correction factor, as done in this study, could be an appropriate method in areas with high quality observations. For example, many locations in the southern Coast range had a strong correlation between modelled and observed HN (Fig. 7) as well as relatively high confidence in the observations (Fig. 3), suggesting precipitation errors were likely spread evenly across all precipitation events. This would be a good candidate for applying a constant bias-correction factor to snowpack simulations. However, a constant bias correction may be less appropriate in the northern Coastal range where observed HN had weaker correlations with modelled depths and the observations had greater uncertainty.~~

~~Bias corrections would only be useful when these biases persist in an NWP model over several seasons, but ideally NWP models continually improve their precipitation forecasts by increasing grid resolution and improving their data assimilation, physics, and parameterizations. A more robust method would be to correct precipitation errors at finer time scaleswith a data assimilation routine like the one presented by Winstral et al. (2018). This dynamic bias correction method would be beneficial in areas with sufficient observation networks to capture short term changes in snow depth, but as found in this study, this could be challenging in large regions where sparse and inconsistent observations do not provide a clear picture of short-term snow depth changes.~~

~~While detailed assimilation was outside the scope of this study, adjusting precipitation inputs by constant factors illustrated some key impacts of precipitation errors~~ Adjusting precipitation inputs illustrated how precipitation input errors could impact the simulated snowpack structure from an avalanche hazard perspective. The greatest change ~~in snowpack structure~~ was observed in the continental ~~Rocky Mountain range~~ Rockies regions (Fig. 10). In cold climates, changes in snow depths ~~had~~ would have a greater impact on temperature gradients in the snowpack, and as a result, the formation of weak faceted layers. Increasing precipitation could result in substantially fewer faceted layers and a less hazardous snowpack structure, while decreasing precipitation could result in substantially more faceted layers and a more dangerous structure. Locations in the Coast range, on the other hand, exhibited fewer differences in their snowpack structure after ~~bias corrections~~precipitation adjustments. Changing snow depths in maritime climates has less impact on the temperature gradients, and as a result the

~~bias-corrected snowpack~~ precipitation-adjusted profiles resembled a stretched or compressed version of the original profiles, usually containing the same weak layers and crusts.

The precipitation-adjustment method applied in this study is likely insufficient for an operational model system, as it assumes a constant bias over time and oversimplifies the causes of snow depth errors. More advanced data assimilation methods have recently been suggested for snowpack models (Largeron et al., 2020), including methods presented by Winstral et al. (2018) and Cluzet et al. (2022) that use similar snow observation networks in Europe. Cluzet et al. (2022) found assimilating snow depth observations improved simulations in areas of France with relatively sparse observations, but that the density of snow observations was correlated with the density of precipitation observations. For large regions in western Canada, sparse and inconsistent observations would pose a challenge to implementing data assimilation methods, especially considering snow depth changes over short time scales were poorly resolved with the upscaled observations (Fig. 7).

5.3 Implications for avalanche forecasting

This study found relatively strong agreement between modelled and observed snow depths in many situations across western Canada. In cases with strong agreement, the simulated snowpack structure can be interpreted by forecasters with a higher degree of confidence. In situations where the modelled and observed snow depths differ, the simulated snowpack structure must be evaluated more critically. Considerations should include the representativeness of local observations, whether the NWP model has a known precipitation bias in that region, and the sensitivity of the snowpack structure to ~~snow depth~~ precipitation errors in that climate.

Assessing the quality of the upscaled snow depth observations identified ~~regions~~ areas where forecasters deal with ~~data sparsity~~ sparse and uncertain observations (Fig. 3). Snowpack models could be particularly valuable in these regions, but to interpret the models it would help to collected targeted snow depth information from automated weather stations, field observers, crowd sourcing platforms (e.g., ~~Mountain Information Network,~~ Community Snow Observations), and satellite-derived snow cover products.

This study focused on regional-scale patterns to provide a general understanding of model performance; however, snowpack models also have potential to simulate finer scale spatial patterns. While such patterns are ~~highly~~ relevant to avalanche forecasters, they become even more difficult to verify with sparse observations. Therefore, we suggest snowpack models forced with NWP ~~model output~~ models should first be understood at coarse regional scales. For example, the correlations in Table 2 indicate how well the model and upscaled observation data agreed on the location of deep and shallow snowpack areas at a sub-regional scale. Sub-regional patterns resolved by snowpack models could be meaningful in regions where these correlations were strong, but further investigation would be necessary in regions with weak correlations to determine whether the patterns suggested by the model are reliable.

Exploring applications at finer scales could be considered in areas with high quality observations and strong agreement between modelled and observed snow depths. Aggregating observations across a common reference treeline elevation provided a consistent approach for regional-scale comparisons, however, more precise aggregations such as sub-grid parametrization of sky view factor (Helbig and van Herwijnen, 2017) would be required for verifying snow depth at smaller scales. The comparisons presented in this study were also point-wise, where modelled and ~~interpolated~~ upscaled observations were compared at the same location, but spatial verification techniques such as neighbourhood and feature-oriented approaches could also be meaningful to understand ~~regional-scale~~ patterns relevant to ~~avalanche~~ forecasters (Gilleland et al., 2009).

6 Conclusions

Forcing snowpack models with ~~output from~~ high-resolution NWP models is a promising method to support avalanche forecasting. However, in many situations there is limited observation data available for verification~~in many situations~~. Furthermore, the skill in modelling mountain precipitation continues to improve to the point where it ~~outperforms~~ rivals observation networks in many contexts. Aggregating snow depth observations from networks ~~such as~~ of automated weather stations ~~operated by government services~~ and manual observations from professional avalanche observers can provide insights into model performance at coarse spatial and temporal scales.

Applying this approach over a winter season across the diverse mountain climates of western Canada provided several insights about using snowpack models for avalanche forecasting. First, the quality and density of weather observations should be considered when verifying NWP and snowpack models. The number and consistency of observations in some areas resulted in relatively large uncertainties when ~~interpolating snow depth~~upscaling snow depths to regional scales, and in ~~roughly~~ almost half of the forecast regions, the ~~avalanche conditions assessed by expert forecasters~~ forecast avalanche conditions were explained with modelled snow depth changes better than with ~~observed~~ snow depth changes from upscaled observations. Despite limitations in the ~~observations~~observation data, the comparison of modelled and observed snow depths strongly suggested ~~a positive precipitation bias existed in the Coast range and a negative bias existed~~ an overprediction of precipitation in the maritime Coast Range and an underprediction in many parts of the continental Rockies. The agreement in snow depth was strongest in the transitional Columbia ~~Mountains~~Range, which also had the highest density of quality observations. The potential impacts of snow depth ~~errors~~ discrepancies were illustrated by comparing the snow profiles produced ~~with NWP models output~~ from the original NWP model data with profiles produced with ~~bias-corrected~~ adjusted precipitation inputs. Precipitation ~~bias-corrections~~ adjustments had the most dramatic effect in cold continental climates where the snow depths heavily influenced the degree of faceting. In warmer maritime climates, ~~bias corrections~~ adjusting precipitation inputs often resulted in stretched or compressed profiles with similar snowpack structure.

These results also highlight how meteorologists and NWP model developers could benefit from the observation networks of avalanche safety operations~~as well as their subjective knowledge of mountain weather~~. Snow depth observations could be a particularly valuable data stream ~~, and NWP models could benefit from additional research into methods for assimilating snow observations. Improving snowfall forecasts from NWP models would have direct benefits to avalanche forecasting and snowpack modelling.~~ to understand and evaluate NWP model performance in remote mountainous areas. From an avalanche forecasting perspective, this study highlights how limitations in observation networks pose a challenge to verifying snowpack models, and how these limitations need to be carefully considered when interpreting simulated profiles~~for avalanche forecasting~~.

This provides a starting point for future research into how operational snowpack ~~models~~ model systems could perform real-time verification and assimilation with available snow observations.

. Code and data are publicly available on the Open Science Framework at https://osf.io/a5pek (Horton and Haegeli, 2022).

Appendix A: Original and ~~bias-corrected~~ precipitation-adjusted profiles for all regions

. Both authors conceptualized the research with SH leading the analysis and writing and PH providing supervision.

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

. Thanks to Avalanche Canada, the Canadian Avalanche Association, and InfoEx subscribers for sharing datasets. Patrick Mair, Florian Herla, and Stan Nowak provided valuable input on data analysis and visualization methods. We also thank the reviewers for their constructive feedback.

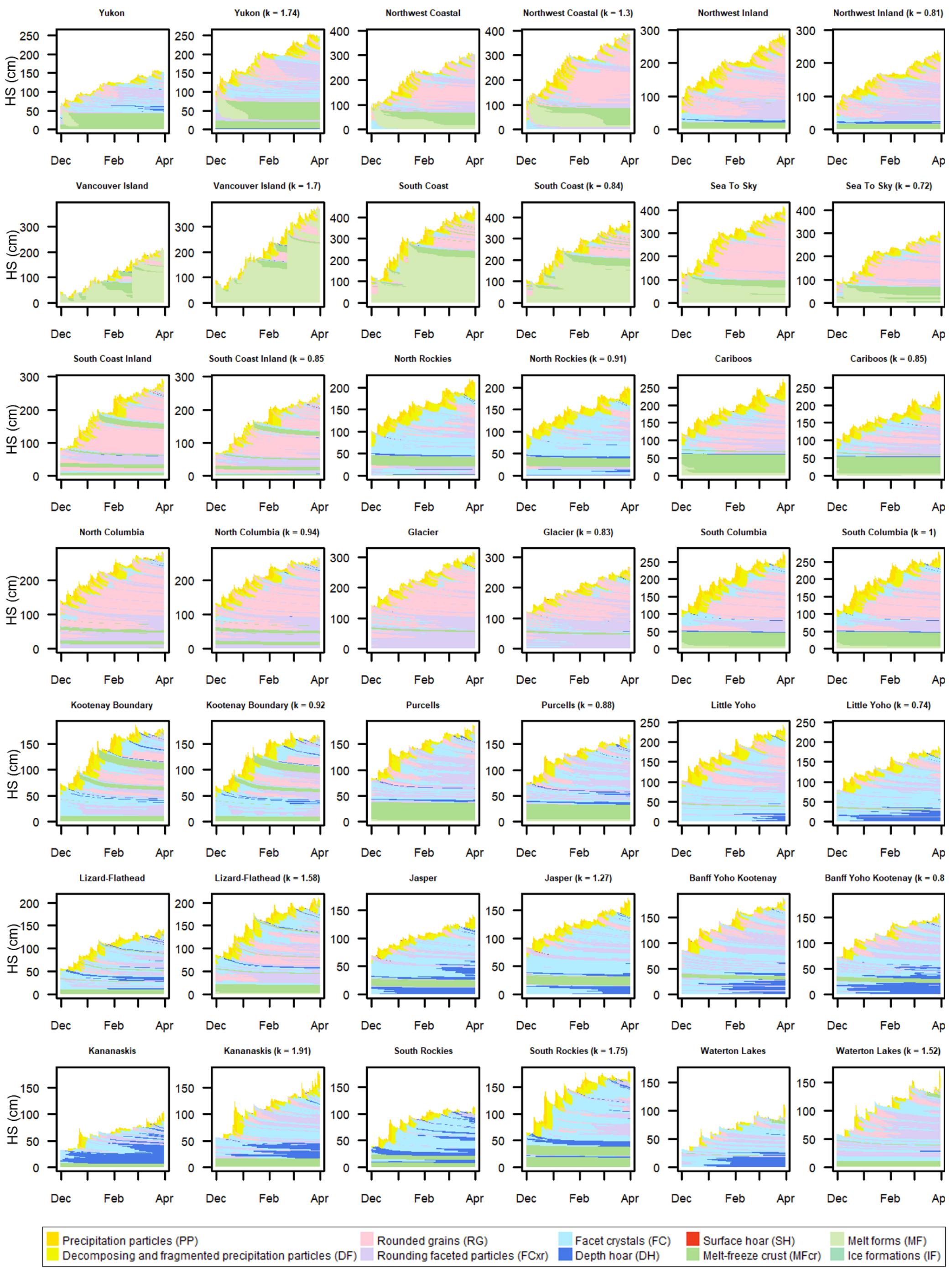

Figure A1. Timelines of simulated snowpack structure before and after precipitation ~~bias-correction~~adjustments. Two profile timelines are shown for each forecast region, with the original profiles to the left and the ~~bias-corrected~~ precipitation-adjusted profiles to the right~~with the bias correction factor~~. The precipitation adjustment factors (k) are shown in the title of the adjusted profiles. Snowpack layers are colour-coded ~~according to~~ by grain type ~~(see Fig~~following Horton et al. ~~8~~(2020a).

References

Bellaire, S. and Jamieson, B.: Forecasting the formation of critical snow layers using a coupled snow cover and weather model, Cold Reg. Sci. Technol., 94, 37–44, https://doi.org/10.1016/j.coldregions.2013.06.007, 2013.

Bellaire, S., Jamieson, B., and Fierz, C.: Forcing the snow-cover model SNOWPACK with forecasted weather data, Cryosphere, 5, 1115–1125, https://doi.org/10.5194/tc-5-1115-2011, 2011.

Bellaire, S., van Herwijnen, A., Mitterer, C., and Schweizer, J.: On forecasting wet-snow avalanche activity using simulated snow cover data, Cold Reg. Sci. Technol., 144, 28–38, https://doi.org/10.1016/j.coldregions.2017.09.013, 2017.

Brun, E., David, P., Sudul, M., and Brunot, G.: A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting, J. Glaciol., 38, 13–22, https://doi.org/10.3189/S0022143000009552, 1992.

Canadian Avalanche Association: Observation Guidelines and Recording Standards for Weather, Snowpack, and Avalanches, Canadian Avalanche Association, Revelstoke, BC, cdn.ymaws.com/www.avalancheassociation.ca/resource/resmgr/standards_docs/OGRS2016web.pdf, 2016.

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

Christensen, R.: ordinal - Regression Models for Ordinal Data, r package version 2019.12-10. https://CRAN.R-project.org/package=ordinal, 2019.

Cluzet, B., Revuelto, J., Lafaysse, M., Tuzet, F., Cosme, E., Picard, G., Arnaud, L., and Dumont, M.: Towards the assimilation of satellite reflectance into semi-distributed ensemble snowpack simulations, Cold Reg. Sci. Technol., 170, 102 918, https://doi.org/10.1016/j.coldregions.2019.102918, 2020.

Cluzet, B., Lafaysse, M., Deschamps-Berger, C., Vernay, M., and Dumont, M.: Propagating information from snow observations with CrocO ensemble data assimilation system: a 10-years case study over a snow depth observation network, Cryosphere, 16, 1281–1298, https://doi.org/10.5194/tc-16-1281-2022, 2022.

Danielson, J. and Gesch, D.: Global multi-resolution terrain elevation data 2010 (GMTED2010), 2011-1073, US Department of the Interior, US Geological Survey Washington, DC, USA, https://doi.org/10.3133/ofr20111073, 2011.

Durand, Y., Giraud, G., Brun, E., Mérindol, L., and Martin, E.: A computer-based system simulating snowpack structures as a tool for regional avalanche forecasting, J. Glaciol., 45, 469–484, https://doi.org/10.3189/S0022143000001337, 1999.

Durand, Y., Laternser, M., Giraud, G., Etchevers, P., Lesaffre, B., and Mérindol, L.: Reanalysis of 44 yr of climate in the French Alps (1958–2002): methodology, model validation, climatology, and trends for air temperature and precipitation, J. Appl. Meteorol. Climatol., 48, 429–449, https://doi.org/10.1175/2008JAMC1808.1, 2009.

Fierz, C., Armstrong, R., Durand, Y., Etchevers, P., Greene, E., McClung, D., Nishimura, K., Satyawali, P., and Sokratov, S.: The International Classification for Seasonal Snow on the Ground, IHP-VII Tech. Doc. Hyrdol. No. 83, IACS Contrib. No. 1, UNESCO-IHP, Paris, 2009.

Gilleland, E., Ahijevych, D., Brown, B., Casati, B., and Ebert, E.: Intercomparison of spatial forecast verification methods, Weather Forecasting, 24, 1416–1430, https://doi.org/10.1175/2009WAF2222269.1, 2009.

Giraud, G., Lafeuille, J., and Pahaut, E.: Evaluation de la qualité de la prévision du risque d'avalanche, Int. Ass. Hydrol. Sci. Publ, 162, 583–591, 1987.

Helbig, N. and van Herwijnen, A.: Subgrid parameterization for snow depth over mountainous terrain from flat field snow depth, Water Resour. Res., 53, 1444–1456, https://doi.org/10.1002/2016WR019872, 2017.

Helfricht, K., Hartl, L., Koch, R., Marty, C., and M., O.: Obtaining sub-daily new snow density from automated measurements in high mountain regions, Hydrol. Earth Syst. Sci., 22, 2655–2668, https://doi.org/10.5194/hess-22-2655-2018, 2018.

Herla, F., Horton, S., Mair, P., and Haegeli, P.: Snow profile alignment and similarity assessment for aggregating, clustering, and evaluating snowpack model output for avalanche forecasting, Geosci. Model Dev., 14, 239–258, https://doi.org/10.5194/gmd-14-239-2021, 2021.

Horton, S. and Haegeli, P.: Using snow depth observation to provide insight into the quality of regional-scale snowpack simulations for avalanche forecasting (Code and data), https://osf.io/a5pek, 2022.

Horton, S., Nowak, S., and Haegeli, P.: Enhancing the operational value of snowpack models with visualization design principles, Nat. Hazards Earth Syst. Sci., 20, 1557–1572, https://doi.org/10.5194/nhess-20-1557-2020, 2020a.

Horton, S., Towell, M., and Haegeli, P.: Examining the operational use of avalanche problems with decision trees and model-generated weather and snowpack variables, Nat. Hazards Earth Syst. Sci., 20, 3551–3576, https://doi.org/10.5194/nhess-20-3551-2020, 2020b.

Lafaysse, M., Morin, S., Coléou, C., Vernay, M., Serça, D., Besson, F., Willemet, J.-M., Giraud, G., and Durand, Y.: Towards a new chain of models for avalanche hazard forecasting in French mountain ranges, including low altitude mountains, in: Proc. Int. Snow Sci. Workshop, Grenoble, France, 7-11 Oct., pp. 162–166, 2013.

Largeron, C., Dumont, M., Morin, S., Boone, A., Lafaysse, M., Metref, S., Cosme, E., Jonas, T., Winstral, A., and Margulis, S. A.: Toward Snow Cover Estimation in Mountainous Areas Using Modern Data Assimilation Methods: A Review, Front. Earth Sci., 8, https://doi.org/10.3389/feart.2020.00325, 2020.

Lehning, M., Bartelt, P., Brown, B., Russi, T., Stöckli, U., and Zimmerli, M.: Snowpack model calculations for avalanche warning based upon a new network of weather and snow stations, Cold Reg. Sci. Technol., 30, 145–157, https://doi.org/10.1016/S0165-232X(02)00073-3, 1999.

Lehning, M., Bartelt, P., Brown, B., Fierz, C., and Satyawali, P.: A physical SNOWPACK model for the Swiss avalanche warning: Part II. Snow microstructure, Cold Reg. Sci. Technol., 35, 147–167, 2002.

Liston, G. and Elder, K.: A meteorological distribution system for high-resolution terrestrial modeling (MicroMet), J. Hydrometeor., 7, 217–234, https://doi.org/10.1175/JHM486.1, 2006.

Lundquist, J., Hughes, M., Gutmann, E., and Kapnick, S.: Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks, Bull. Am. Meteorol. Soc., 100, 2473–2490, https://doi.org/10.1175/bams-d-19-0001.1, 2020.

Milbrandt, J., Bélair, S., Faucher, M., Vallée, M., Carrera, M. L., and Glazer, A.: The Pan-Canadian High Resolution (2.5 km) Deterministic Prediction System, Weather Forecasting, 31, 1791–1816, https://doi.org/10.1175/waf-d-16-0035.1, 2016.

Mo, R., Brugman, M., Milbrandt, J., Goosen, J., Geng, Q., Emond, C., Bau, J., and Erfani, A.: Impacts of hydrometeor drift on orographic precipitation: Two case studies of landfalling atmospheric rivers in British Columbia, Canada, Weather Forecasting, 34, 1211–1237, https://doi.org/10.1175/WAF-D-18-0176.1, 2019.

Morin, S., Horton, S., Techel, F., Bavay, M., Coléou, C., Fierz, C., Gobiet, A., Hagenmuller, P., Lafaysse, M., Ližar, M., Mitterer, C., Monti, F., Müller, K., Olefs, M., Snook, J., van Herwijnen, A., and Vionnet, V.: Application of physical snowpack models in support of operational avalanche hazard forecasting: A status report on current implementations and prospects for the future, Cold Reg. Sci. Technol., 170, 102 910, https://doi.org/10.1016/j.coldregions.2019.102910, 2020.

North American Land Change Monitoring System: Land Cover 30m, 2015 (Landsat and RapidEye), http://www.cec.org/north-american-environmental-atlas/land-cover-30m-2015-landsat-and-rapideye, accessed: 2021-08-15, 2015.

Pebesma, E.: Multivariable geostatistics in S: the gstat package, Comput. Geosci., 30, 683–691, https://doi.org/10.1016/j.cageo.2004.03.012, 2004.

Quéno, L., Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Dumont, M., and Karbou, F.: Snowpack modelling in the Pyrenees driven by kilometric-resolution meteorological forecasts, Cryosphere, 10, 1571–1589, https://doi.org/10.5194/tc-10-1571-2016, 2016.

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

Reuter, B., Schweizer, J., and van Herwijnen, A.: A process-based approach to estimate point snow instability, Cryosphere, 9, 837–847, https://doi.org/10.5194/tc-9-837-2015, 2015.

Richter, B., van Herwijnen, A., Rotach, M., and Schweizer, J.: Sensitivity of modeled snow stability data to meteorological input uncertainty, Nat. Hazard Earth Syst. Sci., 2020, 1–25, https://doi.org/10.5194/nhess-2019-433, 2020.

Roy, G., Stadnyk, T., Koenig, K., Gasset, N., and Mahidjiba, A.: Ten Years of Science Based on the Canadian Precipitation Analysis: A CaPA System Overview and Literature Review AU - Fortin, V, Atmos. Ocean, 56, 178–196, https://doi.org/10.1080/07055900.2018.1474728, 2018.

Schirmer, M. and Jamieson, B.: Verification of analysed and forecasted winter precipitation in complex terrain, Cryosphere, 9, 587–601, https://doi.org/10.5194/tc-9-587-2015, 2015.

Schirmer, M., Lehning, M., and Schweizer, J.: Statistical forecasting of regional avalanche danger using simulated snow-cover data, J. Glaciol., 55, 761–768, https://doi.org/10.3189/002214309790152429, 2009.

Shandro, B. and Haegeli, P.: Characterizing the nature and variability of avalanche hazard in western Canada, Nat. Hazards Earth Syst. Sci., 18, 1141–1158, https://doi.org/10.5194/nhess-18-1141-2018, 2018.

Statham, G., Haegeli, P., Greene, E., Birkeland, K., Israelson, C., Tremper, B., Stethem, C., McMahon, B., White, B., and Kelly, J.: A conceptual model of avalanche hazard, Nat. Hazards, 90, 663–691, https://doi.org/10.1007/s11069-017-3070-5, 2018.

Thornton, P., Running, S., and White, M.: Generating surfaces of daily meteorological variables over large regions of complex terrain, J. Hydrol., 190, 214–251, https://doi.org/10.1016/S0022-1694(96)03128-9, 1997.

Viallon-Galinier, L., Hagenmuller, P., and Lafaysse, M.: Forcing and evaluating detailed snow cover models with stratigraphy observations, Cold Reg. Sci. Technol., 180, 103 163, https://doi.org/10.1016/j.coldregions.2020.103163, 2020.

Vionnet, V., Six, D., Auger, L., Dumont, M., Lafaysse, M., Quéno, L., Réveillet, M., Dombrowski-Etchevers, I., Thibert, R., and Vincent, C.: Sub-kilometer precipitation datasets for snowpack and glacier modeling in alpine terrain, Front. Earth Sci., 7, https://doi.org/10.3389/feart.2019.00182, 2019.

Wang, X. L., Xu, H., Qian, B., Feng, Y., and Mekis, E.: Adjusted daily rainfall and snowfall data for Canada, Atmos. Ocean, 55, 155–168, https://doi.org/10.1080/07055900.2017.1342163, 2017.

Winstral, A., Magnusson, J., Schirmer, M., and Jonas, T.: The Bias Detecting Ensemble: A New and Efficient Technique for Dynamically Incorporating Observations into Physics-based, Multi-layer, Snow Models, Water Resour. Res., https://doi.org/10.1029/2018WR024521, 2018.