Evaluation of Snow Water Storage Estimations for Hydropower Catchments

Edvin Eklund, André Toresson

Department of Space, Earth and Environment
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021

MASTER THESIS 2021: SEEX30

Evaluation of Snow Water Storage Estimations for Hydropower Catchments

A project conducted together with Uniper

Edvin Eklund, André Toresson

Department of Space, Earth and Environment
Division of Microwave and Optical Remote Sensing
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021

Evaluation of Snow Water Storage Estimations for Hydropower Catchments
A project conducted together with Uniper
Edvin Eklund, André Toresson

Supervisors: Beatriz Quesada Montano, Uniper | Jie Zhang, Uppsala University
Examiner: Patrick Eriksson, Department of Space, Earth and Environment

Master Thesis 2021: SEEX30
Department of Space, Earth and Environment
Division of Microwave and Optical Remote Sensing
Chalmers University of Technology
SE-412 58 Gothenburg
Telefon +46 31 772 1000

Typed in LaTeX
Gothenburg, Sweden 2021

Acknowledgements

We would like to thank Uniper for granting us the opportunity to undertake this thesis. It has given us great insight into an important part of the hydropower sector, and specifically Uniper's hydropower production planning procedures.

Furthermore, we would like to extend our gratitude to the Norwegian Water Resources and Energy Directorate for their assistance in the form of making their internal data available for this project. A special thank you goes out to Jess Joar Andersen of the NVE for his personal involvement and going through the trouble of providing us with the large sets of data used in the thesis, and his assistance in helping us interpret the data.

Lastly, our greatest gratitude is directed towards our supervisors Beatriz Quesada Montano of Uniper and Jie Zhang of Uppsala University, who have been a great help to us during this spring. Their feedback and comments have given us both knowledge and encouragement which played a vital role in the progress of this thesis. We would also like to direct a special thank you to our examiner Patrick Eriksson of Chalmers, who has given us continuous feedback and guidance to keep us on track throughout the progression of the thesis.

André and Edvin

Gothenburg, Sweden, June 2021

Evaluation of Snow Water Storage Estimations for Hydropower Catchments

Edvin Eklund, André Toresson

Department of Space, Earth and Environment

Chalmers University of Technology

Abstract

Sweden's power production is in the midst of a paradigm shift, where predictable carbon- and nuclear-based production is being replaced by renewable energy, such as wind and solar. Since both of these energy sources are intermittent, this introduces predictability issues to the energy production sector. Hydropower energy does not directly emit carbon dioxide in its power production, and it is to some extent a predictable source of energy. However, this predictability can be improved with enhanced estimation of the water stored in the form of snow over the catchment areas involved.

The purpose of the project was to evaluate the impacts that topography and wind have on snow distribution over a catchment in northern Sweden and trying to model the snow conditions based on this, and how well snow and weather data-based and satellite-based snow data products could predict the snow water equivalent volumes within it. The analyses and evaluations were based on in situ data collected within the catchment area during one week of the snow season of 2018 to 2020 and consisted of data for, among other things, snow depth and density.

Several topographical features were found to consistently correlate with SWE, up to 69%, over the catchment. Furthermore, models based on all investigated features could, at best, describe up to 51% of the snow variability. Even though the wind direction was quite evenly distributed among all cardinal directions, the accumulation of snow showed correlations with the net wind direction of the season, although with consistent signs of other factors influencing the direction of accumulation as well. Despite evaluating the satellite-based model with the highest resolution available, it displayed very low accuracy in describing the snow's spatial distribution, as well as the aggregated volumes of SWE on the scale tested in this project. The snow and weather data-based data appeared, with slight adjustments, to be the only observational method successful in describing the snow's spatial distribution, as well as accurately predicting the aggregate snow water volume.

Keywords: Catchment, Copernicus, GIS, GPR, Hydropower, Runoff, Snowmelt, SWE

Abbreviations and Acronyms

DEM	Digital Elevation Model
GIS	Geographic Information System
GPR	Ground Penetrating Radar
MLR	Multiple Linear Regression Model
NVE	Norwegian Water Resources and Energy Directorate
SMHI	Swedish Meteorological and Hydrological Institute
SWE	Snow Water Equivalent

Contents

List of Figures

List of Tables

1

Introduction

Hydropower is Sweden's largest renewable source of energy and it plays an important role in the overall balance and stability of the Swedish electricity grid. Since there is an overall trend of drastically increasing intermittent renewable energy sources (such as wind and solar power) introduced to the market, the need for a responsive and predictable energy production will increase, especially as Sweden is currently phasing out its nuclear energy. With this in mind, it is imperative to make the best possible use of the large scale hydropower generation that Sweden already has installed to allow for overall sustainability and predictability in Sweden's power production.

1.1 Background

A large share of Uniper's and Sweden's hydropower plants and dams are regulated on a yearly basis, meaning that the producers plan their operations accordingly. A large part of this planning involves knowing how much to reduce the levels of the reservoirs during winter to free up enough space to store the meltwater that starts filling up the reservoirs during the spring. This ensures that there is water available for production later in the year, especially during winter months when the electricity demand is at its highest. A satisfactory optimization and planning depends therefore on a reliable estimation of the water volumes present in the snowpack.

A common way to estimate the amount of water present in the snowpack is to take manual measurements of the snowpack within a catchment area, read its depth and density, and translate the snow to a Snow Water Equivalent, i.e., how much liquid water that the snowpack corresponds to. This results in an indication of how much water that will be freed up and fed to the system once the snow has melted. However, in today's practice, SWE is only measured on a few occasions and at a rather low number of locations, which makes it difficult to get a clear picture of the SWE's spatial distribution and its variation over time. Furthermore, the manual methods used today are very time consuming and costly, and are showing difficulties in accurately estimating the amount of water stored as snow in mountainous regions and other areas of cumbersome terrain. This is mostly due to factors like inaccessibility, complex topography, or other parameters that do not allow for physical ground measurements. If these estimations are inaccurate, and power producers underestimate the amount of water that will add to the system, reservoirs can overfill, and excess water may have to be wasted. This is also a matter of safety, as an unexpected amount of water in the reservoirs will put an unpredicted strain on the physical structure of the dams. On the other hand, an overestimated amount of meltwater is also a waste of resources as it causes unpredictability in the overall energy system, creating a need for other, not unlikely, less sustainable energy sources to cover for the shortage of power and/or energy. There are, however, other means of performing these estimations in the form of products that continuously monitor the SWE, and in theory these can improve the efficiency and accuracy of this process. It is however not clear how such products perform in certain hydropower applications.

1.2 Purpose

The purpose of this thesis is to evaluate other means of performing snow water storage estimations. This particular project is dedicated to investigating three different main options in doing so.

Firstly, it is of interest to see how well it is possible to model snow conditions, i.e., depth and distribution, based on topography. A successful model of this kind has the main advantage of mitigating the need for manual measurements as it ideally only needs a form of starting value for the snow to model the distribution within an area. It also has the potential to increase the overall accuracy in snow water storage estimations. In addition to topography, it is desirable to retrieve an understanding of how the wind impacts the snow distribution, specifically how it contributes to snow accumulating in certain directions.

Secondly, it is also of interest to evaluate the hydropower catchment specific performance of two different SWE products, both from a spatial distribution perspective and on an aggregated scale. The main reason for this is that a fully functional SWE product could eliminate the need for manual measurement altogether, while also retrieving accurate snow water storage estimations for multiple catchment areas. This project has evaluated two fundamentally different types of SWE products, snow and weather data-based models and satellite-based products.

1.3 SWE estimation products

Since this thesis aims to evaluate two fundamentally different SWE products it is important to have the basic information on how the two types function and how they operate.

Modeled data

One way of estimating the amount of SWE over large land areas is to mathematically model the snow conditions based on snow and weather data. These types of data can often be obtained from meteorological weather stations, or from manually performed snow depth and density measurements. However, the input data often requires information on precipitation, temperature, humidity, atmospheric pressure, and has to originate from a large number of places within the region of interest to be sufficiently accurate. This often makes manual measurements a nonviable option in this application due to practical/economical reasons. Nonetheless, if such data are available and of a high quality, accurate models can be constructed to continuously monitor the snow conditions, including SWE, in large geographical areas.

Most likely due to its practical limitations, being that some form of in situ data of the snow conditions are required, publicly available snow data models that cover vast land areas have not been found. However, there is an interplay between Swedish and Norwegian hydrological resources. As a result, the Norwegian Water Resources and Energy Directorate (NVE) provided snow data generated by their model to be available for this research project. NVE includes observed information from the Swedish side of the border, and therefore also models the snow on a significant land area within Sweden. As the catchment investigated in this study falls within the area of the NVE model it can be utilized in this application.

Satellite data

Satellites can be used to measure SWE by utilizing passive microwave radiation. Microwave radiation is emitted from objects at the Earth's surface and dependent on an objects physical properties such as crystalline structure and atomic composition. This makes it possible to distinguish ice and snow from noncrystalline liquid water, since crystalline structures typically emit more microwave radiation. However, the emitted microwaves are at relatively low levels, resulting in radiation requiring to be collected over large geographical extents [1].

Satellite-based products generally do not have geographical limitations and are therefore quite often used as tools to describe the cryosphere on large scales, such as whole continents or even larger. However, a significant drawback of such products is their inability to capture variations on smaller scales. They are typically useful tools in climate research or similar large scale investigations on cryospherically related phenomena [2], but are rarely presented as useful components in small scale hydrological prediction tools. The data are often generated at relatively large resolutions, with some of the well known products such as Globsnow and AMSR2 having resolutions of 25km x 25km grids, and Copernicus Global Land Service's a 5km x 5km resolution.

1.4 Delimitations

This project will limit its research and data analysis to one specific catchment area. The reason for this is that Uniper have relevant reference data obtained from this catchment, enabling comparative analyses of other snow data applicable to the catchment. However, these data have, at the time of this project being conducted, only been produced for the three previous years. As a consequence, the majority of the results produced by this project will only be valid for the snow conditions of three years.

As the two different SWE products are evaluated, another form of reference data, also provided by Uniper, are applicable. These data were however only available for the twelve previous years, as were the results produced by the SWE product evaluations.

The project will focus only on the satellite product which offers the highest resolution, i.e., the one offered by Copernicus Global Land Services, which has a spatial resolution of 5km x 5km, rather than the products offered by Globsnow or AMSR2, which both have a spatial resolution of 25km x 25km.

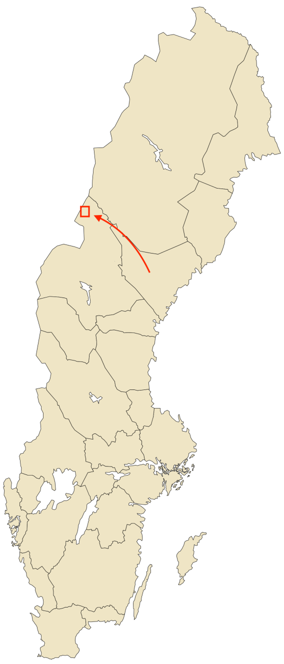
2

Basic Information & Data Description

This project has had its focus on one specific catchment area, several different types of information related to the said catchment, and on two specific SWE estimation products. This chapter aims to inform the reader on all fundamental data and information that has been used as the basis for all analyses that have been conducted throughout the project.

2.1 Catchment area description

The catchment area of interest is located mainly within the northern parts of Jämtland county, but crosses the border between Sweden and Norway in its western and northern parts, and has a surface area of about 940 km^2, see figure 2.1.



(a) Approximate geographical location of the catchment area.

(b) The catchment area with land area in grey and bodies of water in white.

Figure 2.1: Map of Sweden indicating the geographical location of the catchment and the investigated catchment area's shape and size.

The rain and snowfall occurring over this area feeds the power plants downriver of the catchment, i.e., along the river Fax. Since the catchment is located quite far up north, at around 65°N in latitude, and in a region that is prone to precipitation, the snow seasons are quite long with an extensive snow cover often present from some time in October through June or even July, and the amount of snow is typically palpable with snow depths often reaching several meters.

As mentioned, the area is located in a mountainous region, however, not all area is at a higher elevation than the so-called treeline. This means that some share of the area has a forest canopy cover, which is vital information when analyzing the snow distribution. Binary values describe the presence of forest, i.e., shows that a certain geographical location either does or does not contain forest canopy cover, and no information on the specific canopy density. The forest cover over the area is visually represented in figure 2.2.

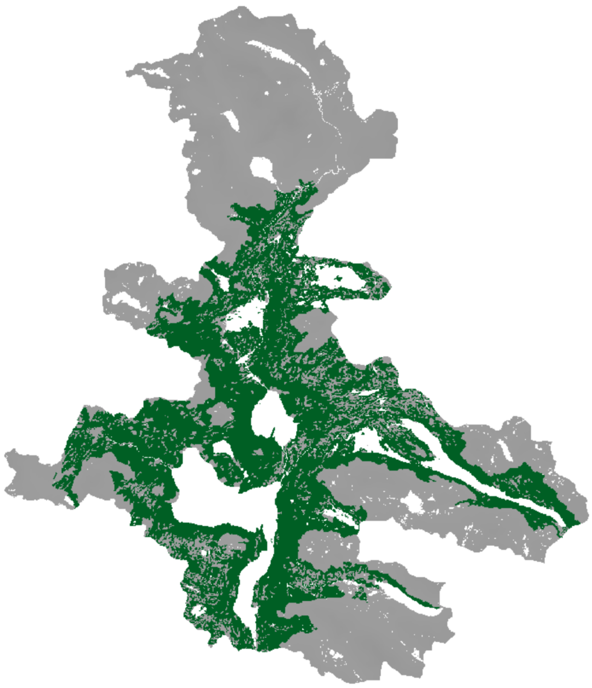


Figure 2.2: Forest cover over the catchment and with a 10x10m cell size.

2.1.1 Topographical information

Information regarding physical properties of the catchment was provided in the form of two Geographical Information System-files (GIS) provided by Uniper. These consisted of the previously mentioned forest cover information and a DEM-file (Digital Elevation Model grid). The DEM over the area is shown in figure 2.3.

The information was provided in raster grids, a form of matrix that contains information on each cell's geographical location, and a value of a relevant property, e.g., altitude. All rasters provided had a 10x10 meter cell size. Since features such as aspect [=cardinal direction (0-360°) a hillside is facing], slope [=incline of a hillside], and surface roughness [≈the terrains ability to capture snow], are all related to the elevation model, these are all accessible through the DEM-raster to further describe the area's topography.

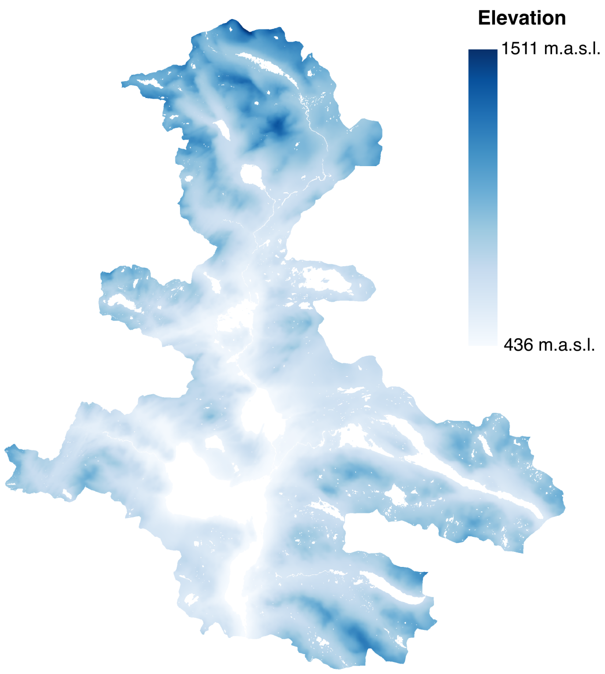


Figure 2.3: The DEM (in meters above sea level) of the area at a 10x10m cell size.

2.2 Hydrological budget within the catchment

The catchment area is the area over which precipitation, whether it being rain or snow, adds to the water budget for the hydropower plants downriver of the catchment. However, it is important to distinguish the total precipitation over the area from the water budget for the hydropower plants. The total precipitation is the accumulated rain, snow, or frozen water that falls over the surface of the area. The water then stays in the soil for some period of time, runs down into the ground water, or drains away. Water partially drains away over the surface, called surface runoff, and partially under the surface, called subsurface runoff. Combining the water drained from both of these factors is called runoff and is what adds to the total amount of accumulated water which is usable for hydropower production. Furthermore, snowmelt is the water contained in the snow over the catchment that eventually melts, and just like precipitation, either remains in the soil for some time, or drains. Assuming that all melting snow drains to the hydropower reservoirs, snowmelt and rain combined would equal runoff [3]. However, a fraction of water is stored in the soil, some evaporates, and some sublimates. As these fractions are unknown and depend on the geological conditions of a particular area, only rough estimations of how much of the snowmelt actually leads to runoff are possible [4].

2.2.1 Runoff

In order to be able to evaluate the precision and the consistency of various SWE models and products, it is important to have a form of reference value. In this particular case there are numbers describing the total water volumes running from the catchment, the inflow, during spring and summer, provided by Uniper. These numbers stem from what Uniper considers the standard snowmelt period of April 16th to July 31st, and includes the time where the vast majority of the snowmelt occurs, including spring floods. The accumulated water volumes simply describe the amount of water flowing to the power plants downstream of the catchment. In other words, this amount should roughly describe the total amount of runoff from melting snow and rain. In order to be able to distinguish between amounts within the runoff that derive from snowmelt and from rain, Uniper also provided data on precipitation in the form of daily weighted average precipitation data from three precipitation stations in the area. Figure 2.4 shows the accumulated amounts of runoff and rain that was added to the system during the snowmelt period for the last twelve years.

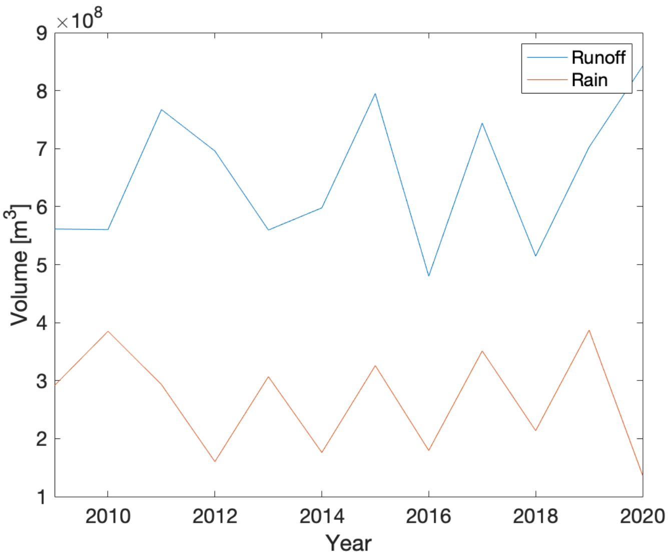


Figure 2.4: Average volumes of water from runoff and precipitation over the catchment from the period of April 16th to July 31st each year.

2.3 Ground measurements

Uniper provided three sets of data containing ground measurement data from 2018, 2019, and 2020. These measurements were taken over a five-day period in mid March, estimated to be when the maximum SWE of the season occurs according to Uniper. The data contain information regarding geographical location, altitude, snow depth, density, SWE and distance driven for every data point. However, other than the fact that the measurements are performed during week 11 of each year, there is no information regarding what time any specific line is measured.

The measurements are taken at strategically positioned snow lines with the purpose of getting an accurate representation of the entire catchment, the validity of which has been proven by a study produced by an external source by the order of Uniper. There are in total eleven snow lines spread over the area, see figure 2.5. The data are collected using ground penetrating radar attached to a snowmobile. This technique has the capability of quickly assessing the characteristics of the snow cover using radar pulses. Measuring the travel time of the radar pulse through the snow cover produces a snow depth. The snow depth is then converted to SWE using a density estimation for the snowpack, which is produced by a number of manual density measurements along each snow line, in total 67 measurements. These measurements and the corresponding snow depth create a linear relationship for the entire line, meaning each snow depth measurement can be converted to SWE using a linear function relating snow depth to density.

The GPR data were generally collected at every driven meter for the years 2018 and 2020, and for every 0.5 meters for 2019, but deviations were present in all three years. However, there were two anomalies, (1) one of the lines (line 7 in figure 2.5) in the 2019 data is missing data in the middle of the line, and (2) that the same line was manually measured at every 100 meters in 2020. The resulting SWE for all snow lines from the 2018 measurement campaign is visualized in figure 2.6. The dotted vertical lines represents the end of each line, with the corresponding line number being in the center of each section.

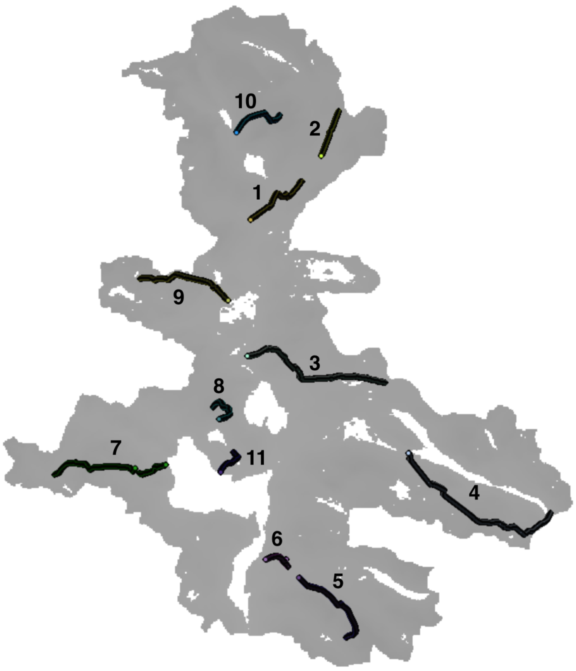


Figure 2.5: GPR data collecting routes and their respective numbering.

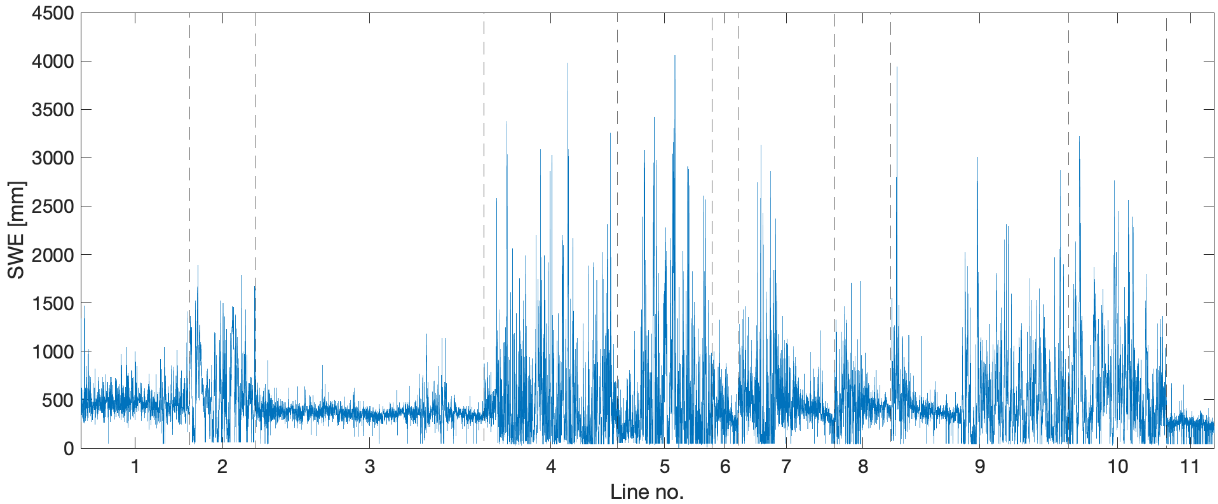


Figure 2.6: SWE obtained using GPR during 2018.

2.4 Wind data

Wind is an important aspect to consider when analyzing the spatial distribution of snow. Unfortunately, the catchment area lacks in-situ measurements for wind. Data is therefore retrieved by accessing a numerical model, namely, the ERA5-land reanalysis set.

ERA5-land delivers information on surface variables with hourly updated, high resolution (9km) information [3]. It contains information from several decades back to a couple of months before the present day. ERA5-land utilizes atmospheric forcing, meaning it uses atmospheric variables as inputs to control the simulated land fields. Since ERA5-land is a numerical model it contains some uncertainty. However, as the number of observations that are available have increased with time, the model estimates have become increasingly accurate, as the quality of the atmospheric forcing becomes higher. The input variables are obtained using a 4D-VAR data assimilation system and a Simplified Extended Kalman Filter. The results are generated during a single simulation and run without data assimilation, which makes it computationally affordable.

The wind data presented in this report is part of the information available from ERA5-land. The data consist of two components, one northward and one eastward component. Combined with each other gives the speed and direction of the wind. The data represents horizontal winds at 10 metres above ground. Since the wind is strongly affected by local terrain and therefore varies on small space and time scales, caution should be taken when comparing with observations as they are represented on average in the ECMWF Integrated Forecasting System. An example of the raw wind data is illustrated in figure 2.7.

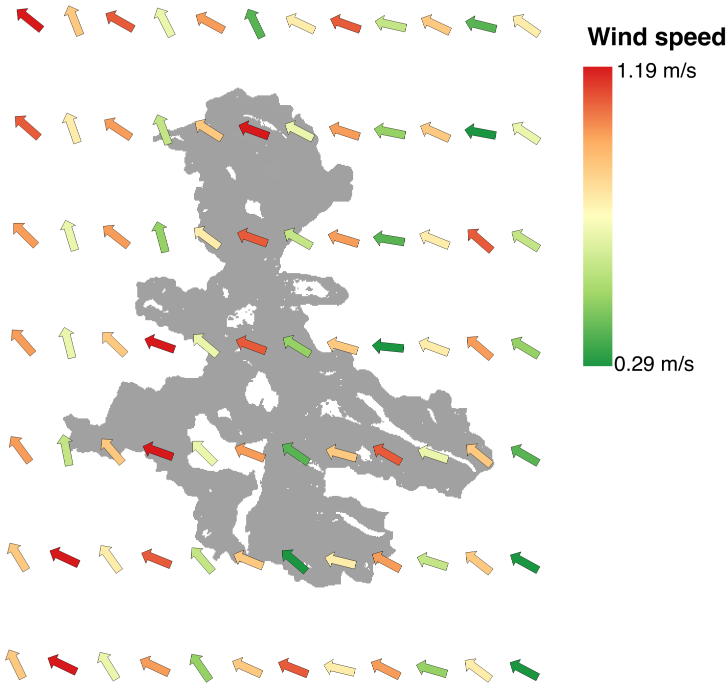


Figure 2.7: Raw wind data extracted from ERA-5 on March 12th, 2018, over the catchment area. The color of the arrows describes the wind speed, and the direction of the arrows shows the wind direction.

2.5 The Norwegian Water Resources and Energy Directorate's SWE product

The Norwegian Water Resources and Energy Directorate, NVE for short, uses various weather and snow observations to create a 1km x 1km grid across all of Norway in order to maintain a better overview of, among other things, their snow conditions and water aspects. The NVE snow model has long used temperature and precipitation as forcing inputs, and estimates SWE based on SMHI's HBV-model [5] and a snow compaction and density module of their own. However, to avoid the overestimation that these two modules alone tend to generate, the NVE model now also applies two extensive sets of in situ snow measurements [6].

NVE provided data of snow conditions, namely SWE, in raster format. The provided data were the full results produced by their model, and the data therefore covered the same area that they model, being all of Norway, a decent share of western Sweden, and even the northern parts of Finland. Isolating the data cells relevant for the catchment, the data could be visualized as in figure 2.8.

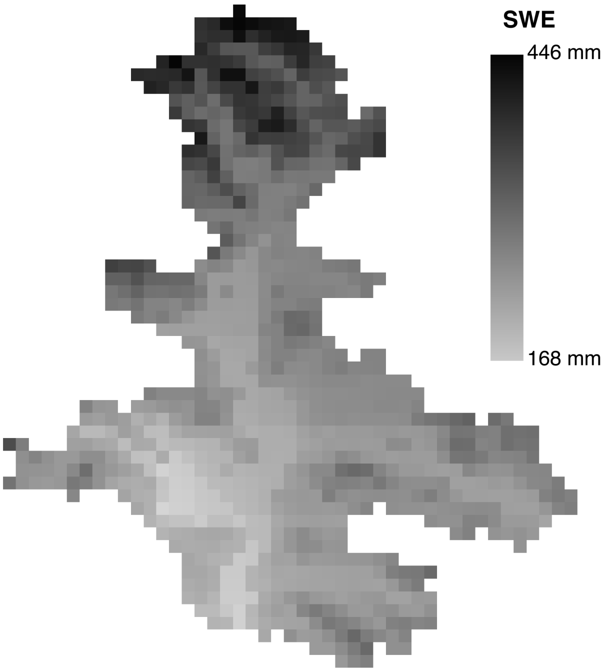


Figure 2.8: NVE data from March 12th, 2018, covering the catchment.

2.6 Copernicus Global Land Services' SWE product

The Copernicus Global Land Service provides a variety of different products within earth observation and cryosphere, including a specific SWE product (hereafter also referred to as Copernicus). According to Copernicus' own SWE product user manual [7], this product utilises a retrieval algorithm which combines information from satellite microwave radiometer and optical spectrometer observations with snow depth measurements acquired from ground-based weather stations. The resolution is enhanced by snow extent information extracted from satellite data, resulting in a spatial resolution of 5km. The data cover land surfaces between latitudes 35°N-85°N, excluding high-mountain regions (not affecting the study area investigated), glaciers, and Greenland.

As previously mentioned, the algorithm uses data from weather stations to calculate both snow depth and snow grain size, resulting in the accuracy being dependent on the density of the network. A static snow density value of 240 kg/m^3 is used. Since snow generally varies between 100 kg/m^3 and 400kg/m^3, the results are subject to biases. Snow estimates are often inflated during early accumulation and underestimated during late winter and melting periods.

The extracted raster data are presented in figure 2.9, describing the SWE distribution over the catchment area. Unfortunately, the Copernicus grid lacks data at a total of eight grid cells located over and close to the largest lake in the southern part of the catchment.

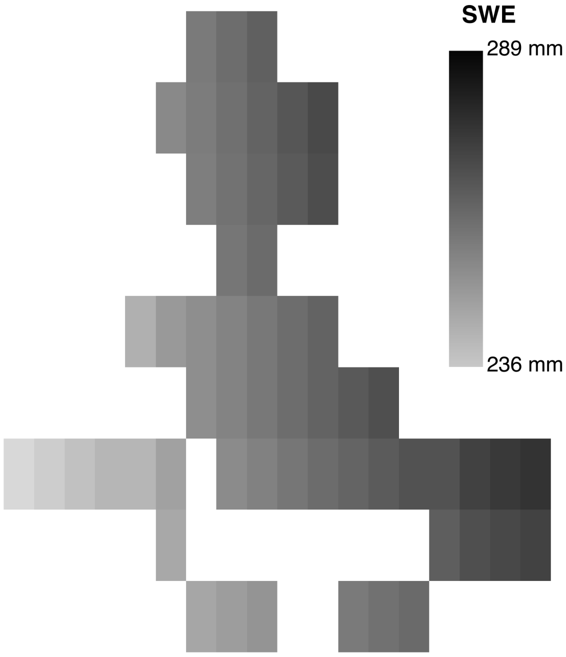


Figure 2.9: Copernicus data from March 12th, 2018, covering the catchment.

3

Method

The project consisted of two main parts, analyzing the impacts that topography and wind has on snow distribution, and evaluating two different SWE products. Figure 3.1 shows how the work was structured and how the various parts of the project relate to each other. Firstly, all relevant and useful data were extracted. The first analytical work focused on the spatial distribution of SWE, and once a proper spatial resolution was found, it was possible to analyze how the distribution of snow within the catchment relates to the topography. This was done with the purpose of being able to construct a regression model that can predict the snow distribution within the area using topographical data and a starting value of the snow depth, as previously mentioned. Moreover, analyzing the spatial distribution of SWE also allows for analyses regarding how wind impacts the snow accumulation within the area, as well as for evaluating how the two SWE products perform on a pixel-wise level.

Since it was also important to evaluate how the two SWE products perform on an aggregated scale, catchment-wide SWE estimations were needed. The two products' performances were evaluated on this scale using two different references, being the GPR data and the snowmelt data.

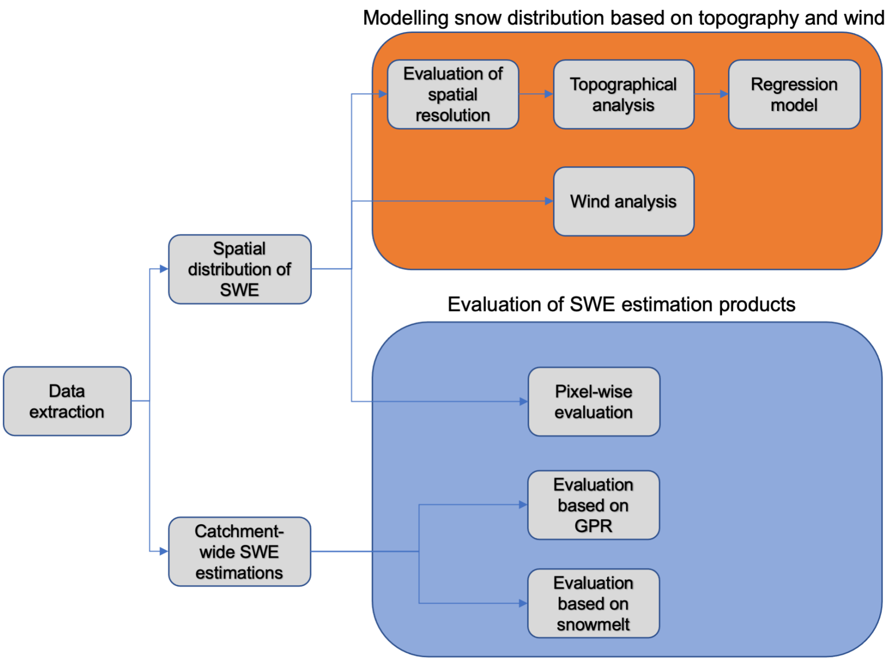


Figure 3.1: Flowchart showing the structure of the work conducted in this project.

3.1 Spatial heterogeneity of snow distribution

Spatial heterogeneity of the snowpack over a landscape is affected by a large number of factors, such as precipitation, topography, forest cover, wind, solar loading, and avalanches. The many factors influencing the spatial heterogeneity of snow makes it variable even at scales of a few meters. It is important to characterize the snow's spatial variability over the catchment to determine whether the SWE products mentioned in this report have the potential to capture the high variability of snow distribution. Furthermore, the spatial heterogeneity affects the predictability of the relationship between SWE and various topographical features [8].

At small scales the relationship between SWE and topography is highly variable and unpredictable. This is due to random factors including micro-topography, fallen logs, and local variation of wind. Increasing the scale of the observations increases the predictability substantially by removing the large local variability [8]. The scale which removes this unpredictability but retains an accurate representation of medium-scale variability is estimated using variograms.

Variograms analyze the spatial dependency of a variable. Two data sets may appear identical when depicted using histograms or similar statistical methods which does not factor in the spatial location of the data. However, despite the two data sets appearing identical, the spatial continuity of the data sets can still differ, making one data set change faster with distance. The spatial continuity substantially affects the sample design, site characterization and spatial prediction, making variograms crucial for the analysis of a data set. Essentially, a variogram visualizes the variability of data points as a function of distance. In general, points at close distance to each other exhibit low variability while points with large distance between them will have a high degree of variability. A range represents the distance at which the data points are still considered related to one another. Beyond this point, the variogram levels out into a "sill", see figure 3.2. Observational methods that use cell sizes significantly larger than this range will fail to fully capture the spatial heterogeneity.

Variogram analyses are comprised of an experimental semi-variogram derived from the data, and a variogram model which fits a mathematical model to the data. For all pairs of observations, the difference squared is averaged and then divided by two to create an experimental variogram, i.e., creating an expression for the variance between each observation. This is done using a specified separation distance and direction. The model for fitting the data is based on the shape of the curve of the experimental variogram, and there are three fundamental models for fitting the variogram: (1) A linear model that is suitable for experimental variograms which lacks a sill, i.e., never levels out, (2) an exponential model which is appropriate if the experimental variogram levels out but shows a sinusoidal behaviour, and (3) a spherical model that is preferable if the experimental variogram starts linearly and shows a sharp bend at the sill. A representation of an experimental semi-variogram, its corresponding variogram, and how to read the variogram is shown in figure 3.2.

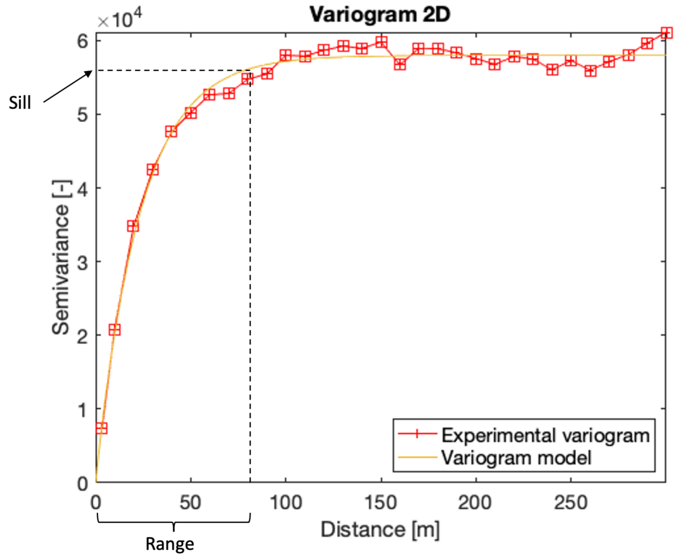


Figure 3.2: Explanation of variogram constructed with an exponential model.

In this project, variograms were produced for all separate snow lines to analyze how the spatial heterogeneity varies as a result of location in the catchment. If dependency on location is found to be a decisive factor in spatial heterogeneity, the requirement for accuracy of observational methods also varies across the catchment. Additionally, variograms for all lines together were created. These gave a general estimate of a spatial resolution which reduces small scale variations across the entire catchment.

3.2 Topography- and wind-based impacts on snow distribution

Snow distribution can not be accurately depicted solely by snowfall data since it is highly affected by the characteristics of the location, surrounding topography, and wind. For example, in areas of high wind speed, the snow is eroded and transported to areas of lower wind speed. Locally varying wind speed is a direct result of the topographic features of the landscape, where tall vegetation and mountains serve as shelter. The resulting accumulation should therefore occur at the face of topographical features in the direction of the wind, with the main accumulation being located behind the shelter and to some extent in front of it [9].

Topography's exact effect on snow distribution is somewhat ambiguous, and the correlation between different terrain properties and SWE can vary substantially within a season and in different locations. In essence, the distribution of snow is mainly dependent on two factors, differential ablation and differential accumulation, but how the underlying factors affect these two was to some extent unknown. As an immediate step in analyzing these underlying distribution factors, a total of five different topographical features' effect on the amount of SWE were analyzed: (1) A correlation between elevation and SWE can occur due to higher levels of precipitation at higher elevations, while thermal effects cause snowmelt or rain at lower elevations.

(2) High canopy density leads to increasing accumulation compared to sites with no forest at the basin scale. This is due to differences in ablation, with the forest providing shade. Furthermore, open areas with surrounding forested sites increase accumulation as a result of wind-interception. However, accumulation due to interception decreases as the season progresses. (3) Aspect relates to ablation as a result of varying solar loading depending on the direction the topography is facing, where the south and west aspects exhibit the highest loading. Furthermore, accumulation differs on aspects in lee and windward slopes as snow is redistributed by wind. (4) Slope mainly affects the local distribution of SWE as a result of lateral flow of meltwater within a snowpack and redistribution by avalanches and creep. Its correlation with SWE is also expected due to wind redistribution. (5) Roughness is defined as the surface's degree of irregularity. It serves as a characteristic to describe the area's ability to retain snow, mainly as a result of snow redistribution from wind [10].

3.2.1 Topographical impacts on snow distribution

Characteristics which exhibit high correlation with SWE are useful as predictors for estimating SWE where no measurements are available. Using the variogram analyses, a proper raster resolution could be determined and the GPR and topography data could be rasterized accordingly. Another benefit of rescaling the topographical data was that the original binary forest canopy raster data could be averaged, meaning that the canopy density would no longer show a binary behaviour, but any intermediate value from 0 to 1. From the rasterized ground measurements of the snow, together with the matching GIS-extracted information, several correlation analyses were carried out. What was initially considered of interest was whether the SWE was significantly affected by any one specific topographical feature, e.g., if an increasing elevation corresponds with an increasing amount of SWE, or if the presence of forest canopy directly relates to a lower SWE. A correlation coefficient was calculated to describe how heavily the SWE value was impacted by each feature, for all three years, for all lines combined and for all lines individually.

To, as accurately as possible, model the SWE distribution over the catchment, multiple linear regression models using all topographical features as explanatory variables were created. In total, nine models were created, each describing a specific year and resolution. Additionally, a generalized model able to describe the SWE over the snowlines with only topographical information and a mean SWE over the catchment as input was created. This model meant that estimating the snow distribution could be less cumbersome than using information from many manual measurements, as only one value of SWE is needed. To evaluate the models' performance, two different metrics were used, coefficient of determination and correlation between the model and the GPR data. The coefficient of determination, R^2, explains the models ability to capture the variability present in the GPR data. A value of zero means that the model explains none of the variability of the data, while a value of one means that the model fully depicts the variability of the data.

3.2.2 Wind-based impacts on snow distribution

It is clear that the wind has an impact on snow distribution, however, the degree to which it affects this particular catchment was unknown. The significance of wind on snow distribution was therefore studied for each snow season of 2018 through 2020 using the GPR data. The analysis compared the direction of the main accumulation of snow to the dominating wind direction and average magnitude. Correlation of wind direction and direction of accumulation on topographical features served as an indication of the wind's impact. In order to obtain the aspect information related to the GPR data, the DEM provided by Uniper was converted to the corresponding aspect raster, and the raster cells geographically intercepted by the snow lines were extracted.

3.3 Evaluation of Copernicus and NVE SWE in the catchment

Both the NVE model (section 2.5) and the satellite-based observational product, Copernicus (section 2.6), are used as fully functional tools in some applications. However, their respective accuracies and behaviours still needed evaluation for this specific application and catchment size. The evaluation was conducted through several analyses with both ground measurement data and snowmelt volume.

3.3.1 Evaluation of spatial variations based on ground measurements

To gain a better understanding of how the NVE and Copernicus products capture the spatial heterogeneity of the SWE, the two products were compared to the ground measurement data at their respective pixel-wise levels. Using the GPR data as references, the corresponding SWE data for the snow lines provided by the products can be evaluated to appreciate their general accuracy. To make such comparisons, the two sets of data (the ground measurements and one of the two products) need to be of comparable formats. Knowing the coordinates of the snow lines, the products' rasters could be cut to only contain the cells geographically intersected by the snow lines. Furthermore, the GPR data had to be rasterized to the same size and extent as each product. Since the ground measurements are produced over a five-day period, the model and satellite data were extracted for the same five days and averaged. Once the two sets of data are reformatted into two geographically and temporally identical rasters, they can be evaluated based on individual cell values.

3.3.2 Evaluation of aggregated performance based on ground measurements

Analyzing the aggregated values is useful to describe the differences in catchment-wide average SWE values from the GPR data, the NVE model, and the Copernicus product. The purpose of the GPR data was to accurately depict the snow conditions over the terrain in the entire catchment. The data can therefore be used to evaluate the other products' catchment-wide accuracy. Naturally, the NVE and Copernicus data need to be temporally extracted in the same way as for the pixel-wise comparisons, while the GPR data were not manipulated in any way. However, for this specific analysis, the lake surfaces were also of importance. The GPR data are only meant to depict the terrain properties within the catchment, and any aggregated values from other products should therefore not include the SWE data from the area over larger lakes. In this case, the NVE raster pixels over the two largest lakes were removed. In the case of Copernicus, the largest lake was already removed beforehand, and the smaller lakes did not interfere with enough area of any one pixel for anything else to be removed.

3.3.3 Evaluation of aggregated performance based on snowmelt

The main limitation with the GPR data was that the measurements were only produced for three time instances over three years. However, numbers for runoff and precipitation were available for 12 years. By subtracting the rain from the runoff for the whole season, an estimate for total snowmelt over the catchment was obtained. In combination with the size of the catchment, the snowmelt volume was translated to a corresponding average SWE value for the entirety of the catchment. The accuracy of the NVE and Copernicus products could then be evaluated on an aggregate level for a significantly longer time span than when using the GPR data. The snowmelt is compared to the SWE value at the time which each product yields its maximum average SWE over the catchment for each year. This gives an estimation of each product's accuracy regarding their aggregated values, assuming that no substantial amount of snow has melted and flowed downriver before the date of maximum SWE, or equivalently that the SWE does not increase after this date. However, as explained in section 2.2, the snowmelt numbers can only provide rough estimations of the actual snowmelt due to the complex hydrology of the system, and the estimations can therefore not be used as absolute references.

4

Results & Discussion

The chapter is split into three major parts and mostly follows the outline of the previous chapter. The first part analyzes the spatial heterogeneity of snow distribution, providing a foundation for the following section dedicated to the analysis of impacts on snow distribution from topography and wind. Lastly, a section evaluating the accuracy of Copernicus and NVE based on their ability to capture spatial variations and estimating the aggregated SWE values for the entire catchment is presented.

4.1 Spatial heterogeneity of snow distribution

Analyzing the spatial heterogeneity of snow within the catchment was vital to produce accurate models and evaluating the spatial resolution of observational methods. A suitable resolution was found using variograms. The shapes of said variograms were very similar between snow lines and years, which led to the fitted models utilizing an exponential function to fit the model to the experimental variograms in all cases. The different lines exhibited large variations in terms of both range- and sill-values. However, between years, the value of each line was similar, indicating that the interannual variation of spatial heterogeneity was low, implying that the characteristics of the landscape are a determining factor for spatial heterogeneity.

Unfortunately, the variograms for some lines never plateaued and therefore failed to provide an accurate estimation of range and sill. Furthermore, the measurements over line 7 during 2020 were manually measured, meaning the amount of data was greatly reduced compared to the other lines, and only consists of 110 measurements. This made the variogram quite pointless as there was a lack of variability within the small data set. The data from 2018 had two lines that failed to produce values for range and sill. Both lines suffered from highly erratic data, likely as a result of measurement errors or extreme variability at small scales. While a few faulty measurements did not significantly impact the mean SWE value for the line, it did have an impact on the variogram as the variance was affected by the significant changes in the data.

Variograms depicting the semivariance for all lines combined were created from the data of all lines. Similar to the individual lines, these variograms exhibited large similarities between the years, with all ranges being within 59 and 102 meters, and sills from 49,000 to 205,000. This indicated that 100 meters was a sufficiently large distance to consider the data points unrelated to one another for every year. All sill and range values generated in this analysis can be seen in tables 4.1 to 4.3.

Table 4.1: 2018 variogram values.

Line #	Sill [-]	Range [m]
1	11312	7.98
2	109977	131.18
3	2499	20.96
4	104230	190.92
5	-	-
6	19607	38.05
7	45377	75.55
8	22294	25.91
9	118624	43.86
10	215460	81.96
11	-	-
All	51489	65.78

Table 4.2: 2019 variogram values.

Line #	Sill [-]	Range [m]
1	18011	20.24
2	240776	105.54
3	2319	42.95
4	104230	64.06
5	139457	75.54
6	6062	7.55
7.1	31485	33.16
7.2	18204	26.67
8	41830	29.55
9	300491	389.54
10	340126	129.14
11	712	112.22
All	49670	59.76

Table 4.3: 2020 variogram values.

Line #	Sill [-]	Range [m]
1	22212	14.37
2	351979	100.69
3	3030	31.36
4	428232	144.30
5	363795	104.73
6	5324	12.33
7	-	-
8	127677	37.91
9	572840	459.76
10	447527	121.81
11	2323	40.03
All	205076	101.54

4.2 Topography- and wind-based impacts on snow distribution

As previously stated, topography and wind are heavily intertwined when describing the distribution and redistribution of snow during a snow season. However, a goal of this project was to create a regression model that could describe the snow conditions over the catchment using topography and a mean SWE value as input, and without involving wind data. For this reason the analytical work involving wind was separated from that involving the topography.

4.2.1 Topographical impacts on snow distribution

Based on the range values given by the variograms (see tables 4.1 to 4.3), a 100m cell size appeared to be sufficient to remove most small-scale random factors inhibiting predictability. However, to ensure accurate results, an analysis of raster resolution was conducted. Three different cell sizes of 100m, which was a starting value obtained by the variogram analysis, 500m, which was a decent step up while not compromising too much of the variability of the data, and 1000m, which was considered the largest possible cell size before losing a vital share of the data's variability, were tested. These cell sizes were analyzed based on the correlation between SWE and individual topographical features, standard deviations, and performance of multiple linear regression models based on coefficient of determination and correlation with GPR data.

The correlation coefficients for each topographical feature, for all three years and at the three different resolutions were summarised in tables 4.4 to 4.6. Evaluation of each individual feature reveals that elevation has the highest correlation with SWE. The overall correlations increased with increased cell size, indicating that some small-scale variations remain despite increasing the size of the cells to over 100m. In this case, a 1000m cell size yielded higher correlations than the two lower cell sizes, but more importantly it was the only cell size to yield interannual consistency in the correlations of all features, as the correlations of any feature had the same sign and were in the same order of magnitude between the years.

Table 4.4: Correlation between SWE and each separate feature using a cell size of 100m.

Year	Elevation	Aspect	Roughness	Slope	Forest
2018	0.2267	0.065	0.1136	0.1336	-0.1914
2019	0.3874	-0.0185	0.2706	0.2639	-0.2778
2020	0.3668	-0.0955	0.2414	0.2271	-0.3233

Table 4.5: Correlation between SWE and each separate feature using a cell size of 500m.

Year	Elevation	Aspect	Roughness	Slope	Forest
2018	0.3901	-0.1149	0.0099	-0.0153	-0.3150
2019	0.5380	0.0160	0.247	0.2067	-0.4136
2020	0.4717	-0.0987	0.2849	0.2662	-0.4449

Table 4.6: Correlation between SWE and each separate feature using a cell size of 1000m.

Year	Elevation	Aspect	Roughness	Slope	Forest
2018	0.4359	-0.2961	0.2694	0.2109	-0.3111
2019	0.6903	-0.2987	0.4150	0.2952	-0.5577
2020	0.5915	-0.2912	0.3693	0.2599	-0.5530

Increasing the level of aggregation (cell size) smooths the variability of SWE, i.e., reduces the variability. In figure 4.1, the dependency of cell size on variability of SWE over the catchment is illustrated. The variability experiences a sharp decrease between cell sizes of 100m and 500m. Increasing the cell size to 1000m decreases the variability further, however, not as drastically as the previous step.

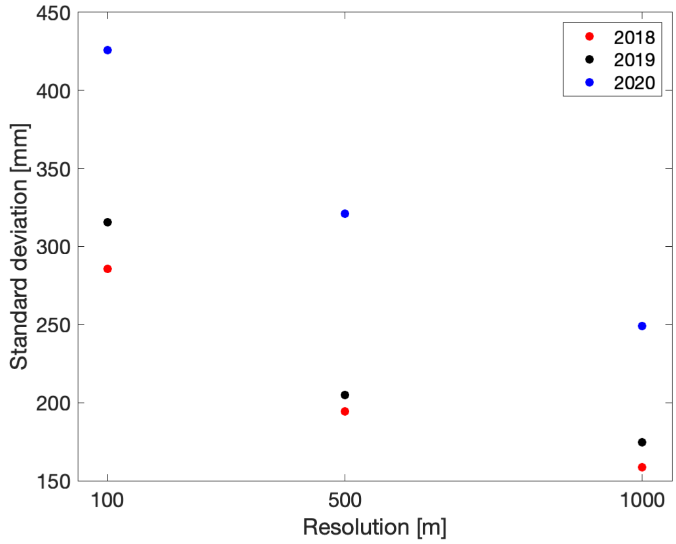


Figure 4.1: Relationship between level of aggregation (cell size) and variability of SWE in the form of standard deviation.

The predictability of SWE based on topography was analyzed using individual multiple linear regression models for each year. The regression models used all topographical features and the respective years' mean SWE value as input. The performance of the models was, unlike the variability, positively correlated with an increase of cell size, see figure 4.2. At a cell size of 1000m, the models are able to explain between 21% to 51% of the variability of the GPR data, with 2019 having the highest performance of all years regardless of resolution.

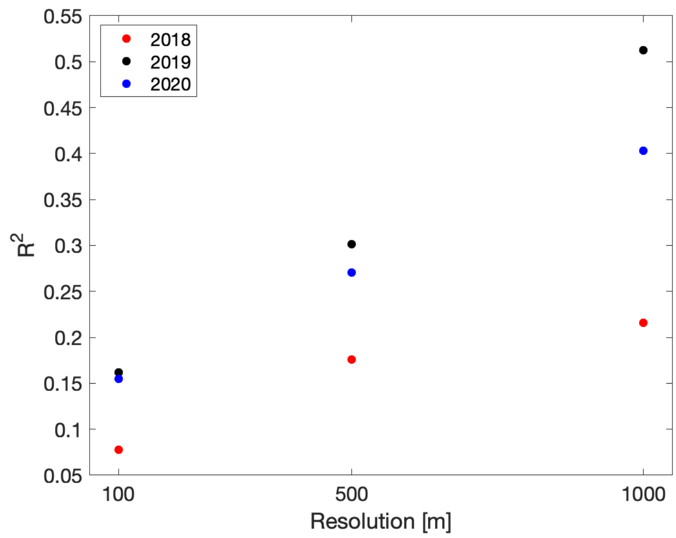


Figure 4.2: Relationship between regression model performance and level of aggregation (cell size) expressed as R^2.

Similarly to the R^2 value, the correlations between the GPR data and the regression models were significantly increased during each increase of cell size, see figure 4.3.

The regression models strongly correlated with the GPR data at 1000m, from 46% to 72% for all years. Once again, the performances of the regression models were at its highest for 2019.

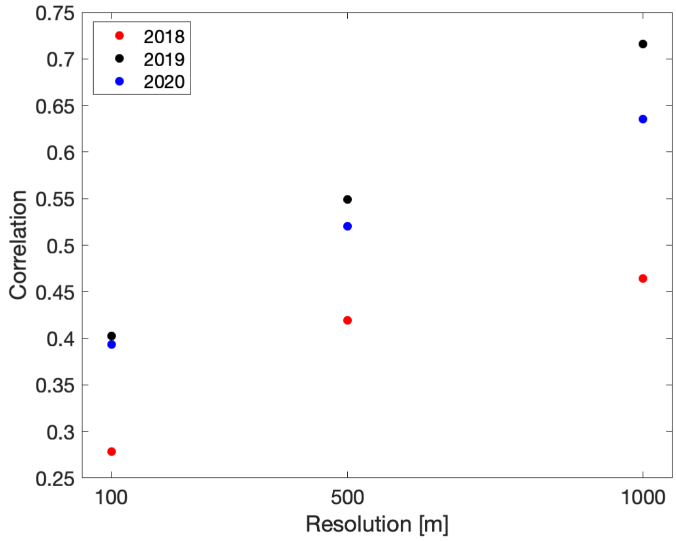


Figure 4.3: Relationship between regression model performance and level of aggregation (cell size) expressed as correlation between GPR data and regression model.

The results from the regression models indicated that a cell size of 1000m would provide the most successful model, at the cost of losing some of the variability of the data. A generalized model describing the SWE over the catchment is constructed by combining all GPR data into one large data set. The resulting regression model uses topographical information and mean SWE as explanatory variables. By scaling the regression model with a seasonal mean SWE value, a description of snow distribution over the snow lines, or even the full catchment, for that year is obtained. The resulting equation for the regression model, equation 4.2.1, uses all topographical features, albeit weighted depending on respective correlation with SWE. This gave a model expressed as

$$\begin{aligned} \mathrm{SWE} = \mathrm{SWE}_m * (&0.229421928701963 + 0.00102861752704567 * E- \\ &- 0.000288765747370705 * A + 0.00166654724033376 * R- \\ &- 0.0658412026692389 * S - 0.0674834050845746 * F) \end{aligned} \tag{4.2.1}$$

where SWE_m is the seasonal mean SWE value in mm, E is the elevation in meters above sea level, A is the aspect in 0-360°, R is the roughness in meters, S is the slope in 0-90°, and F is the forest canopy cover in 0-1.

Illustrated in figure 4.4 is the generalized model depicting the SWE distribution over the snow lines during 2018. Due to the interannual variability being so low, the generalized model performs well and follows the general behavior of the GPR data. However, the variability is lower and as a result, the model can not replicate the large spikes and dips present in the GPR data, even though it too is rasterized to 1000m cells.

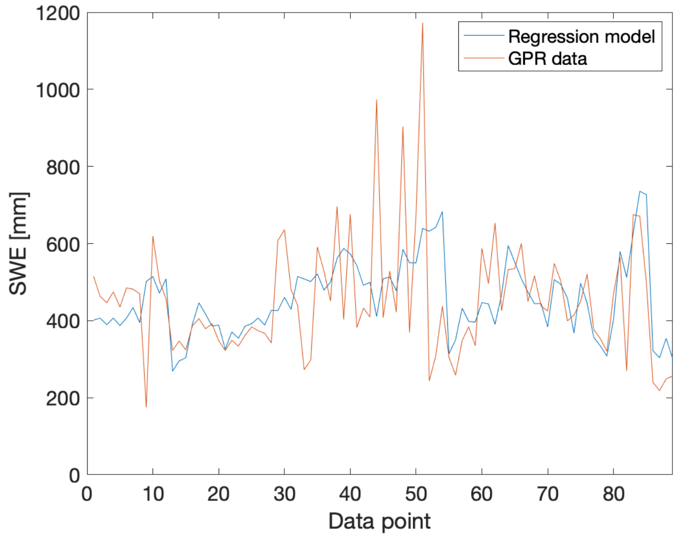


Figure 4.4: Generalized regression model of the SWE values over the snow lines for 2018. The GPR and modelled data are both rasterized to a 1000m cell size.

The coefficient of determination, R^2, for the generalized regression model was 0.3573, i.e., the model could explain 36% of the variability of the data set containing the GPR data for all years. The correlations between the GPR data at 1000m cell size and the generalized regression model are shown in table 4.7. Unsurprisingly, the generalized model performed similarly to the previous regression models, with 2019 showing the highest correlations while 2018 yielding the lowest. This is likely due to the increased amount of data points during 2019.

Table 4.7: Correlation between GPR data and generalized regression model, both with a cell size of 1000m.

Year	2018	2019	2020
Correlation	0.4511	0.7126	0.6179

4.2.2 Wind-based impacts on snow distribution

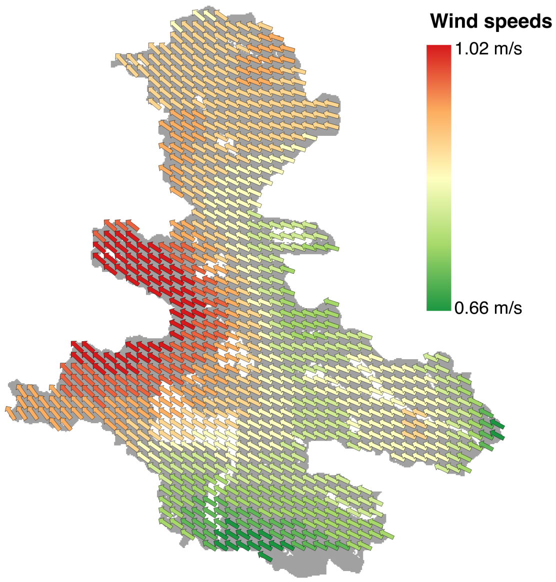
The wind analysis was conducted for the snow season of 2018 to 2020. The data was extracted over the relevant snow seasons, meaning it describes the winds over 5 months, from November 1st to March 31st each year, and three different plots visualize the results from the analyses related to each year's wind data.

Wind maps were created to describe the averaged wind data, i.e., the net wind's velocity and direction. The color of the arrows depends on the wind magnitude and the direction which the arrow is pointing describes the net wind's direction.

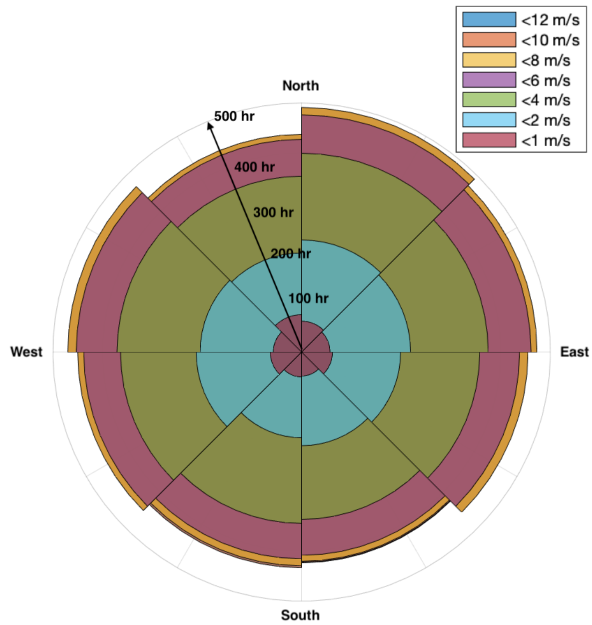
Polar histograms illustrate the direction and absolute magnitude of all winds during the season. It was divided into eight bins, each describing a specific wind direction. The radius illustrates the number of readings in a particular bin, which equals the amount of time that the wind was blowing in that direction. The colors represent the share of winds below a certain wind speed, giving an overview of the absolute magnitude of the wind speeds throughout the season.

Lastly, histograms depicting the direction of accumulation of snow on topographical features together with the GPR data were constructed. Analyzing the wind map in combination with the histogram related the dominating wind direction and speed to the spatial distribution of snow over the catchment.

During 2018, the dominant wind direction was about north-west, i.e., the wind originated from south-east and blew in a north-west direction, see figure 4.5a. The winds were quite evenly distributed cardinally but with a slightly higher frequency on the northern half of the histogram. The winds ranged from 0m/s to 10m/s with the mean wind speed being 2.6m/s, see figure 4.5b.



(a) Seasonal net wind velocity and direction.



(b) Polar histogram of seasonal wind speeds and directions.

Figure 4.5: Wind velocities and directions during the snow season of 2018.

A dominant wind direction to the north-west would in theory result in snow accumulating mainly on the north-west side of topographic features, or, in terms of aspect, at ~310°. Analyzing the relationship between SWE and aspect based on the ground measurements, see figure 4.6, reveals that the accumulation mainly occurred on the aspects of ~345°, but also to some degree aspects of ~45° and ~105°. This meant that the main direction of accumulation correlates with the wind direction to a high extent. However, the accumulation at ~45° is perpendicular to the wind direction, making conclusions regarding wind's impact on snow distribution uncertain during this year.

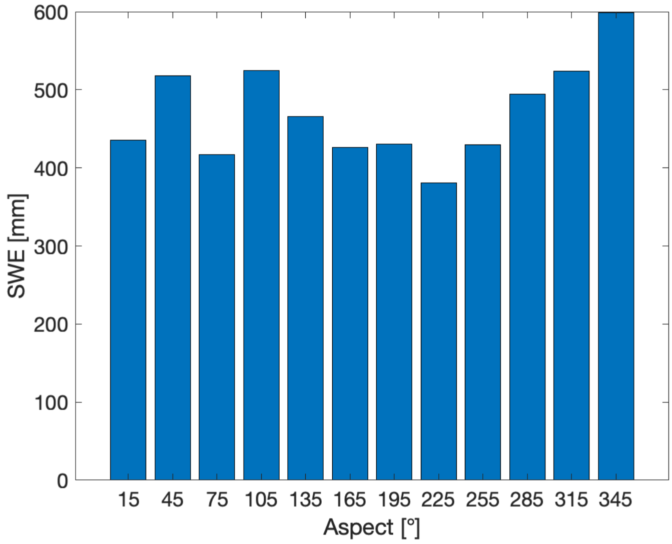
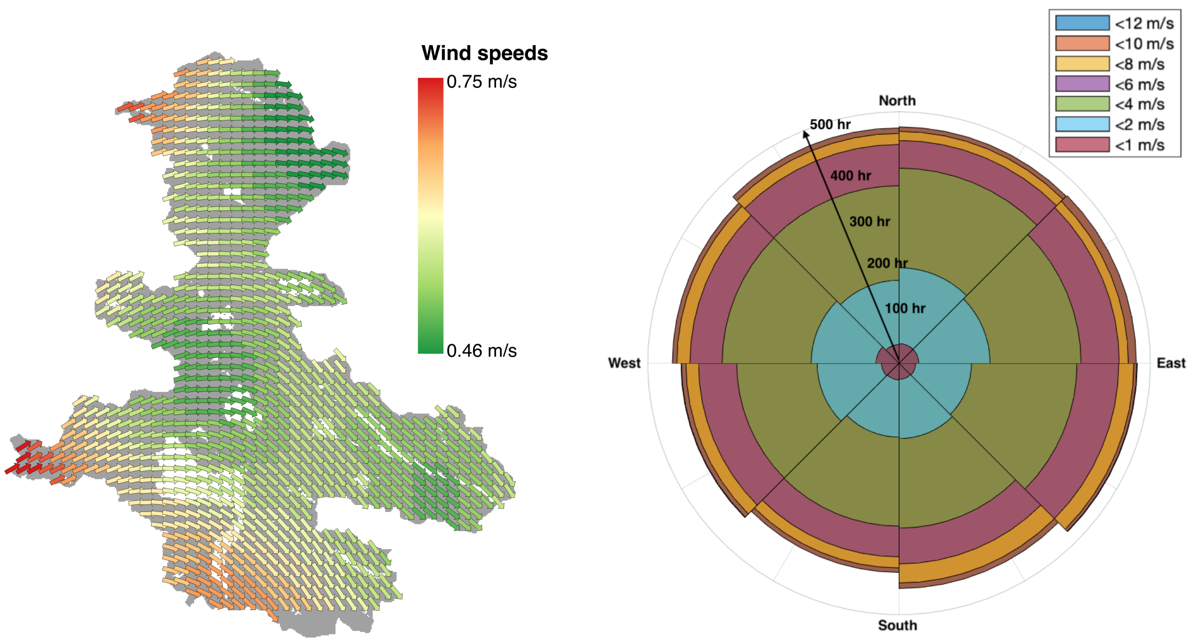


Figure 4.6: Relationship between average SWE and aspect based on ground measurements from 2018.

The wind speeds during 2019 were slightly higher with a mean wind speed of 3.0m/s. However, the wind direction was more evenly distributed among all cardinal directions, see figure 4.7b, yielding a wind direction eastward over the majority of the catchment and with significantly lower net wind velocities, see figure 4.7a.



(a) Seasonal net wind velocity and direction.

(b) Polar histogram of seasonal wind speeds and directions.

Figure 4.7: Wind velocities and directions during the snow season of 2019.

Since the net wind direction of 2019 was eastward, the main snow accumulation should occur in this direction, or in terms of aspect ~90°. The main accumulation occurred at aspects around ~105-135°. However, additional accumulation occurred at ~45°, ~195°, and ~315°, making the results ambiguous and inconclusive.

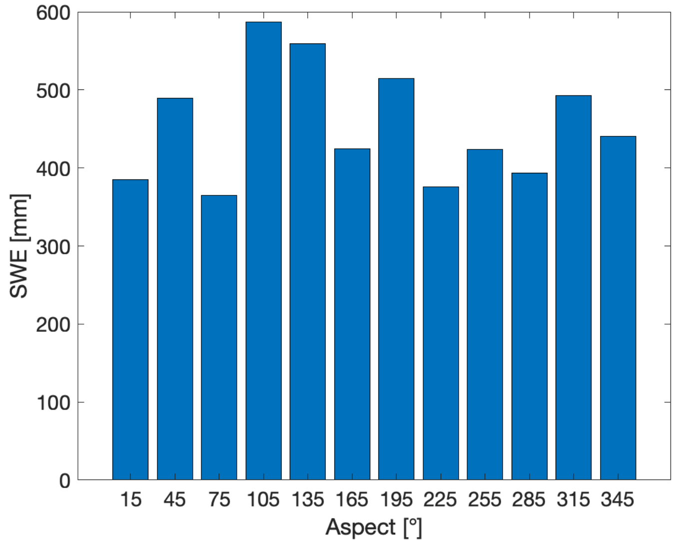
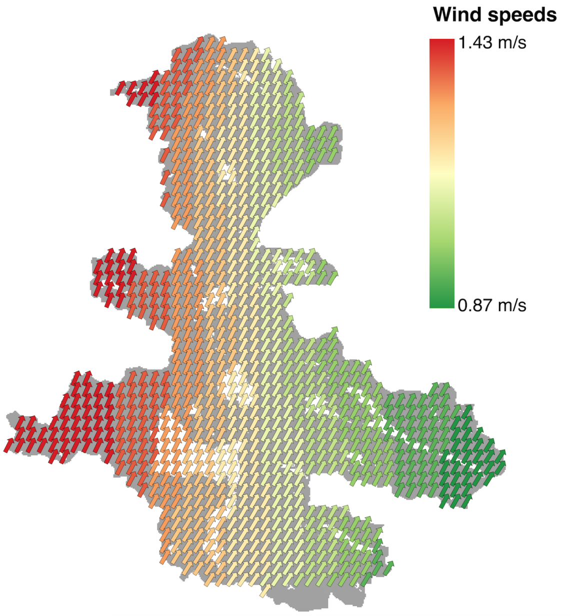
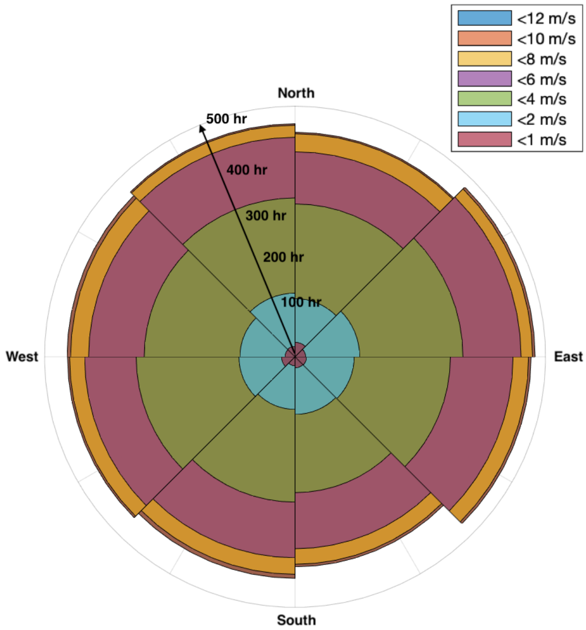


Figure 4.8: Relationship between average SWE and aspect, based on ground measurements from 2019.

During 2020, the catchment experienced stronger winds, with a mean of 3.3m/s, and compared to 2019 the wind direction was not as evenly distributed, see figure 4.9b. The net wind blew in a north-eastward direction and with significantly higher velocities than previous years, as can be seen in figure 4.9a.



(a) Seasonal net wind velocity and direction.



(b) Polar histogram of seasonal wind speeds and directions.

Figure 4.9: Wind velocities and directions during the snow season of 2020.

Winds blowing in a north-east-ward direction should result in snow mainly accumulating on the north-east side of topographical extremes, or in terms of aspect ~45°. Analysis of the ground measurements shows accumulation in aspects of ~45° and ~105°, once again showing accumulation in the direction of the wind but also in a direction dissimilar to the wind direction, see figure 4.10.

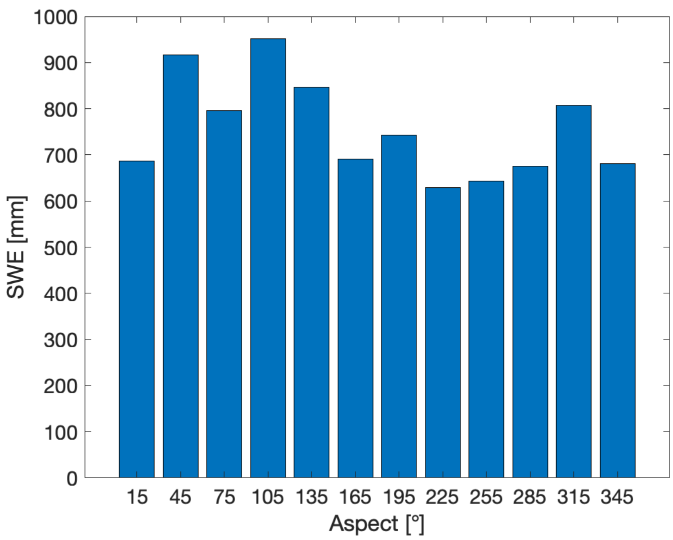


Figure 4.10: Relationship between average SWE and aspect, based on ground measurements from 2020.

All in all, the direction of accumulation of snow shows some correlation with the main wind direction during some years. However, additional accumulation occurs at other directions, making the results inconclusive. As can be seen in polar histogram figures 4.5b, 4.7b and 4.9b, the wind direction was quite evenly distributed, meaning there generally was no distinct dominating wind direction in any of the years. This naturally reduces the likeliness of a consistent snow accumulation, making the results vague. Furthermore, the wind data were obtained from estimations and can not guarantee high accuracy at small space and time scales. Solar loading is another factor that affects the accumulation based on aspect, making the results even more uncertain.

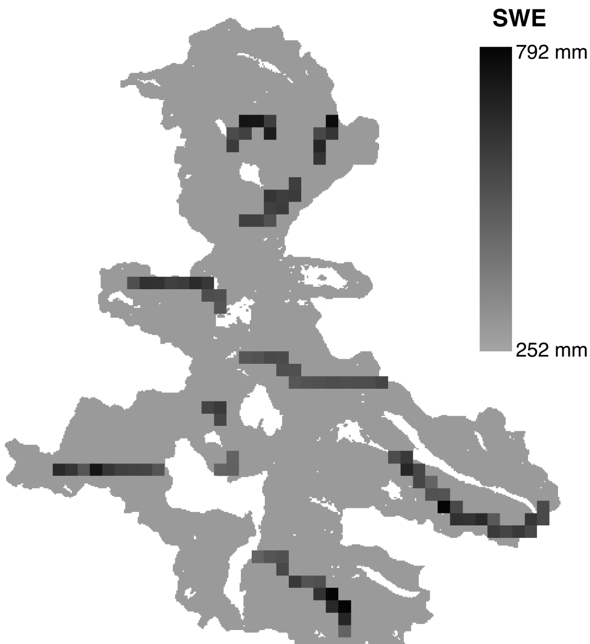
4.3 Evaluation of Copernicus and NVE SWE in the catchment

The following section is divided into three parts in accordance with section 3.3. The first part consists of pixel wise comparison, analyzing the satellite's and model's ability to capture spatial heterogeneity. The second part evaluates the two products capabilities to produce accurate estimations for the aggregated SWEs over the entire catchment.

4.3.1 Evaluation of spatial variations based on ground measurements

The analysis presented in section 4.1 concluded that cell sizes substantially higher than 100m will fail to capture spatial heterogeneity to some extent. Due to both products having higher cell size than 100m, some spatial heterogeneity was bound to be lost. However, to what degree each product was able to capture the spatial heterogeneity was important for evaluating their potential for SWE estimation at this scale.

Copernicus' coarse resolution and the snow lines close proximity to each other causes the individual lines to intersect. Since the empty Copernicus cells can not be included in the snow line analysis, Copernicus' product was effectively missing 4 cells that would have been included had the cells not been empty. The NVE grid was fairly coarse, however, the corresponding snow lines are still well separated from one another. The two products' snow line rasters from week 11, 2018, are shown in figure 4.11.

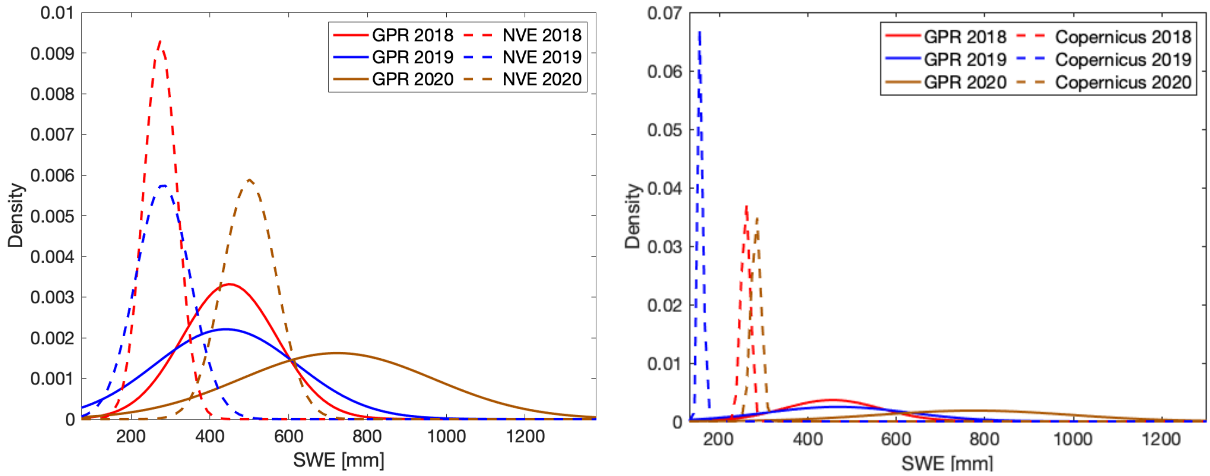


(a) NVE's 94 raster pixels covering the snow lines.

(b) Copernicus 35 raster pixels covering the snow lines.

Figure 4.11: The two products' respective raster data from week 11, 2018, at the position of the ground measured snow lines.

To make the GPR data comparable with each product, the data were upscaled to a grid with a resolution identical to each product's resolution. In total, there were 4 different data sets, the Copernicus data, the NVE data, and two sets of GPR data with 1000m and 5000m cell sizes. To visualize the data, normal distribution curves were fitted to each respective resolution, see figure 4.12. Comparison between the products and the corresponding GPR data illustrates their respective potential to depict the absolute variability. Ideally, each respective data set would exhibit identical behavior to the GPR data, however, due to estimation and measurement errors there are differences between them. The GPR data and NVE data show large similarities. While the variation of the NVE data was significantly lower compared to the GPR data, how the shape and position of NVE's curves related to one another was very similar to how the GPR curves related to each other. This implied that NVE quite accurately captured the interannual variability of SWE within the catchment. Unfortunately, the same can not be said for the Copernicus data. Its standard deviations were vastly lower compared to the GPR data's, and the interannual variability showed substantial discrepancies from the GPR data.



(a) GPR and NVE snow line data.

(b) GPR and Copernicus snow line data.

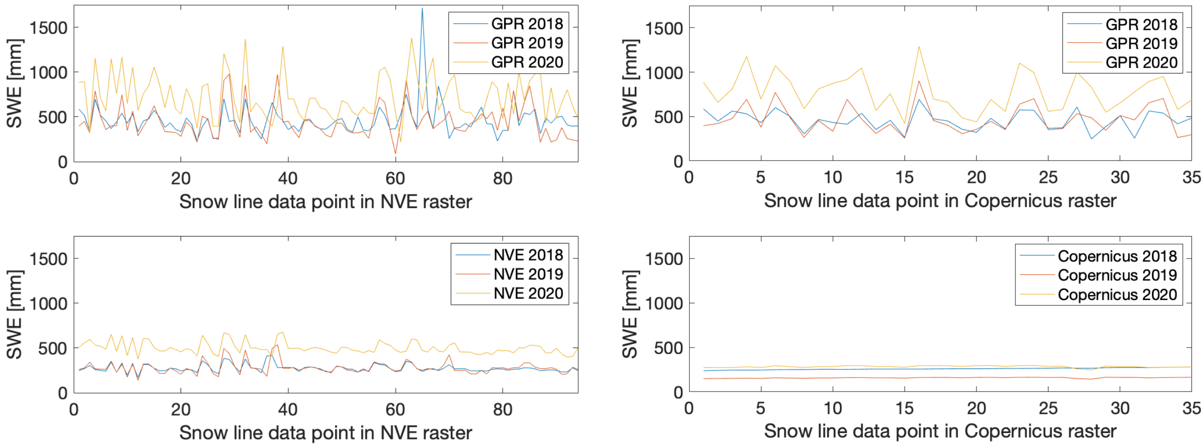
Figure 4.12: Normal distribution curves fitted to accordingly rasterized GPR data and each products snow line raster data. NOTE: the two versions of GPR distribution contain identical ground data, but due to the different rasterizations, the GPR curves have slight differences between the two figures.

Values for each product's yearly standard deviation, indicating their respective variations, is shown in table 4.8, further illustrating the Copernicus product's inability to depict the GPR data's behavior. These values also showed slight differences in the differently rasterized GPR data as a result of the different resolutions.

Table 4.8: Standard deviation σ in mm SWE.

Year	GPR 1km x 1km	NVE	GPR 5km x 5km	Copernicus
2018	120.3	43.0	109.2	10.5
2019	180.6	69.3	160.8	5.2
2020	245.8	67.8	216.6	11.2

Figure 4.13 shows each product's data and the correspondingly upscaled GPR data sets, comparing the SWE values on a pixel-wise level. It is important to note that the two subfigures contain the very same GPR data, and that the different number of data points and behaviour of the curves is a result of the two different resolutions that the data has been rasterized to. Despite each product not being able to fully replicate the spatial variations of the GPR data, it was evident that NVE captures it to a significantly higher extent than Copernicus. Furthermore, NVE not only shows spatial variations, but shows a clear similarity to the behaviour of the GPR data's variations. For example, in figure 4.13a the GPR data show 4 clear spikes in the SWE value between data points 20 and 40, which is something that the NVE data clearly replicate, although with lower magnitudes.



(a) Pixel wise SWE values from NVE and ground measurements.

(b) Pixel wise SWE values from Copernicus and ground measurements.

Figure 4.13: NVE and Copernicus pixel wise spatial heterogeneity over the snow lines, as compared to the correspondingly rasterized GPR data.

A correlation analysis was conducted as a representation of each product's ability to replicate the behavior of the GPR data, see table 4.9. Due to the NVE product being able to reflect the spatial variability, the correlations were high. This indicates that the NVE product is a capable tool for estimating the spatial variability, even in smaller areas. The corresponding data from the Copernicus product did not correlate with the GPR data, proving its incapability in depicting the spatial heterogeneity.

Table 4.9: Correlation coefficients between the GPR data and the respective SWE products.

Year	Correlation NVE	Correlation Copernicus
2018	0.1777	-0.1710
2019	0.6965	0.0730
2020	0.6667	0.0700

4.3.2 Evaluation of aggregated performance based on ground measurements

Regardless of the products' ability or inability to capture the snow's spatial heterogeneity, it was also important to know whether they could perform well on a catchment scale. This was done by comparing the catchment-wide average SWE values from each of the products to that of the GPR data. The snow lines are positioned to give an accurate description of the snow conditions of the entire catchment's terrain. Therefore, the mean value of the GPR SWE data gives a precise estimate for the aggregate catchment-wide SWE. When looking at the products' data, however, the rasters not only had to be temporally extracted at the same time as the GPR data, but they also had to be cut to fit the catchment shape as precisely as possible and to remove the larger lakes. The results of the three average SWE values are plotted in figure 4.14.

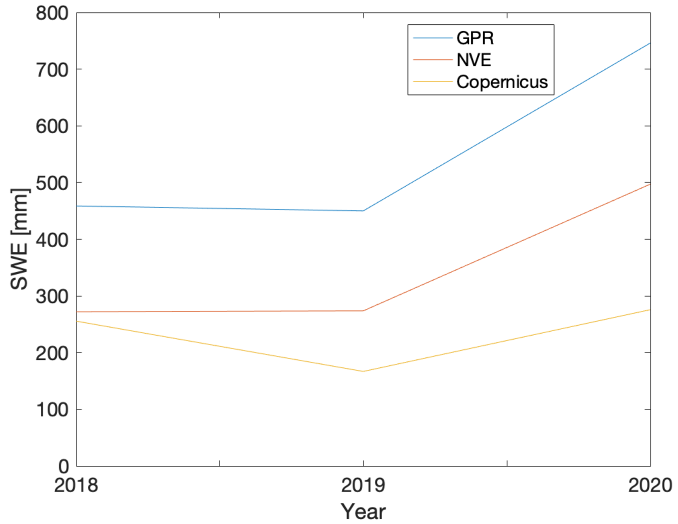


Figure 4.14: The three catchment-wide average SWE values for week 11, 2018 to 2020.

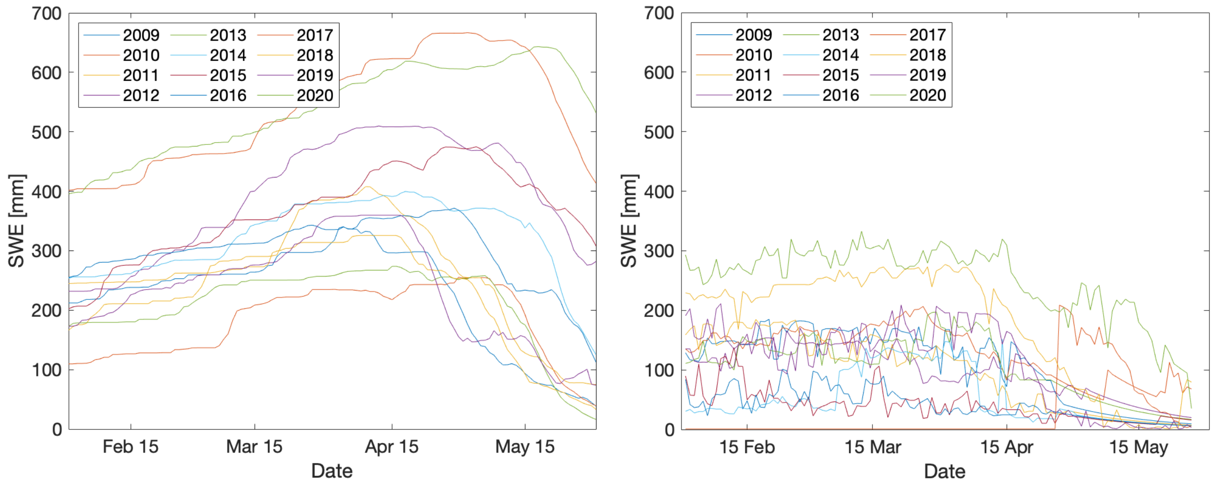
Looking at the absolute catchment-wide average SWE values, the two products, especially the NVE model, performed well compared to the pixel-wise comparisons. Table 4.10 shows by how much each product underestimated the total amount of SWE. In this case, the ratios showed that the NVE data underestimated the in situ data by 33-41%. Since the ratios were similar between the years, these data could be used in estimating the aggregated snow water volumes in this catchment. Copernicus' data, on the other hand, displayed a large difference in the ratio from one year to the next.

Table 4.10: Ratios describing how each product relates to the GPR data in absolute values.

Year	Ratio NVE/GPR	Ratio Copernicus/GPR
2018	0.5887	0.5572
2019	0.6049	0.3708
2020	0.6633	0.3696

4.3.3 Evaluation of aggregated performance based on snowmelt

The accuracy of modeled and satellite-based SWE products were also evaluated based on snowmelt volumes, which was derived from runoff and precipitation values obtained from Uniper. These volumes were based on snowmelt from the entire catchment, meaning that the snowmelt volumes over the lakes were included. Unfortunately, the Copernicus data lacked measurements over and close to the large lake at the bottom of the catchment, which would to some extent impact the results. However, seeing as Copernicus showed such a homogeneous behaviour, this impact should not be significant. The analysis covers 12 years and compares the seasonal snowmelt with the corresponding aggregated average SWE value from the two products. To obtain the aggregated SWE value from the snow products, the maximum average SWE value of the season was acquired. The aggregated average SWE over the entire catchment was therefore extracted for each day and for each full year. Since the maximum amount of SWE consistently occurred during spring, the aggregated values of February 1st to May 31st were visualized in figure 4.15 in order to get a better overview of the aggregated values' behaviour. Ideally the products would yield equal results, e.g., show the same day of maximum SWE each year. However, since the products utilize fundamentally different observational methods and inevitably contain estimation errors, the values differed.



(a) NVE's average daily SWE values. (b) Copernicus' average daily SWE values.

Figure 4.15: NVE's and Copernicus' average daily SWE values over for the full year and the majority of the snow season respectively, and for each year from 2009 to 2020.

The differences between the two products are evident when looking at figure 4.15. For example, the NVE model reached a significantly higher SWE, up to 3.7 times the maximum of Copernicus. Furthermore, the maximum day of SWE occurs later in the season for the NVE model compared to Copernicus. Table 4.11 shows each product's maximum aggregated SWE value for each year, as well as the date at which this maximum occurs. Included in the table is also each product's ranking of highest to lowest SWE value, and the difference in when the maximum occurs for the two products in number of days.

Table 4.11: The two products' respective day of maximum average SWE and the corresponding SWE values. The 'Rank' column shows how they rank internally amongst the product's values, and the 'Diff' column shows the difference in when the day of maximum occurs between the two products in number of days. The last row shows the average of the above values for each column respectively.

Year	NVE			Copernicus			Diff
	Date	SWE [mm]	Rank	Date	SWE [mm]	Rank	
2009	Apr 4	340.4	9	24 Feb	184.9	6	39
2010	May 6	254.8	12	31 Mar	206.6	4	36
2011	Apr 10	407.6	5	21 Feb	184.7	7	48
2012	Apr 20	509.3	3	28 Feb	165.9	8	51
2013	Apr 15	273.7	11	3 Apr	197.3	5	12
2014	Apr 18	399.7	6	18 Apr	151.9	9	0
2015	May 4	474.9	4	16 Feb	137.4	10	46
2016	Apr 29	371.6	7	27 Feb	101.2	11	61
2017	May 2	666.7	1	1 May*	208.7	-	-
2018	Apr 9	325.6	10	5 Apr	277.1	2	4
2019	Apr 15	359.7	8	13 Feb	211.3	3	61
2020	May 18	643.2	2	17 Mar	332.8	1	62
All	Apr 23	418.9	-	16 Mar	196.6	-	38

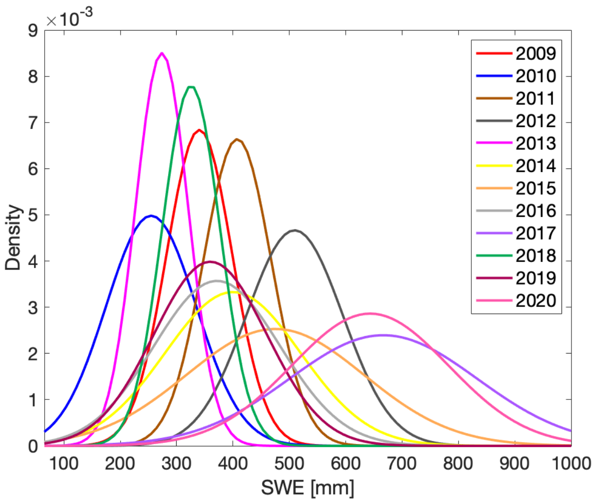
*Due to errors in the data set, this date is likely much later than the actual day of maximum SWE.

As is shown in table 4.11, the two products yielded large differences in when the time of maximum SWE occurred, as well as how the different years ranked. In other words, there was very little that agreed between the two data sets obtained from this analysis.

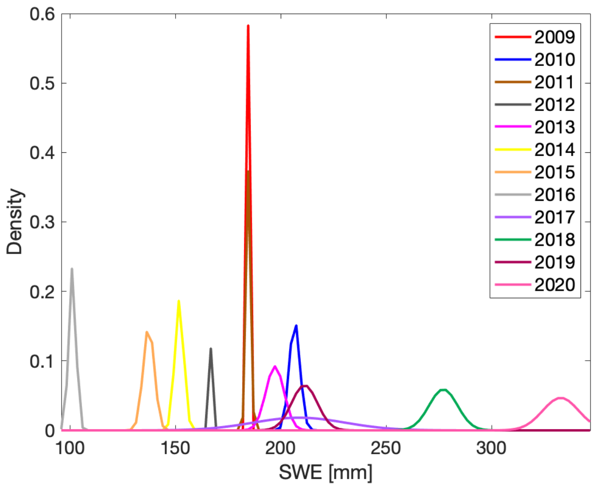
An issue found with Copernicus was that data from a significant share of the snow season of 2017 was missing, which can be seen in figure 4.15b. This lead to 2017 yielding a date and SWE value later and lower, respectively, than the full data set probably would have given. For this reason the Copernicus' 2017 data were removed from the ranking and difference analysis shown in table 4.11. Copernicus' SWE values also decreased and increased before and after its day of maximum SWE on multiple occasions over the years. In some years, NVE also showed a decrease in average SWE prior to the day of the maximum or an increase in SWE after the day of maximum. As mentioned in section 3.3.3, this phenomena meant that the maximum average SWE underestimated the total amounts of SWE corresponding to snowmelt. This premaximum melting or postmaximum addition of snow was not accounted for since it was not possible to determine whether the changes of SWE derives from actual changes in snow conditions or if it stems from estimation errors. Nevertheless, it has to be taken into consideration when comparing the products' values to the snowmelt values.

Normal distribution curves were fitted to the data from Copernicus and NVE for the day of maximum SWE of the season to further visualize the products' respective temporal variations, see figure 4.16. Unfortunately, no direct comparison could be made with GPR, since these data only cover 3 years and the wrong time frame. However, the shape of the bell curves and the standard deviations of the normal distribution curves gave indications on how the products behaved on a temporal scale.

NVE displayed a similar behaviour between all years. The shape of the bell curve remained similar, with slight changes in mean SWE, see figure 4.16a, and standard deviations ranging from 46.9 to 166.6 mm for all 12 years. Due to Copernicus' coarse grid, the variation of SWE over the catchment was low. Furthermore, the interannual variations were significant, resulting in the bell curves hardly intersecting, see figure 4.16b. The standard deviations showed some variations between them, but were overall very low for all years, ranging from 0.25 to 21.89 mm. For reference, the GPR data with the same grid size as NVE showed standard deviations that ranged between 120.3 and 245.8 mm, and the standard deviations of the GPR data with the same grid size as Copernicus ranged between 109.2 and 216.6 mm. This part of the analysis also showed that a year with higher SWE values, both in terms of average and overall, tends to generate higher variations in SWE, which was also a trend in the GPR data.



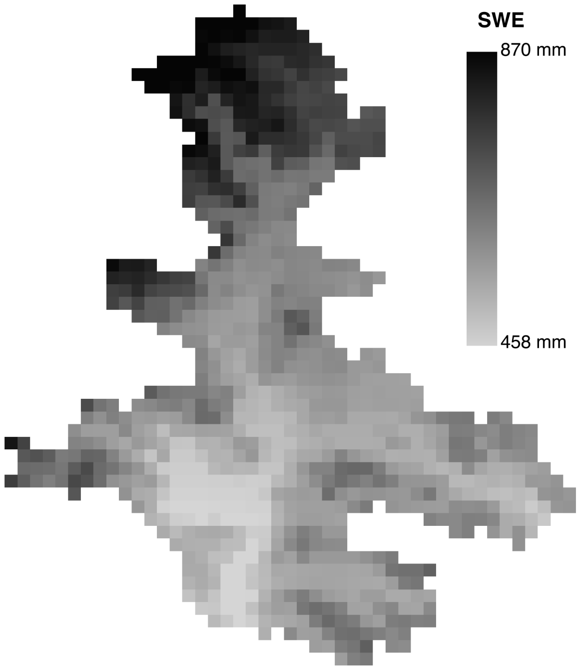
(a) NVE's normal distribution curves.



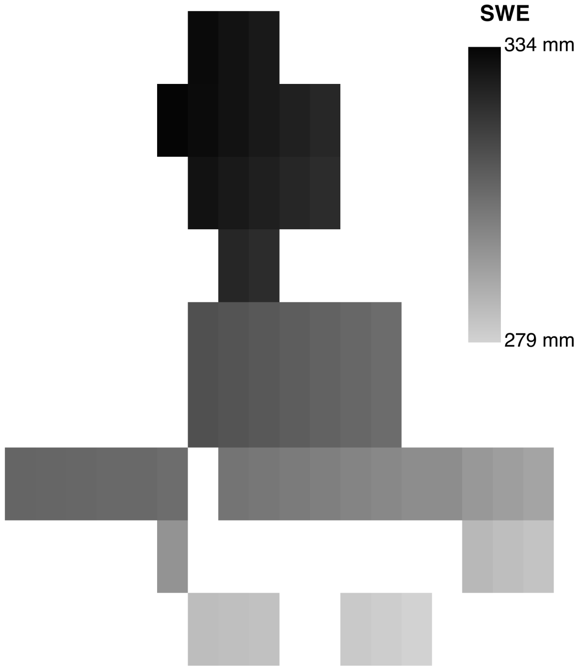
(b) Copernicus' normal distribution curves.

Figure 4.16: Normal distribution functions fitted to the two products' data from their respective yearly day of maximum SWE. NOTE: as the two data sets yield significantly different curves, both in terms of mean values and in distribution, the two figures are plotted independently with different axes.

To further visualize an example of the distribution of the SWE obtained from the two products, figure 4.17 shows the values of NVE and Copernicus on their respective days of maximum average SWE in 2020.



(a) NVE's distribution of SWE on May 18th, 2020.



(b) Copernicus' distribution of SWE on March 17th, 2020.

Figure 4.17: The two products' distribution of SWE over the catchment on their respective day of maximum average amount of SWE in 2020.

Figure 4.18 shows the yearly aggregated maximum SWE values for all products, and as previously stated, this also includes the volume-to-SWE translation of the values obtained from the snowmelt estimations. Both the aggregated values and the inter-annual trends between NVE and snowmelt closely resembled one another, proving that the NVE model estimated the total snowmelt volume well. The Copernicus product severely underestimated the SWE in the catchment during all years except 2010, but showed a behaviour that indicated that it is more accurate during years of lower amounts of snow. Furthermore, its trend showed very little similarity to the snowmelt values.

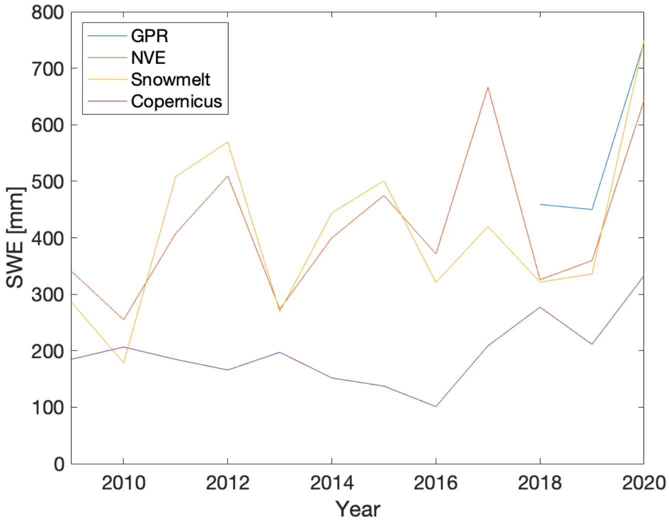


Figure 4.18: All products average SWE values from the day of maximum aggregated average SWE within the catchment, but with the GPR data from w.11.

In table 4.12, the ratios showing how the two products and the GPR data related to snowmelt are summarized. The average ratio between snowmelt and NVE was 1.0724, meaning that the NVE product on average estimated the snowmelt accurately. The average ratio between Copernicus and snowmelt was substantially lower at 0.5464, clearly showing this product's inability to estimate the aggregated SWE values in this application. However, caution should be taken when comparing these numbers since the snowmelt values were estimated, and on average underestimates the SWE over the catchment by 20% compared to the GPR measurements. Additionally, the GPR measurements were taken during a period of estimated maximum SWE and not at the actual time of maximum, meaning that they do not represent SWE values that correspond to a seasonal snowmelt. Nevertheless, the interannual trends of snowmelt gave valuable information regarding the general behavior of the snowfall over the catchment between the years.

Table 4.12: Ratios describing how the two products and the ground measurements relate to the snowmelt.

Year	Ratio NVE	Ratio Copernicus	Ratio GPR
2009	1.1915	0.6473	-
2010	1.4279	1.1580	-
2011	0.8019	0.3633	-
2012	0.8944	0.2914	-
2013	1.0162	0.7324	-
2014	0.9005	0.3422	-
2015	0.9481	0.2743	-
2016	1.1576	0.3152	-
2017	1.5880	0.4972	-
2018	1.0128	0.8619	1.4276
2019	1.0710	0.6289	1.3397
2020	0.8585	0.4442	0.9956
All	1.0724	0.5464	1.2543

5

Conclusions

The purpose of this thesis was on one hand to evaluate the possibility of modelling snow distribution using topographical features and forest cover as forcing inputs, and on the other hand to evaluate the catchment specific performance of two different SWE estimating products. This work was based on data obtained from a Swedish catchment area called Stora Blåsjön, an area in the mountainous region close to the Norwegian border and with a size of approximately 940km^2.

All investigated features, being elevation, slope, aspect, roughness, and forest canopy cover, show correlations with SWE over the snow lines, with elevation and forest having the highest. Combining every topographical feature to form a multiple linear regression model proved capable to estimate the spatial distribution of SWE to some extent. The regression models for individual years were able to explain 21% to 51% of the GPR data's variability, while the generalized model could explain 36% of the variability of the complete data set. The correlations for the individual regression models had correlations ranging from 46% to 72%, while the generalized model's correlations ranged from 45% to 71% between the three years. The small differences in performance between the individual models and the generalized model stem from the low interannual variability of the snow distribution. In conclusion, the models show high correlation with the GPR data. Despite the variability being low, it is evident that topography has a significant impact on SWE distribution and is a useful tool for predicting the SWE over the catchment. The lack of a consistent dominating wind direction during the snow seasons reduced the effect of wind on snow distribution. This resulted in the accumulation of snow showing sporadic correlations with wind over the catchment and the results remain inconclusive.

The satellite-based observational method, Copernicus, showed a complete inability to describe spatial distribution of SWE over the catchment. Furthermore, the aggregated SWE value for the entire catchment disagreed with both the GPR and snowmelt data. In order for a satellite-based snow product to be useful in this form of small-scale application, the resolution needs to be increased substantially and the pixel-wise accuracy needs to improve in order to capture spatial variability to any extent. The accuracy of the observations also need to improve in order to provide accurate estimates for the aggregated SWE values.

When summarizing the results produced by the NVE product to the GPR data, it becomes clear that it has substantial potential even at geographical extents as small as the one involved in this project. It shows high potential in replicating the values obtained from the GPR runs, both on an aggregate level and on a pixel-wise scale. Since the snowmelt volumes that were used as a reference are uncertain they can not be used as absolute indicators, but the fact that the two agree to such a high extent only strengthens the previous results of the model. However, further evaluation with additional in situ measurements is necessary to conclude if the GPR data can be replaced by the modeled data. The NVE model indicates that the maximum SWE occurs later in the season than when the ground measurement campaigns were conducted, evaluating the date when the measurements are taken could therefore prove to be beneficial to increase the accuracy.

Since the snow and weather data-based model used in this project showed such large potential in estimating the area's snow conditions, and since it has not been designed with the Swedish land area in its focus, it would be desirable to see how well a model similar to NVE's could perform if designed with the purpose of describing the snow conditions within the Swedish catchment areas. For example, it could be of interest to see how successful NVE's Swedish counterpart, SMHI, would be in producing such a model. Naturally, constructing such a model would require many hours of work, not to mention the amount of in situ input data, but it could be something to consider moving forward with.

Bibliography

[1] National Snow & Ice Data Center, *Remote Sensing: Passive Microwave*, 2020. [Online]. Available: https://nsidc.org/cryosphere/seaice/study/passive_remote_sensing.html (visited on 03/10/2021).

[2] Copernicus Global Land Services, *Cryosphere*, 2021. [Online]. Available: https://land.copernicus.eu/global/themes/cryosphere (visited on 05/10/2021).

[3] J. Muñoz Sabater, *ERA5-Land hourly data from 1981 to present*, 2019. [Online]. Available: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview (visited on 05/20/2021).

[4] D. L. Kane, "Snowmelt Infiltration into Seasonally Frozen Soils," 1980. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0165232X80900208.

[5] SMHI, *HBV*, 2009. [Online]. Available: https://www.smhi.se/forskning/forskningsenheter/hydrologisk-forskning/hbv-1.1566 (visited on 04/15/2021).

[6] T. M. Saloranta, "Simulating snow maps for Norway: description and statistical evaluation of the seNorge snow model," 2012. [Online]. Available: https://tc.copernicus.org/articles/6/1323/2012/tc-6-1323-2012.pdf.

[7] K. Luojus, J. Pulliainen, M. Moisander, M. Takala, J. Cohen, and J. Lemmetyinen, "Copernicus Global Land Operations," 2017. [Online]. Available: https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS2_PUM_SWE-NH-5km-V1_I1.02.pdf.

[8] F. G. Watson, T. N. Anderson, W. B. Newman, S. E. Alexander, and R. A. Garrott, "Optimal sampling schemes for estimating mean snow water equivalents in stratified heterogeneous landscapes," *Journal of Hydrology*, vol. 328, no. 3, pp. 432–452, 2006, The ICWRER - Symposium in Dresden, Germany, ISSN: 0022-1694. DOI: https://doi.org/10.1016/j.jhydrol.2005.12.032. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022169406000138.

[9] C. A. Hiemstra, G. E. Liston, and W. A. Reiners, "Snow Redistribution by Wind and Interactions with Vegetation at Upper Treeline in the Medicine Bow Mountains, Wyoming, U.S.A.," 2002. [Online]. Available: https://www.jstor.org/stable/1552483.

[10] B. T. Anderson, J. P. McNamara, H.-P. Marshall, and A. N. Flores, "Insights into the physical processes controlling correlations between snow distribution and terrain properties," *Water Resources Research*, vol. 50, no. 6, pp. 4545–4563, 2014. DOI: https://doi.org/10.1002/2013WR013714. eprint: https://

agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013WR013714. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013WR013714.